Toddlers' inference of people's desires for objects: The effect of gender-stereotype knowledge

Julie A. Eichstedt

A Thesis

in

The Department of Psychology

Presented in Partial Fulfilment of the Requirements

for the Degree of Doctor of Philosophy at

Concordia University

Montreal, Quebec, Canada

May, 2002

ABSTRACT

Toddlers' inference of people's desires for objects:
The effect of gender-stereotype knowledge

Julie A. Eichstedt, Ph.D.
Concordia University, 2002

Children as young as 18 months of age are capable of reasoning about other people's desires for objects and food items. During the same age range, children acquire concepts of male and female, including knowledge of the activities, toys, and other objects associated with each gender. The present study assessed whether children's knowledge of gender stereotypes plays a role in their inferences about men's and women's desires for objects. Using an object request task, 20- and 24-month-old children were shown a series of object pairs, each consisting of one feminine- and one masculine-stereotyped item. A male or female experimenter expressed a desire for one of the two items, as revealed by a happy vs. disgust facial expression, then requested that the child give him or her one item. Half of the children observed the experimenter expressing preferences for the gender-appropriate objects, and the other half, preferences for the gender-inappropriate objects. A stereotyping task was also included to ensure that the objects were gender stereotyped by the children. Children were hypothesized to be more likely to select the desired item to hand to the experimenter if it was gender consistent than if it was gender inconsistent, beginning at 24 months. This effect was hypothesized to be stronger for girls than boys. Children were also administered a task to assess their understanding of other people's desires for food items, in addition to various gender-concept and gender-typing measures.

Results suggested that some children show a rudimentary ability to reason about others' desires for gender-typed objects in the second year of life. Specifically, boys who demonstrated adequate gender-category knowledge and a basic understanding of others'

desires for food successfully offered the experimenter the desired item on the majority of trials. Although these boys also displayed significant gender-stereotype knowledge on the stereotyping trials, they based their desire inferences on the experimenter's affective cues, rather than on the gender stereotyping of the objects. Girls with similar gender and food desire knowledge, however, did not consistently offer the experimenter the desired items. Rather, they offered the experimenter cross-gender-typed items, regardless of which objects were desired.

Acknowledgements

I would like to express my sincere thanks and appreciation to my supervisor, Dr. Lisa Serbin, and co-supervisor, Dr. Diane Poulin-Dubois. Their continuous support, encouragement, and guidance have been invaluable to the completion of this research, as well as to my development as a researcher and psychologist. I would also like to thank my committee members, Dr. Nina Howe, Dr. Tannis Arbuckle-Maag, Dr. Kang Lee, and Dr. Ellen Jacobs for their helpful suggestions and comments on this project. Many thanks also go to my fellow experimenters, Shawn Gates and Alex Frenette, for their tireless work and many, many weekends devoted to this project, as well as to Martine Habra, Irfan Yaqub, Maya Sen, Natacha De Genna, Naomi Grunzeweig, Valentina Munoz, Michelyne Verville, Johanna Vynke, and Sarah Frenkiel. I would also like to extend my thanks to the "child scientists" and their parents who participated in the study.

Additionally, I would like to thank the many people who provided support and assistance to me during this graduate school journey. Special thanks go to the staff and students in the Department of Psychology at Concordia University, the members of the Early Social Development Laboratory, members of the Poulin-Dubois lab, my classmates, and friends. You have been like a second family to me and the times we spent together will be cherished always. I would like to thank my family for keeping the porch light burning for all of these years, and for their never-ending support and encouragement. Without you, I would not be where I am today. Finally, I would like to thank my husband Jason, who deserves a Ph.D. in his own right, for all of the essays and theses he has had to proofread, all of the psychology lectures I have dragged him to, and for the million other ways he has lent his full support to my every endeavour.

Table of Contents

List of Figures

List of Tables

List of Appendices

Toddlers' inference of people's desires for objects: The effect of gender-stereotype knowledge

From the moment of birth, children enter into a complex social world. They are surrounded by family members and other individuals in their environments. They have to learn the "dance" of social interactions, and begin the process of learning the complex rules that govern social behaviour. Due to their dependence on others for survival, it makes sense that infants are biologically predisposed to learn about social beings. In the first few weeks of life, infants preferentially attend to face-like stimuli (Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis, & Morton, 1991), they are able to differentiate their mother's face from a stranger's face (Bushnell, Sai, & Mullin, 1989), and prefer their mother's voice over that of another female adult (DeCasper & Fifer, 1980; Walker-Andrews, 1997). Given their selective attention to social stimuli and the richness of their social environments, it is reasonable to assume that children may also learn to understand and predict others' behaviour at an early age, as well as abstract rules and regularities from their observations of others.

Children's Theory of Mind

In the last 15 years, a great deal of research has been conducted on the development of a theory of mind in childhood (Astington, Harris, & Olson, 1988; Frye & Moore, 1991; Wellman, 1990). Theory of mind refers to the attribution of mental states such as desires, thoughts, beliefs, and intentions to oneself and others, and the ability to use these mental states to both predict and explain people's behaviour. Adults' everyday understanding of human behaviour relies largely on mental states as an explanatory framework. If Peter goes to the ice cream parlour, for example, we assume he did so because he *desired* ice cream and *believed* he could buy some at the ice cream parlour. This naive or folk psychology has been described as a belief-desire psychology as beliefs

and desires are jointly used to predict and explain behaviour.

Research on theory of mind development suggests that children possess a similar, though less elaborate, belief-desire psychology by the time they reach 4 or 5 years of age (see meta-analysis by Wellman, Cross, & Watson, 2001). Specifically, children in this age range display a fairly sophisticated understanding of such mental states as desires, beliefs, and intentions and their relationship to emotions and actions. For example, they understand that if a person believes a particular outcome will occur and it does not, the person will experience surprise (Wellman & Bartsch, 1988). They also understand that information about a person's desires and beliefs is necessary to predict and explain his or her actions, and that the same desire can lead to different actions depending on the person's beliefs (Wellman & Bartsch, 1988). Children in this age range further understand the representational nature of mental states. They are able to understand that mental states such as belief may not correspond with external reality, as shown by their understanding of false beliefs. In the standard false-belief task, the child observes the story protagonist place an object in one of two locations, then leave the scene. In his absence, a second character unexpectedly transfers the object from its original location to a second location. In the critical test question, the child is asked to predict where the story protagonist will look for the object upon his return. Numerous studies have shown that the majority of children above the age of 4 will correctly predict that the protagonist will look for the object in its original location, rather than in its new location, correctly attributing a false belief to the protagonist (Clements & Perner, 1994; Hogrefe, Wimmer, & Perner, 1986; Wellman et al., 2001). In contrast, younger children typically predict that the protagonist will look for the object in its present location, thus failing to take into account the protagonist's false belief.

Although 3-year-olds have also been shown to have an understanding of belief (e.g., Wellman & Bartsch, 1988), they continue to reason about others' actions largely

with respect to desires, with belief playing a more minor role in their explanations. Thus, children in this age range have been termed to have a desire-belief psychology (Bartsch & Wellman, 1995; Wellman, 1993), to emphasize the primacy of desires in their causal explanations of human action. During this stage of theory of mind development, children's understanding of mental states also undergoes transition. Initially, mental states are construed as reflecting objective reality; there is no concept of mental states as representations of reality, subject to inaccuracies or misperceptions. During the fourth year of life, however, some children begin to acquire a representational understanding of mind, with some 3-year-olds showing an early understanding of false belief when they are given the motivation to consider another person's belief, or task demands are simplified (Wellman et al., 2001).

Children's Understanding of Desires

Recently researchers have begun to examine the origins of theory of mind in infancy and toddlerhood (Flavell, 1999; Meltzoff, Gopnik, & Repacholi, 1999; Wellman, 1993). This research suggests that in contrast to the belief-desire (or desire-belief) psychology demonstrated by older children, younger children's reasoning about mental states seems to rely primarily on the concept of desires, with no complementary understanding of beliefs (Wellman & Woolley, 1990). Hence, their understanding of mind has been termed a "simple desire psychology". Studies have shown that by the third year of life, children talk about desires in ways that make reference to both their own and others' internal mental states (Bartsch & Wellman, 1995; Bretherton & Beeghly, 1982; Bretherton, McNew, & Beeghly-Smith, 1981; Brown & Dunn, 1991). They also seem to understand the relation between desire and action and show an awareness of the emotional consequences of fulfilled and unfulfilled desires (Bartsch & Wellman, 1995; Wellman & Woolley, 1990). For example, when given information about a story character's desires, children as young as 2½ years of age were able to accurately predict that the character

would continue searching for the desired object until it was attained and would cease searching for the desired object once it was located (Wellman & Woolley, 1990). The children also correctly predicted that the character would be happy if the desired object was found and sad if the desired object was not found. Other studies have similarly documented an understanding of desire-dependent actions and emotions in 3-year-olds using story-based paradigms (e.g., Hadwin & Perner, 1991; Stein & Levine, 1989; Wellman & Banerjee, 1991; Wellman & Bartsch, 1988; Yuill, 1984). Empirical evidence for an understanding of desires in younger children is, however, scarce.

Tilden and Poulin-Dubois (Poulin-Dubois, 1999; Tilden, 2000) were one of the first to examine desire understanding in infancy. Using an adaptation of the preferential looking paradigm, 18- and 24-month-old children were shown films on a television monitor in which an actress expressed a desire for one of two objects by pointing to the object and vocalizing, "I want that one!". The children were then presented with two still-frame test photos, one depicting the actress holding the desired object and the other depicting the actress holding the undesired object. Visual fixation times to the test photos indicated that children at both 18 and 24 months of age looked longer at the action that was inconsistent or incongruent with the actress' previously expressed desire (i.e., the picture of the actress holding the undesired object). In total, 65% and 62% of 18- and 24-month-olds, respectively, showed a preference for the incongruent picture. Thus, these results suggest that children as young as 18 months of age are able to understand that people act in order to fulfill their desires.

In the above study, the children were provided with visual (i.e., direction of gaze), gestural, and verbal cues about the actress' desires. In a follow-up study conducted with a separate group of 18- and 24-month-old children, gaze and gesture cues were controlled by having the actress also look at, and extend her arm toward the undesired object. Thus, equal attention was given to both the desired and undesired objects in this task, forcing

the children to rely on verbal cues to infer the actress' desires. The results indicated no difference in looking times to the congruent versus incongruent pictures, suggesting that the children were not surprised to see the actress reaching for either of the objects in the test phase. This suggests that attentional cues (i.e., gestures or gaze direction) may be more powerful cues to others' desires than verbal cues for children in this age range. However, it is possible that other modifications to the design (i.e., the location of the two objects was switched between the information and test phase) may also have contributed to the lack of significant findings.

Finally, these researchers also demonstrated that children as young as 18 months of age understand the link between desires and emotions. In a third study, the same children who participated in the second desire-action study were shown videotaped scenarios showing actors expressing desires for objects in the information phase, and different emotional outcomes in the test phase. In the information phase, an actor expressed a desire for one of two objects by looking and pointing at the object while saying, "I want that one! I want that one!". A second actor then replied, "I will give it to you", holding his hand in between the two objects. In the test phase, two photos appeared showing the first actor holding the desired object (positive outcome) on one screen, and the undesired object (negative outcome) on the other screen. The actor displayed a happy facial expression on both screens for two trials and a sad facial expression for the other two trials. When the actor displayed a happy facial expression, the results showed that children in both age groups looked longer at the screen showing the positive outcome (i.e., the actor holding the desired object). In contrast, when the actor displayed a sad facial expression, children in both age groups were found to look longer at the screen showing the negative outcome (i.e., the actor holding the undesired object). Thus, these results suggest that children as young as 18 months of age understand that fulfilled desires lead to happiness, and unfulfilled desires, to sadness.

Recently, Repacholi and Gopnik (1997) examined desire reasoning in 14- and 18-month-old infants, using a food request procedure. In their study, infants were presented with either cheese-flavoured crackers or broccoli flowerets and were allowed to sample each to determine their initial food preferences. After this baseline period, the experimenter tasted each food and produced an affective response (either positive or negative) to indicate her preference, then requested that the child give her some more food. The results indicated that 18-month-old infants, but not 14-month-olds, could reason about desires. Specifically, the 14-month-olds were found to respond egocentrically, offering the food item they themselves preferred (the crackers), regardless of the experimenter's previous emotional display. The 18-month-olds, in contrast, were able to correctly infer that the experimenter wanted the food she had expressed a preference for, even when her desires differed from their own. This suggests that some understanding of people's desires as subjective mental states emerges around 18 months of age.

McKoy and Poulin-Dubois (1999) later attempted to replicate and extend Repacholi and Gopnik's (1997) findings. Eighteen-month-old infants were tested using a variation of the food request task in which either the same experimenter expressed a desire for, and requested, some food (same agent condition), or one experimenter expressed a desire for the food and a second experimenter requested the food (different agent condition). Results indicated that the infants behaved differently in the two conditions. Specifically, the majority of infants were found to give the experimenter the desired food in the same-agent condition, even when the expressed desire differed from their own preference (66.7% and 60.9% of children gave the desired food, when crackers and broccoli were desired, respectively), consistent with Repacholi and Gopnik's findings. Of interest, in the different agent condition, infants did not extend the desires expressed by the first experimenter to the second experimenter. Rather, the majority of the infants

gave the second experimenter the food they themselves preferred, the crackers, in both preference conditions (i.e., when the first experimenter expressed a preference for crackers or broccoli). Furthermore, the latency to give food was significantly longer in the different agent condition, suggesting that the infants did not readily infer the desire of the requester from the affective cues provided by the other experimenter. Although the proportion of children offering the requester the desired food in the same agent condition fell short of statistical significance, likely due to a lack of power, these results suggest some understanding of the subjectivity of desires (i.e., that desires are not transferrable from one person to another) in children as young as 18 months of age.

Researchers examining young children's understanding of mental states within the natural context of family interactions have also proposed that an understanding of desires develops early in life (e.g., see Dunn, 1988; Dunn, 1991; Dunn & Munn, 1985). For example, during the second year of life, children have been reported to show a practical understanding of their siblings' likes and dislikes (i.e., their desires) and will remove or destroy siblings' cherished possessions to upset or annoy them (Dunn, 1988). Such teasing behaviours may suggest that young children can reason about others' desires and are aware that thwarting one's desires leads to dissatisfaction, anger, and frustration. These teasing behaviours increase in elaboration and frequency during the second year, with 43% of 18-month-olds and 90% of 24-month-olds found to tease their family members (Dunn, 1988). According to Dunn, children in the second and third years of life also demonstrate their understanding of others' desires in prosocial ways, such as offering family members an object they want in an attempt to comfort them or make them happy, voluntarily helping parents accomplish desired tasks (e.g., setting the table, helping with household chores), and engaging in cooperative play with their siblings. Although the above behaviours may indicate an early understanding of others' desires, other interpretations are also possible. Specifically, children in this age range may have learned

social scripts or routines, based on their interactions with family members in the past, that allow them to predict certain outcomes when they engage in particular behaviours. That is, a child who removes a sibling's cherished doll may do so not because of an understanding of the sibling's desire for the doll, but rather because the sibling displayed negative affect when the child removed the doll on previous occasions. Thus, these children may be operating on a simple, behavioural rule, rather than attributing mental states to others. However, as Dunn (1988) argues, the fact that the children displayed the teasing and prosocial behaviours described above in varied contexts lends support to the notion that these children were able to reason about others' desires, and were not merely relying on well-learned social routines.

The above studies suggest that children show an understanding of desires as early as 18 months of age. Moreover, even children in the first year of life have been shown to have an understanding of goal-directed actions. Studies using the habituation paradigm have shown that infants link gaze toward, or grasping, an object with an intent or desire to have that object (Spelke, Phillips, & Woodward, 1995; Woodward, 1998). Thus, an understanding of desires appears to emerge early in life. However, a number of questions still remain unanswered. A central question concerns young children's understanding of the factors influencing one's desires.

Children's Understanding of Desire Formation

Adults' conception of desires encompasses the understanding that desires can be generated by physiological states, such as hunger, thirst, and fatigue. For example, if a person is hungry, they will likely desire food. Desires can also be generated by experiences that affect a person's attitudes (e.g., experiences producing pleasure, pain, or fear). For instance, if a person has an enjoyable experience engaging in a particular activity, they may develop a liking for it and may want to participate in that activity again. Finally, character or personality traits can determine desires, resulting in desires that are

relatively stable across situations and time, and that vary across individuals. For example, an extroverted person may desire and seek opportunities to perform in front of a crowd, whereas an introverted individual may prefer to avoid such activities. Understanding the factors that influence desires can provide children with greater explanatory power, increase their accuracy for predicting future desires, and affords them greater control over their own, or others', desires and behaviour.

Research by Moses, Coon, and Wusinich (2000) suggests that children as young as 3 years of age have a rudimentary understanding that experiences affecting attitudes can generate desires, with this understanding increasing with age from 3 to 5 years. For example, they understand that a boy who had an enjoyable experience playing with a dog would want to play with the dog again more than a boy who had previously been bitten by the dog. An awareness of the link between physiological states and desires was not evident until age 4, however. For example, 3-year-olds had difficulty understanding that a girl who had not eaten since breakfast would be hungrier and would want to eat lunch more than a girl who had just eaten a large meal. The authors hypothesized that the children's difficulty in understanding physiology-based desires may result because of the differences in valence (e.g., a positive experience of eating leads to the absence of desire for food because the individual is satiated), which is not true of attitude-based desires (e.g., a positive experience leads to a desire to engage in the activity again). Of interest, children at all ages tested in this study seemed to have difficulty understanding that different experiences would lead to different desires. That is, although they predicted that a child who had a positive experience with an object would want to play more with the object than a child who had a negative experience, both children were seen as having a desire for the object that merely differed in degree. This suggests that children in this age range do not have a full understanding of the causal links between physiological states or affective experiences and desires.

With respect to character traits influencing desires, research by Yuill and Pearson (1998) suggests that children from the age of 5 have an understanding of traits as organized internal states, based on desires, that cause behaviour. In their study, children between the ages of 4 and 7 were presented with stories describing contrasting pairs of traits (e.g., generous-selfish, honest-dishonest), and then were asked a series of questions about each trait. For example, one story described a girl named Mary who always stopped her brother from playing with her toys and refused to share her candy with anyone. One of the critical test questions asked whether Mary would feel sad or happy about giving the children at her party some of her birthday cake. To predict the character's emotion, the children would need to identify the relevant trait from the story (i.e., selfishness), infer the character's desires based on this trait (i.e., a desire to not share the cake), and understand the link between desires and emotions (i.e., that sharing the cake will lead to sadness). The results indicated that children 5 years of age and older were able to correctly predict the story character's emotions in these situations based on the trait information provided, at a level that significantly exceed chance. Thus, children in this age range, but not younger children, were found to have an understanding of traits as relatively stable states of mind that generate desires.

Although younger children did not seem to understand the link between personality traits and desires in the above study, the traits tested were fairly abstract. It is possible that children would demonstrate an earlier understanding of traits as generating desires if more salient social characteristics were used, such as traits related to one's gender identity.

Children's Gender Schemas

From a very young age, children learn to see the world along gendered lines. People are divided into "mommies" and "daddies" and "boys" and "girls", and different roles, traits, interests, and abilities are ascribed to each of these gender categories based on their observation of males and females in their environments (Martin & Halverson,

1981; O'Brien, 1992; Powlishta, Sen, Serbin, Poulin-Dubois, & Eichstedt, 2001; Reis & Wright, 1982; Ruble & Martin, 1998). According to gender-schema theory (Martin & Halverson, 1981), the resulting network of ideas or beliefs about gender (the "gender schema") influences what children attend to and recall, what activities and interests they engage in, and even the occupations they may later pursue as adults. This organizational aspect of schemas has been supported by research demonstrating that school-age children typically display superior recall for schema-consistent information relative to schema-inconsistent information (e.g., Koblinsky, Cruse, & Sugawara, 1978; Liben & Signorella, 1993; Ruble & Stangor, 1986), although some exceptions have been noted in the literature (see Jennings, 1975; Kropp & Halverson, 1983). Moreover, in some cases, children will even distort gender-inconsistent information to conform with gender stereotypes, misremembering a male nurse to be a female nurse or a male doctor, for example (Cordua, McGraw, & Drabman, 1979; see also Martin & Halverson, 1983). Gender-schema theory also suggests that children will show greater attention to, recall, and knowledge of gender-appropriate than gender-inappropriate activities and behaviours (the "own-sex" schema; Martin & Halverson, 1981). This has been demonstrated in studies reporting that children show greater interest in novel toys or objects that are labelled as being for their own sex over those labelled as being for the opposite sex (Bradbard, Martin, Endsley, & Halverson, 1986; Cobb, Stevens-Long, & Goldstein, 1982; Martin, Eisenbud, & Rose, 1995; Thompson, 1975).

Gender schemas have also been thought to direct children's behaviours and interests. Specifically, children's gender identity (i.e., their ability to identify themselves and others as male or female) and their knowledge of gender stereotypes were initially proposed as the cognitive basis for the development and expression of gender-typed behaviours and interests. However, gender stereotyping is evident quite early in life in children's preferences for toys and activities, with infants preferring to play with toys that are consistent with gender stereotypes as early as 14 to 20 months of age (Fein, Johnson,

Kosson, Stork, & Wasserman, 1975; O'Brien & Huston, 1985). The observation of gender-typed behaviours prior to the age at which knowledge of gender stereotypes is thought to emerge has lead some researchers to speculate that gender knowledge and gender-typed behaviours may initially develop independently (Blakemore, LaRue, & Olejnik, 1979; Perry, White, & Perry, 1984). Other researchers have found little relationship between the amount of children's gender-related knowledge and the degree of gender-typed preferences exhibited (Hort, Leinbach, & Fagot, 1991; Weinraub et al., 1984). However, gender knowledge has been shown to increase school-aged children's desire to engage in gender-consistent behaviours and activities in other studies (see review by Ruble & Martin, 1998).

Thus, schemas function (1) to organize and structure experiences by influencing what is attended to, encoded, and recalled; and (2) to regulate behaviour by guiding children toward gender-appropriate activities and interests. Schemas also have a third function, which is to guide inference and interpretation in situations in which gender-related information is deficient or ambiguous. Several studies have shown that young school-aged children will use their knowledge of gender stereotypes to make inferences about people, of unspecified sex, described as having a masculine- or feminine characteristic. For example, when told about sex-unspecified target children who liked either masculine or feminine toys, 4- to 6-year-old children predicted that the targets would like other toys from the same gender category (Martin, Wood, & Little, 1990). Of interest, the children only made these stereotypic inferences for target children described as having interests like their own, and not for target children with cross-gender-typed interests, perhaps due to greater knowledge of own-sex stereotypes. Similarly, Bauer, Liebl, and Stennes (1998) presented 4½-year-old children with descriptions of sex-unspecified individuals with stereotypically masculine or feminine personality traits or occupations, and found that the children predicted that the target would wear clothing consistent with that stereotype (e.g., a dress versus a suit coat). Thus, these results suggest

that young school-aged children use the gender-related information contained in their developing schemas to guide their inferences in situations where gender-category information is deficient.

Other research has examined the flexibility of children's gender schemas (see review by Ruble & Martin, 1998). This research suggests that gender stereotypes are typically held quite rigidly until children reach approximately 7 or 8 years of age. Prior to this age children are not thought to recognize the individual variability in masculinity and femininity that exists in males and females. With age, however, as their stereotype knowledge increases, children also begin to learn that such norms are relative and somewhat flexible, and become more willing to accept cross-gender behaviour. Sex differences in the flexibility of children's gender schemas have also been noted. Specifically, research has suggested that girls are both more knowledgeable about and, in the preschool years, more flexible in their application of gender stereotypes than boys (Ruble & Martin, 1998).

Although much research has been conducted on the gender schemas of older children, until recently, little research examined the origins of gender schemas in infancy and toddlerhood. Given the importance of gender in the social world and its impact on children's development, it is important to determine how children first learn about gender as well as how this information affects how they perceive and reason about people.

Development of Gender Categories and Stereotypes in Infancy and Toddlerhood

Research has shown that infants form the categories of "male" and "female" in the first two years of life, largely based on such perceptual features as voice pitch, hair length, and clothing style (Leinbach & Fagot, 1993). They can discriminate male and female voices by 7 months of age (Miller, Younger, & Morse, 1982) and male and female faces as early as 9 months of age (Leinbach & Fagot, 1993). Infants begin to associate male and female voices with the corresponding gendered faces in dynamic displays by 6 months of age (Walker-Andrews, Bahrick, Raglioni, & Diaz, 1991) and in static displays by 9

months of age (Poulin-Dubois, Serbin, Kenyon, & Derbyshire, 1994), reflecting the development of intermodal gender categories. Labels for these gender categories, such as "lady" and "man", are acquired by 18 months of age, with some studies showing greater knowledge of gender labels in girls compared to boys (Poulin-Dubois, Serbin, & Derbyshire, 1998; Thompson, 1975; Weinraub et al., 1984). Gender-category knowledge continues to develop in the second year of life. Recent research by Johnston, Bittinger, Smith, and Madole (2001), for example, examined children's sequential touching of various dolls and found a significant increase in children's tendency to categorize male and female dolls according to gender between the ages of 18 and 22 months.

Once these rudimentary gender categories are formed, children begin to develop various associations around them, including knowledge of the activities and objects commonly associated with males and females. For example, by 18 months of age, girls have been shown to associate dolls with girls and vehicles with boys, although similar knowledge of toy stereotypes was not yet evident in the boys, even at 24 months of age (Serbin, Poulin-Dubois, Colburne, Sen, & Eichstedt, 2001). Knowledge of gender-typed activities has been demonstrated in children as young as 2 years of age. Using the preferential-looking paradigm, Serbin, Poulin-Dubois, and Eichstedt (in press) presented 24-month-old children with a series of paired photographs displaying a man and a women engaged in identical masculine, feminine, or gender-neutral activities. The results showed that the children looked significantly longer at the stereotype-incongruent photos of the male actor engaging in feminine activities (e.g., the photo of the man putting on make-up) than would be expected by chance. Other research, using picture sorting or imitation tasks, have similarly found knowledge of gender-typed objects and activities in children as young as 2 years of age, with children found to pair such objects and activities as a suit, a shirt and tie, shaving, and fixing cars with males, and dresses, putting on make-up, cooking, and vacuuming with females (Poulin-Dubois, Serbin, Eichstedt, Sen, & Beissel, 2002; Weinraub et al., 1984). Finally, various interests, traits, and characteristics have

been found to be associated with males and females early in life. Kuhn, Nash, and Brucken (1978), for example, found that 2-year-old children believed that girls like to play with dolls, like to help their mother, talk a lot, never hit, and say, "I need some help." Conversely, they believed that boys like to help their father, play with cars, fight, and say, "I can hit you."

To summarize, by the end of the second year of life, children have already begun to classify people as "boys" and "girls" or "mommies" and "daddies" and also attribute various interests and preferences differentially to the gender categories.

Use of Gender Information in Children's Inferences about Other People's Desires

Children's knowledge about gender emerges around the same time as does various theory of mind concepts. Both types of knowledge help children to understand and predict other people's behaviour, and these sources of information are likely to be used conjunctively. For example, knowledge of the preferences, interests, and traits that characterize males and females could serve as a basis for making relatively consistent, cross-situational predictions about people's desires. As stated previously, past research on children's understanding of desire formation suggests that children do not have an understanding of personality traits as relatively stable states of mind that generate desires until the age of 5 (Yuill & Pearson, 1998). However, given the salience of gender concepts in early childhood, it is possible that young children may use knowledge about masculine and feminine characteristics as a basis for making inferences about other people's desires or preferences. Research on the development of gender schemas has shown that school-aged children will use their knowledge of gender stereotypes to make inferences about the preferences, interests, and characteristics of sex-unspecified individuals (e.g., a person described as having one feminine interest is often predicted to have another feminine interest; Bauer et al., 1998; Martin et al., 1990). Similarly, when given information about a person's gender, research suggests that school-aged children will use this information to make predictions about his or her likely preferences, traits, or

characteristics (Berndt & Heller, 1986; Martin, 1989). In fact, younger school-aged children were found to use a person's sex as a primary basis for making predictions about his or her preferences for gender-typed objects, ignoring other individuating information, such as the person's previous interests (Berndt & Heller, 1986; Martin, 1989). For example, Martin presented children with descriptions of boys and girls with either stereotypic, counter-stereotypic, or neutral interests and asked them to predict the target children's preferences for a series of gender-typed toys. Children between the ages of 3½ and 6 years of age judged that boys would like masculine-typed toys and girls, feminine-typed toys even when the target children were described as previously displaying cross-gender-typed interests (e.g., a girl described as having a boy for a best friend and liking airplanes was still predicted to prefer dolls over cars). Only children aged 8½ and older consistently used both the target children's stated interests and their sex to predict their preferences for gender-typed toys. These studies suggest that gender knowledge is used to make inferences about others' desires in school-aged children, and may override other sources of information about a person's desires, including his or her past interests. Whether gender-based inferences of other people's desires could be demonstrated in even younger children is currently unknown.

A related question concerns the relative weight given to gender information by young children in their determinations of other people's preferences. In the gender studies above, young school-aged children used gender information over information about an individual's past interests to make predictions about their likely preferences for gender-typed toys. However, this may not be the case when they are given more immediate cues about the individual's desires. Various studies have suggested that children are able to make inferences about people's desires based on such physical cues as direction of eye-gaze and gesturing (e.g., see Baron-Cohen & Ring, 1994), as well as affective cues, such as facial expressions with accompanying verbal prompts (e.g., Repacholi & Gopnik, 1997). The way children use these different sources of information about people's desires

in conjunction with gender cues to make social inferences is of importance, as well as the relative weight given to each of these cues. Of particular interest, is whether information about a person's gender will override other immediate cues to his or her desires in children's inferences when these different sources of information conflict. That is, how do children reason about other people's desires when the desires expressed conflict with traditional gender norms?

It is possible that toddlers, like the young school-aged children described in Martin's (1989) study, may rigidly apply gender stereotypes in reasoning about males' and females' desires. Thus, they may assume that all males will display preferences for traditionally masculine objects and all females, preferences for traditionally feminine objects. Alternatively, these children may apply their gender knowledge more flexibly, using their knowledge of gender stereotypes to guide their inferences, but able to consider that individuals may deviate from this stereotypical pattern. The ability to make gender-based inferences about other people's desires is likely to be age-dependent. Although rudimentary knowledge of gender stereotypes has been demonstrated in children as early as 18 months of age (Eichstedt, Serbin, Poulin-Dubois, & Sen, in press; Serbin et al., 2001), gender-typed knowledge is not consistently reported in children until 24 months of age (Poulin-Dubois et al., 2002; Serbin et al., in press; Thompson, 1975; Weinraub et al., 1984). Thus, by 24 months, children are likely to have sufficient gender knowledge to guide their inferences about other people's desires for objects.

Differences between girls and boys in the use of gender information to make desire predictions are also possible. Some research suggests that girls have more advanced gender-stereotyped knowledge than boys (see Ruble & Martin, 1998), showing greater knowledge of gender labels (Poulin-Dubois et al., 1998; Thompson, 1975; Weinraub et al., 1984) and gender stereotypes for toys and activities in infancy and toddlerhood (Poulin-Dubois et al., 2002; Serbin et al., 2001). This suggests that girls may be more likely than boys to make gender-based inferences. However, other studies with children in

the same age range have found either equal gender-stereotype knowledge in both sexes, or superior knowledge in boys relative to girls, depending on the method used and gender concepts tested (Eichstedt et al., in press; Serbin et al., in press; Weinraub et al., 1984). Moreover, studies examining the flexibility of gender stereotypes in school-age children have generally found that girls are more flexible in their gender-typed knowledge than boys (see Ruble & Martin, 1998). This may suggest that girls may be more willing to consider individuating information to determine others' desires, rather than relying exclusively on gender information, compared to boys. However, the flexibility of younger girls' gender schemas is unknown.

The Present Study

The objectives of the present study were to determine (a) whether children use information about a person's gender to make inferences about his or her specific desires or preferences; (b) the relative weight given to gender information when provided with immediate, conflicting cues to an individual's desire (i.e., how flexibly gender concepts are applied in making desire inferences); (c) the factors underlying the ability to make gender-based inferences (e.g., age, sex, level of desire understanding, level of gender-typed knowledge, etc.). Using an adaptation of Repacholi and Gopnik's (1997) food request task, 20- and 24-month-old children were shown a series of toys or objects, half of which were feminine-stereotyped (e.g., doll, purse), the other half were masculine-stereotyped (e.g., toy car, baseball glove). A male or female experimenter expressed a preference or desire for one of two items, one masculine- and one feminine-stereotyped, as revealed by a happy versus disgust facial expression. The experimenter then requested that the child give him or her one item. There were two experimental conditions: a gender-consistent condition and a gender-inconsistent condition. In the gender-consistent condition, the experimenter expressed a preference for the gender-appropriate object (e.g., the male experimenter expressed a preference for the masculine object). In the gender-inconsistent condition, the experimenter expressed a preference for the gender-

inappropriate object (e.g., the male experimenter expressed a preference for the feminine object). Stereotyping trials, in which the experimenter requested an item without previously expressing a desire for one of the objects, were also included to ensure that the items were gender stereotyped by the children. Knowledge of gender stereotypes would be demonstrated if the children correctly selected the item consistent with the gender of the experimenter.

To examine whether reasoning about desires for gender-typed objects was related to an understanding of desires for other types of items (e.g., food), the children were also given two trials in which the experimenter expressed a preference for either raw broccoli flowerets or cheese-flavoured crackers using a food request procedure (Repacholi & Gopnik, 1997). The children were then asked to give the experimenter some more food. Children's performance on these two tasks (food and object request) were compared to determine whether these abilities were related.

In the food request task, the children were expected to give the experimenter the desired food by 20 months of age, even when the experimenter's preference differed from their own, consistent with Repacholi and Gopnik's (1997) and McKoy and Poulin-Dubois' (1999) findings. In the object request task, the children were also expected to correctly infer the experimenter's desires based on his or her emotional display. However, a developmental pattern was expected in children's responses to the two gender conditions. At 20 months of age, the children were expected to show little knowledge of gender stereotypes and consequently, to be guided primarily by affective cues in their desire inferences. Thus, they were predicted to respond similarly in both the gender-consistent and gender-inconsistent conditions, selecting the desired item to give to the experimenter regardless of the item's gender stereotyping. In contrast, research has consistently suggested that by 24 months of age, children have knowledge of gender stereotypes (Poulin-Dubois et al., 2002; Serbin et al., 2001; Serbin et al., in press; Weinraub et al., 1984). Thus, it was predicted that 24-month-old children may use this gender information

in reasoning about the experimenter's desires. Specifically, the 24-month-old children were hypothesized to be more likely to select the desired object to hand to the experimenter if it was gender consistent than if it was gender inconsistent. Alternatively, the children were predicted to take longer (i.e., have an increased latency) to hand the experimenter a desired item if it was gender inconsistent than if it was gender consistent. As gender stereotypes for males are typically more rigidly held (Archer, 1984; Hort, Fagot, & Leinbach, 1990), the 24-month-old children were also hypothesized to view a male expressing a desire for a feminine-stereotyped object as more inconsistent or incongruent than a female expressing a desire for a masculine-stereotyped object. As a result, children in this age group were expected to be least likely to give the desired object when a male experimenter expressed a preference for a gender-inappropriate, feminine-typed item. Finally, since girls have been shown in some studies to have more advanced gender knowledge than boys (Poulin-Dubois et al., 1998; Poulin-Dubois et al., 2002; Serbin et al., 2001; Signorella, Bigler, Liben, 1993), girls were expected to show greater knowledge of stereotypes in the current study, and to be more likely to make gender-based inferences of others' desires than boys.

The above predictions rest on the assumption that children's knowledge about gender will be used to guide their inferences about the experimenter's desires, when other sources of information about the experimenter's desires are also available. Another possibility is that children may possess gender stereotypes by 24 months of age, but may rely more heavily on the more immediate, affective cues to the experimenter's desires than on his or her gender in making their desire inferences. This would result in children at both ages giving the experimenter the desired objects in the gender-consistent and -inconsistent conditions.

To examine the knowledge underlying the ability to make gender-based desire inferences, various gender-concept and gender-typing measures were administered. Children's gender-category knowledge was assessed using an adaptation of Leinbach and

Fagot's (1986) gender-labelling task. Children were shown a series of paired photographs of children and adults and were asked to identify male and female targets in response to such labels as "mommy" and "daddy", "boy" and "girl". Children's exposure to, and preference for, gender-typed toys and objects was assessed in the object request task baseline trials. In these trials, the child was allowed to play with each masculine-feminine pair for 35 s. The total time spent playing with each item was recorded to determine children's initial preferences. The children's exposure to, and preference for, gender-typed toys and objects was also assessed through parental questionnaire data. Finally, parents' adherence to traditional gender roles was evaluated using the Sex-Typed Child Care and Household Tasks Questionnaire (Orlofsky, 1981; Serbin, Powlishta, & Gulko, 1993). It was predicted that children with greater knowledge of, and exposure to, gender-typed roles, activities, and interests would be more likely to make gender-based inferences of other people's desires, selecting the desired item to give to the experimenter when it was gender consistent, but not when it was gender inconsistent.

To summarize, the major hypotheses of the current study were:

(1) Both age groups were expected to have an understanding of other people's desires. Thus, they were predicted to give the desired food to the experimenter in the food request task, even when the experimenter's expressed preferences differed from their own, replicating with older children the studies by Repacholi and Gopnik (1997), and McKoy and Poulin-Dubois (1999).

(2) The 20-month-old children were expected to have an understanding of others' desires, but limited knowledge of gender stereotypes. Therefore, they were expected to give the desired object to the experimenter in the object request task, irrespective of the gender stereotyping of the item or the sex of the experimenter. On the stereotyping task, both boys and girls were expected to be at chance levels. Although girls have been shown to have gender stereotypes linking dolls with girls, and vehicles and the colour blue with boys by 18 months of age (Eichstedt et al., in press; Serbin et al., 2001), knowledge of the

gender stereotypes for the five remaining objects used in the present study has not been previously demonstrated in children younger than 24 months of age (Picariello, Greenberg, & Pillemer, 1990; Weinraub et al., 1984).

(3) By 24 months of age, children were expected to show higher levels of gender-typed knowledge, as assessed by their performance on the gender-stereotyping task and their knowledge of gender labels. Their performance on both of these tasks was expected to be above chance. On the object request task, these children were expected to use their gender knowledge to guide their desire inferences. Thus, it was predicted that the 24-month-olds would give significantly more desired objects to the experimenters when they were gender consistent than when they were gender inconsistent. Alternatively, the children were predicted to show a longer latency to give the experimenter the desired objects when they were gender inconsistent versus gender consistent. The above effects were expected to be most prominent when the male actor expressed a preference for the objects. It was also predicted that girls may show greater knowledge of gender stereotypes and may be more likely to make gender-based desire inferences than boys.

(4) Children with greater gender knowledge were expected to be more likely to make gender-based desire inferences. That is, the child's level of gender knowledge was predicted to be positively related to the amount of desired, gender-consistent objects given to the experimenters, and negatively related to the amount of desired gender-inconsistent objects given to the experimenters. As children's exposure to gender-typed roles, activities, and interests were also expected to advance their gender knowledge, these measures were also predicted to be positively related to children's tendency to make gender-based desire inferences.

Method

Participants

Sixty-four children participated in the study: 32 20-month-old children and 32 24-month-old children, with equal numbers of boys and girls in each age group. The mean ages of the two groups were 20.06 months (range: 19.47 to 20.57) and 24.12 months (range: 23.47 to 24.57), respectively. These children were recruited from birth lists provided by the Régie Régionale de la Santé et des Services Sociaux de Montréal-Centre, after approval from the Commission d'Accès à l'Information du Québec. The children were from predominately English-speaking, middle-class, intact families. Eighty-six percent of the children were of European-Canadian origin, 3% were of Arabic descent, and 11% were African-Canadian. None of the children had any visual or auditory impairments, as reported by their parents. Most of the children (69%) were cared for at home, typically by their mothers; the remaining children attended daycare.

An additional 24 children (14 20-month-olds, 10 24-month-olds; 17 boys, 7 girls) participated in the study, but were excluded from the final analyses due to noncompliance or fussiness (n = 18), parental interference (n = 1), experimenter error (n = 3), or because the children were unable to come for a second testing session (n = 2). Children were excluded from the final analyses because of noncompliance or fussiness if (a) they did not complete more than 50% of the trials in any of the experimental conditions (i.e., male desire, female desire, male stereotyping, or female stereotyping; n = 15); (b) there were significant delays in responding prior to the child offering the experimenter a food item or object on the majority of trials (i.e., the child's mean latency to respond to the experimenters' requests across all trials was over 3 standard deviations from the group mean; n = 2); or (c) the child was visibly upset during the task and therefore, testing had to be discontinued (n = 1). There were similar numbers of children excluded from the two experimental groups (gender-consistent group, n = 13; gender-inconsistent group, n = 11). However, there were disproportionately more boys than girls among the excluded cases.

relative to the final sample. In the final sample, there were equal numbers of boys and girls, whereas boys comprised 71% of the excluded cases. This difference approached statistical significance, χ^2 (1, N = 88) = 3.07, p < .10.

Materials

Warm-up task. For the warm-up task, three objects were presented on a red plastic tray: a red plastic cup, a small stuffed gray cat, and a small picture book.

Food request task. Two identical off-white, plastic bowls were used to hold the food. One bowl contained a handful of Pepperidge Farm Goldfish crackers (cheese-flavoured) and the other, a handful of raw broccoli flowerets. Two children were unable to eat the cheese crackers due to food allergies, so cookies or another type of cracker were substituted for the cheese crackers.

Object request task. Four pairs of items were used, each including one feminine- and one masculine-stereotyped item: (a) a doll and a car; (b) a purse and a baseball glove; (c) a necklace and a tie; and (d) a pink and a blue stuffed bear. (Pictures of the items used are presented in Appendix A). Ten adult judges (5 men, 5 women) rated each object on a 5-point scale, with a score of 1 indicating that the object was likely to be *preferred much more by boys*, a score of 3 indicating that the object was likely to be *preferred equally by boys and girls*, and a score of 5 indicating that the object was likely to be *preferred much more by girls*. All of the objects included in the present study, with the exception of the blue stuffed bear, were judged to be highly gender stereotyped by the adult raters (see Appendix B for means, standard deviations, and t values for the object ratings). However, the pair of blue and pink bears was retained for use in the study as the pink bear was judged to be highly feminine stereotyped and significantly more feminine than the blue bear, $t(9) = 7.97$, $p < .001$. Moreover, although the blue bear was not given a high masculine rating when evaluated independently, it was felt that the adult judges would have consistently selected the blue bear as the masculine-typed item and the pink bear as the feminine-typed item if they had been presented with both bears and asked to select one

for the boys and one for the girls. Such a forced-choice procedure was used with the children in the present study. The items comprising each pair were matched as closely as possible for size, material, and saliency.

Gender-labelling task. As a warm-up task, the children were first shown coloured pictures of a teddy bear and a bunny rabbit on facing pages of a looseleaf binder. These pictures were taken from a children's picture book and birthday card, respectively, and were displayed against a white background. Twelve paired, coloured photographs of children and adults followed, the first six depicting highly stereotypical boys and girls and the latter six, highly stereotypical men and women. The photographs displayed the head and shoulders of school-aged children and adults, all fully clothed, displayed on a white background. The photos were taken from magazines and mail-order catalogues and the individuals comprising each pair were matched as closely as possible with respect to apparent age, size of face, appearance (e.g., race, hair and skin colour), pose, and facial expression displayed. Two versions of the gender-labelling task were used: the majority of the participants (n = 60) received a version of the task containing pictures of children and adults of European-American descent. A second version containing pictures of children and adults from various ethnic backgrounds, including African-, Asian-, East Indian-, and European-American descent, was used for children who were members of visible minority groups (n = 4). Examples of the child and adult photograph pairs used in the European-American and Visible Minorities versions of the gender-labelling task are presented in Appendix C. Ten adult judges (5 men, 5 women) rated the gender stereotypicality of each of the photographs on a 5-point scale, with a score of 1 indicating that the child or adult pictured was *highly masculine*, a score of 3 indicating that the child or adult pictured was *gender neutral or ambiguous*, and a score of 5 indicating that the child or adult pictured was *highly feminine*. Mean ratings for the photographs were all significantly different from the gender-neutral score of 3, all ts(9) $\geq |2.45|$, ps $< .05$ (see Appendix D for means, standard deviations, and t values for the picture ratings).

Design

The experiment consisted of two 45 min testing sessions, spaced a week (maximum of two weeks) apart. Each testing session began with a warm-up task, intended to familiarize the child with the demands of the experimental tasks and to assess the child's motivation or willingness to share the objects with the experimenters. Following the warm-up trials, two tasks were administered: (a) the food request task, consisting of a single baseline and desire trial; and (b) the object request task, consisting of four baseline trials, four desire trials, and four stereotyping trials. The food request task always preceded the object request desire trials as the former task was assumed to be the easier of the two, and thus, served as a warm-up for the more difficult object request desire trials. For the object request task, each of the four pairs of items was presented three times, once during the baseline trials, once during the desire trials, and once during the stereotyping trials. A summary of the experimental design is presented in Table 1. One female and two male adults served as the experimenters throughout the course of the study. The male experimenters each tested half of the boys and half of the girls in each age group and experimental condition. The same male experimenter tested the child in both testing sessions.

On each visit, one experimenter administered both the food request trials and the object desire trials, and the other experimenter administered the stereotyping trials. Since the same object pairs were used in both the object desire and stereotyping trials, different experimenters were needed for each of these tasks to prevent carry-over effects from one task to the other. Thus, if the male experimenter administered the food request and object desire trials, the female experimenter administered the stereotyping trials and vice versa. On the second visit, these roles were reversed. Following from the previous example, on the second visit, the female experimenter would administer both the food and object desire trials and the male experimenter, the stereotyping trials. Half of the children received the female desire and male stereotyping trials first, the other half received the

Table 1

Summary of Experimental Design

Experimental Group	Task Order	Visit 1		Visit 2	
		Order 1	Order 2	Order 1	Order 2
gender consistent (n = 32)	stereotyping trials first (n = 16)	1. warm-up (3) 2. OR: female stereotyping (4) 3. food request task (1) 4. OR: male desire (4)	1. warm-up (3) 2. OR: male stereotyping (4) 3. food request task (1) 4. OR: female desire (4)	1. warm-up (3) 2. OR: male stereotyping (4) 3. food request task (1) 4. OR: female desire (4) 5. gender-labelling task	1. warm-up (3) 2. OR: female stereotyping (4) 3. food request task (1) 4. OR: male desire (4) 5. gender-labelling task
	desire trials first (n = 16)	1. warm-up (3) 2. food request task (1) 3. OR: male desire (4) 4. OR: female stereotyping (4)	1. warm-up (3) 2. food request task (1) 3. OR: female desire (4) 4. OR: male stereotyping (4)	1. warm-up (3) 2. food request task (1) 3. OR: female desire (4) 4. OR: male stereotyping (4) 5. gender-labelling task	1. warm-up (3) 2. food request task (1) 3. OR: male desire (4) 4. OR: female stereotyping (4) 5. gender-labelling task

Note. OR = object request task. The numbers in brackets indicate the number of experimental trials for that task. The procedure used in the gender-inconsistent condition was identical to that used in the gender-consistent condition.

male desire and female stereotyping trials first.

Food request task. For the two desire trials of the food request task, there were two preference conditions: (a) a match condition, in which the preference expressed by the experimenter was assumed to match the children's own preferences (i.e., a preference for the crackers over the broccoli); and (b) a mismatch condition, in which the experimenter expressed the reverse preference (i.e., a preference for the broccoli over the crackers). Each child received both preference conditions, one during the first visit and the other during the second visit, with a different experimenter expressing the desires for the foods across the two testing sessions (i.e., the male experimenter first visit, followed by the female experimenter second visit, or vice versa). These conditions were administered in a counterbalanced order across participants. In addition, the side on which the desired item was presented was counterbalanced across the two trials and across children, and the experimenter began with the item on the left for half of the children and with the item on the right for the other half of the children. The emotion (pleasure or disgust) expressed first was also counterbalanced across children. The side of presentation of the bowls was reversed from the baseline to the desire trial, and was counterbalanced across the two visits and across children.

Object request task. The order of presentation of the object request desire and stereotyping trials was counterbalanced across participants, such that half of the children received the desire trials first, followed by the stereotyping trials, and the other half received the reverse order. The desire and stereotyping trials were presented in the same order across both testing sessions. The baseline object trials were administered alternately with the first set of experimental object trials, either desire or stereotyping. For example, for those children receiving the desire trials first, the children would receive the baseline trial for a particular item pair (e.g., doll and car), following which the experimenter would administer the desire trial using the same pair. The experimenter would then proceed with the baseline trial for the next pair (e.g., purse and baseball glove). The side on which the

items were presented was switched from the baseline trial to the stereotyping or desire trials.

In both the desire and stereotyping trials of the object request task, the four pairs of items were presented in one of 24 possible orders. For each child, the order of presentation of the items differed in the desire and stereotyping trials, and across visits. Thus, each child was randomly assigned four different object pair orders: two different presentation orders were used for the desire and stereotyping trials on Visit 1, and two additional presentation orders were used for the desire and stereotyping trials on Visit 2.

For the desire trials, an equal number of boys and girls in each age group were randomly assigned to two experimental conditions: (a) a gender-consistent desire condition; (b) a gender-inconsistent desire condition. In the gender-consistent desire condition, the experimenter always expressed a preference or desire for the gender-appropriate objects (e.g., the male experimenter expressed a desire for the masculine-typed objects and disgust for the feminine-typed objects). Conversely, in the gender-inconsistent desire condition, the experimenter always expressed a preference or desire for the gender-inappropriate objects (e.g., the male experimenter expressed a desire for the feminine-typed objects and disgust for the masculine-typed objects). The emotion expressed first and the side on which the desired item was presented were counterbalanced across trials, as was the side on which the masculine- and feminine-typed objects were presented. In addition, the experimenter began by expressing the designated emotion for the item on the left for half of the trials, and for the item on the right for the other half of the trials.

Each child observed both the male and female experimenters expressing a preference for the objects (i.e., both male and female desire conditions). One experimenter expressed preferences for the four object pairs during the first visit, and the second experimenter expressed his/her preferences for the same four object pairs during the following visit. The order of administration of the female and male desire conditions

was counterbalanced across participants.

Gender-labelling task. Following the food and object request tasks in the second visit, children were given a gender-labelling task, similar to that used by Leinbach and Fagot (1986), to assess their gender-category knowledge. As a warm-up task, the children were first presented with pictures of a rabbit and a teddy bear on facing pages of a loose-leaf binder. This warm-up trial was used to familiarize the children with the task and to assess their ability to perform the gender-labelling task. The children were then shown a series of 12 paired photographs of highly stereotypical school-aged children and adults, arranged in male-female pairs, displayed on facing pages of a loose-leaf binder. The first set of six paired photos were of boys and girls, and the remaining six pairs, men and women. The photos were presented in the same order for all of the children and the side on which the target picture appeared was counterbalanced across picture pairs. For half of the children, one photo in each pair served as the target; for the other half of the children, the other photo in the pair served as the target.

Procedure

Children were tested individually in a laboratory room, seated at a table in a booster chair, facing a male and female experimenter. The parent (usually the mother) was seated directly behind the child and was instructed to remain neutral and to limit his or her interaction with the child during the course of the study. Thirteen children refused to sit in the booster chair and therefore, were tested at the table while sitting on their parents' laps during one or both visits. The majority of the children were tested in either English (n = 60) or French (n = 1), depending on which was their mother tongue, as indicated by the parent. Three additional French-English bilingual children were tested using both of their languages to facilitate their comprehension. One camera videotaped the children's behaviours during the testing sessions, while a second camera videotaped the experimenters' facial expressions. The output from each of these cameras was recorded as a frame-in-frame image on the videotape (i.e., the experimenters' facial expressions

appeared in the top left-hand corner of the screen, with the child's behaviours during the testing sessions occupying the remainder of the screen). The recordings of the experimenters' facial expressions were subsequently coded by adult judges to ensure that they were reliably expressing pleasure and disgust.

For both the food and object request tasks, a second experimenter was used to help with the task set-up. The second experimenter took out the objects and food bowls, placed them on the cardboard tray, and presented them to the child or placed them in front of the first experimenter, as needed. This procedure ensured that the experimenter whose desires for the items were to be inferred did not have previous contact with the objects or food items, thus maximizing the ecological validity of the task.

Warm-up task. For the warm-up task, the child was presented with three objects, a cat, a cup, and a book, on a plastic tray and allowed to explore them briefly. The first experimenter, who was seated directly in front of the child, then asked the child to give him or her each of the three items, in turn, saying "Can you give me the cat/cup/book?" Following compliance, the experimenter thanked the child. If the child did not give the item when requested, the parent was asked for the item and thanked following compliance in order to model the appropriate behaviour. The object was then returned to the plastic tray and the experimenter asked the child again for the item.

Food request task: baseline preference trial. The procedure used for the food request task was similar to that used by Repacholi and Gopnik (1997) in their study of desire understanding in 14- and 18-month-old infants. A bowl of broccoli and a bowl of cheese-flavoured crackers were first presented to the child on the plastic tray by the second experimenter. The children were encouraged to sample the food and were given 35 s to examine and/or taste the food items, following which the bowls were removed from their immediate reach. This baseline trial provided a measure of the children's initial food preferences.

Food request task: desire trial. The desire trial followed the baseline trial. The second experimenter placed the two bowls shoulder-width apart, on a black sheet of cardboard, in front of the first experimenter. The first experimenter tasted each food, expressing pleasure or happiness for one type of food and disgust or distaste for the other type of food by producing designated facial and vocal expressions. The expressions of pleasure and disgust each lasted approximately 5 s. The first experimenter then placed his or her hands, palms facing up, midway between the two bowls and requested some food, saying "I want some. Can you give me some?" Immediately following the food request, the second experimenter pushed the two bowls toward the child on the sheet of cardboard. Thus, the request was made before the child had reached for either bowl. If the child did not immediately offer any food, additional requests were made. Precautions were taken to make additional requests only when the children had no food in their mouths or hands or were not in the process of reaching toward a bowl to avoid biasing the children's responses (e.g., they might have assumed that the referent of a request was the food in their hands). If the child failed to comply with the request in a reasonable amount of time (i.e., about 1 min), the desire expressions were repeated and the experimenter again requested some food. If the child failed to comply with additional requests, the experimenter proceeded with the next trial.

In the match condition, the experimenter facially expressed happiness after tasting the crackers and said, "Mmm! Crackers! I tasted crackers! Mmm!" After sampling the broccoli, the experimenter displayed a disgusted expression and pushed the bowl away, saying, "Eww! Broccoli! I tasted broccoli. Eww!" In the mismatch condition, the preferences were reversed, with the experimenter expressing a preference or desire for the broccoli, and disgust for the crackers. The facial expressions of happiness and disgust displayed by the experimenter were consistent with the descriptions provided by Ekman and Friesen (1975). To avoid biasing the child's response, no specific behavioural cues (i.e., gaze or gesture toward one of the two bowls) accompanied the request. A schematic

representation of the procedure used for the food request desire trials (mismatch condition) is presented in Figure 1.

Object request task: baseline preference trials. During the object baseline trials, the second experimenter placed a masculine and feminine item pair, shoulder-width apart, in front of the child and allowed him or her to play with the items for 35 s. These trials served to familiarize the child with the objects and also provided a measure of the child's initial object preferences. The items were then removed from the child's immediate reach and the experimenter proceeded with the corresponding stereotyping or desire trial, using the same item pair.

Object request task: desire trials. In the object desire trials, the second experimenter placed a masculine and a feminine item shoulder-width apart in front of the first experimenter, on the black sheet of cardboard, out of the child's reach. The first experimenter expressed a preference for one of the objects by holding the item close to his or her body (near his or her face), rocking side to side, and smiling, while exclaiming in an animated voice, "Wow! A (label for object). Wow!" The experimenter expressed disgust for the other item by grimacing and pushing the object away from him or her, saying, "Eww! A (label for object). Eww!" The same experimenter then placed both hands, palms facing up, midway between the two items, saying, "I want one. Can you give me one?" Immediately following this request, the second experimenter pushed the objects toward the child on the sheet of cardboard. Following compliance, the items were removed and the child was thanked or praised by the second experimenter. Additional requests were made following the same procedure used in the food request task. If the child gave the experimenter both objects, they were again placed in front of the child and the child was requested to give the experimenter "only one." A schematic representation of the procedure used for the object request desire trials (gender-consistent condition) is presented in Figure 2.

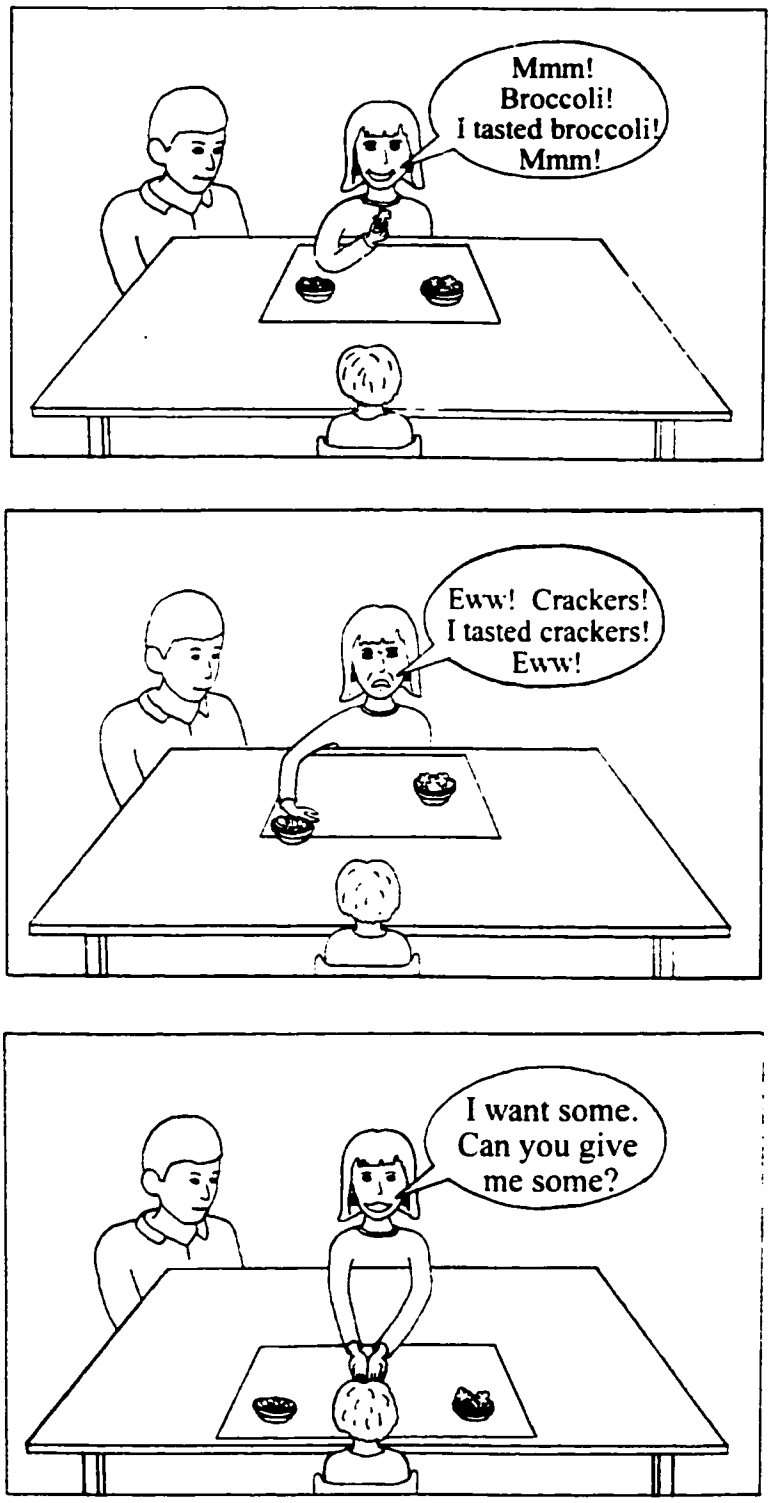


Figure 1. Schematic representation of the procedure used in the food request desire trials: Mismatch condition.

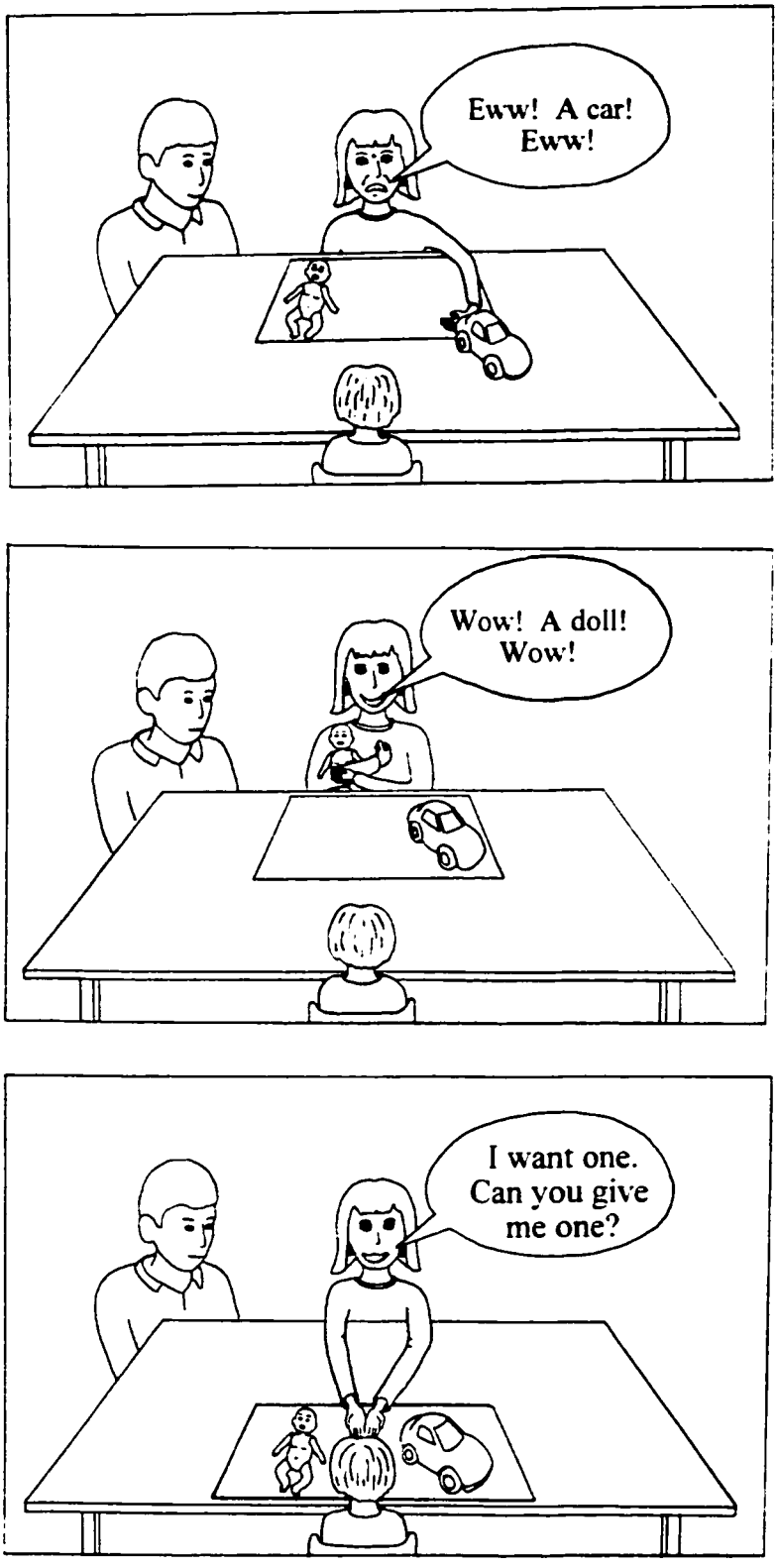


Figure 2. Schematic representation of the procedure used in the object request desire trials: Gender-consistent condition.

Object request task: stereotyping trials. In order to ensure that the items used in the study were gender stereotyped by the children, four stereotyping trials were also included. During each of these trials, the first experimenter placed a masculine- and feminine-typed object shoulder-width apart on the cardboard in front of the second experimenter (who now sat in front of the child). The second experimenter then placed his or her hands, palms facing up, in between the two objects, saying, "I want one. Can you give me one?" During these trials, no desires for the objects were expressed. If the child was aware of the gender stereotyping of the objects, he or she was expected to give the experimenter the item appropriate for his or her gender (e.g., the child was expected to give the baseball glove to the male experimenter). All children received both the male and the female stereotyping conditions, one on each visit to the laboratory.

Gender-labelling task. The children were first shown pictures of a teddy bear and a rabbit on facing pages of the looseleaf binder, and were asked to point to the "teddy bear" or "bunny rabbit." Following compliance, the child was praised. If the child selected the incorrect picture, the experimenter corrected the child. Children were then shown each of the 12 paired photographs of boys and girls, and men and women, in turn, and were asked to point to or touch the "boy" or "girl", "mommy" or "daddy" (e.g., the experimenter asked the child, "Where is the mommy/daddy? Can you show me the mommy/daddy?"). If the child did not respond to the labels "Mommy" and "Daddy", other labels such as "lady" and "man" were substituted.

The primary experimenter (female) administered the gender-labelling task to all of the participants. The female experimenter sat at the table facing the child during the task. To avoid cueing the child, the female experimenter held the binder upright with the photographs facing the child, so that she was unable to see the location of the target photographs. As she was unable to see the child's responses, a second experimenter sat next to the child and indicated when the child had made a response (i.e., pointed to, or touched, one of the pictures). Children who failed to complete a minimum of three of the

six child trials and three of the six adult trials were excluded from the analyses involving this measure ($n = 4$). An additional three children received the inappropriate version of the gender-labelling task for their racial group; thus, their data were also excluded.

Questionnaire Measures

Parents were asked to complete the Sex-Typed Child Care and Household Tasks Questionnaire (see Appendix E), which is a measure of the degree to which parents divide child care and household tasks according to gender-role norms. This questionnaire was originally developed by Orlofsky (1981) and was adapted by Serbin et al. (1993). The Serbin et al. adaptation of the questionnaire consists of 49 items (28 masculine-typed and 21 feminine-typed) and was shown to have adequate internal consistency (.87 and .88, for the masculine and feminine scales respectively). For the purposes of the present study, one masculine household task ("tuning a hard-to-get T.V. station") and one feminine child-care task ("regulating eating") were removed as they were considered to be outdated and/or confusing, and four feminine-typed "filler" items were added to achieve a better balance between the masculine and feminine items (e.g., "disciplining the girls" and "talking to girls about their problems" were added to match the masculine-equivalent questions, "disciplining the boys" and "talking to boys about their problems", that were already included in the questionnaire). Thus, the revised questionnaire had a total of 51 questions (27 masculine-typed and 24 feminine-typed). Since the four added feminine-typed items were not part of the original questionnaire, they were not included in the calculation of the scale scores. For both the child-care and household tasks, parents were asked to specify who typically does each activity at home, self or spouse, using a 5-point Likert-type scale (1 = *spouse only*, 3 = *both spouses equally*, 5 = *self*). Infrequently, an individual other than the parents was reported to regularly perform the activity at home (e.g., nanny, relative). In these cases, the parents were asked to specify the gender of the person who performed the activity. Examples of traditionally masculine tasks included "mowing the lawn" and "playing catch with the children", whereas traditionally feminine

tasks included "doing housework" and "infant care."

From this questionnaire, measures of both the fathers' and mothers' participation in traditional child-care and household tasks were computed. Items were coded in the same manner as in the Serbin et al. (1993) study, such that lower scores indicated greater participation by fathers on masculine tasks and by mothers on feminine tasks. These ratings were averaged across the masculine items and across the feminine items to create two scores. An overall gender-typing score was also computed by averaging parents' ratings across the masculine and feminine items. As the questionnaire data on the division of child-care and household tasks requires both parents to be living at home to be meaningful, the data from the one single parent family in the sample were excluded from analyses involving this measure. An additional family was missing more than 50% of the items comprising the masculine tasks subscale and therefore did not receive a mean score for this subscale.

Parents were also asked to complete an Item Familiarity Questionnaire that assessed which of the toys or objects used in the study the children had at home or at daycare and the frequency with which the children played with, or were exposed to, these items (see Appendix F). Questions inquired whether the child possessed the exact items used in the study (i.e., same make and colour) or other items from the same class (e.g., other types of dolls). Parents were also asked to rate the frequency with which their children played with, or wore, these items on a 5-point Likert-type scale, ranging from *never* to *always*. Additional questions assessed whether, and how frequently, parents wore necklaces or ties in the homes, as well as about the colour of the children's toys, clothes, and bedroom decor. Parents' ratings of the frequency with which their children played with gender-appropriate toys, or observed the gender stereotyping of the items used in the study at home, were used to generate an overall gender-typing score for each child. This score ranged from 0 (*non-gender-typed*) to 1 (*gender-typed*).

Behavioural Coding

The testing sessions were videotaped and were subsequently coded by the primary experimenter.

Food request task: baseline preference trials. Children received one of four food preference codes: (a) no preference; (b) prefers broccoli; (c) prefers crackers; (d) neither food tasted/touched. If the child sampled one or both food items, the child's food preference was based on the amount of each food eaten, the child's facial expression in response to tasting each food (i.e., pleasure or disgust), and any behaviours indicating rejection of the food item (e.g., food removed from mouth, thrown on floor). If neither food was sampled, the amount of time spent touching each bowl or food item was used to determine the child's food preference.

Food request task: desire trials. For the desire trials, the food offered by the child in response to the experimenter's request was recorded, as well as the latency of the child's response. Latency was measured in hundredths of a second from the time the request was made to when the offered food touched the experimenter's hand. The primary observer also recorded frequency counts for any emotional expressions such as surprise, smiling, and laughter displayed by the child, as well as behaviours that may be indicative of teasing (e.g., a child offering a food item then removing it from the experimenter's hand; see Appendix G for the complete coding scheme). These emotional responses or behaviours could be displayed by the child during the experimenter's affective displays or after the experimenter's request for a food item.

Object request task: baseline preference trials. The baseline preference trials were coded for the total time spent touching or playing with each object, and for behaviours indicating rejection of the object (e.g., pushing the object away, giving the object back to experimenter, throwing the object on the floor). Four preference categories were used: (a) no preference; (b) prefers masculine object; (c) prefers feminine object; and (d) touches neither object.

Object request task: desire and stereotyping trials. The child's responses during the desire and stereotyping trials were analyzed in terms of which object the child offered the experimenter, as well as the latency of response for each trial (measured in the same way as for the food request task desire trials). The primary observer also recorded any emotional expressions and teasing behaviours displayed by the child.

Gender-labelling task. The pictures that the child pointed to, or touched, in response to the experimenter's queries were coded from the videotapes. The children received a score of 1 for each target picture identified correctly. Scores were summed across the child and adult pictures to create three separate scores, a child gender-labelling score, an adult gender-labelling score, and a total score.

Interobserver Reliability

Interobserver reliability for each of the above measures was assessed by having a second coder, naive to the hypotheses of the study, code 20% of randomly selected testing sessions from the videotapes. Only the child was visible from the videotapes (i.e., the portion of the screen displaying the experimenters' facial expressions was covered) and coders were blind to the child's assignment to experimental conditions. Reliability for the food offered during the food request task and for the objects offered during both desire and stereotyping trials of the object request task were assessed by coding these trials without sound. Coding was completed without sound to ensure that the coder remained blind to the child's assignment to conditions (i.e., the coder was unaware of which food item/object was desired by the experimenter). To assess interobserver reliability for the child's response latency, the experimenters' desire expressions were not audible. Thus, the coder heard only the request for a food item or object (i.e., "Can you give me some/one?). The reliability coding for the gender-labelling task was also completed without sound to ensure that the coder was unaware of which picture served as the target for a particular trial.

For the baseline food and object preference measures, a Cohen's kappa coefficient

of .90 was obtained. Kappa coefficients for the food or object offered during the experimental trials were similarly high, with obtained kappas of .93 for the food offered during the food request task, .98 for the object offered during the object request task desire trials, and 1.00 for the object offered during the object request task stereotyping trials. To determine the reliability of the response latency measures, intraclass correlations were computed between the latency ratings of the two judges. An average intraclass correlation of .92 was obtained. For the coding of the children's emotional expressions and teasing behaviours, interobserver reliability was assessed by calculating the percentage agreement (agreements / agreements + disagreements) between coders' judgements that a particular emotional response or behaviour was present. Interobserver agreement was computed for each expression or behaviour category separately, and was above .88 in all cases. Finally, for the gender-labelling task, a kappa of .95 was obtained.

Reliability of Experimenters' Facial Expressions

Videotapes showing only a frontal view of the experimenters' faces (i.e., the portion of the screen displaying the child's behaviours during the testing session was covered) were used as a check to ensure that the facial expressions of happiness and disgust were clearly recognizable. Two naive adult coders examined a random subset of the testing sessions on the videotapes without sound, so that only the experimenters' faces were used in the emotion judgements. It was necessary to code the tapes without sound as the vocal expressions of pleasure and disgust ("Wow" and "Eww") would have served as obvious cues to the emotion displayed. The experimenters' facial expressions were categorized in terms of their overall hedonic tone (positive, negative, neutral), the predominant emotion displayed (happiness, interest, surprise, anger, disgust, fear, sadness, neutral, other), as well as the perceived intensity of the emotion (1 = *very mild* to 5 = *very intense*). This was done for each of the experimenters used in the study (1 female, 2 males). Interobserver kappa coefficients for overall hedonic tone and discrete emotion displayed were all above .88 for the three experimenters and indicated that the

appropriate emotions (i.e., happiness and disgust) were reliably identified from the experimenters' emotional displays. Ratings of the perceived intensity of the emotion were not reliable, however, and therefore were not analyzed further. Such difficulties in establishing reliability for emotion intensity ratings were similarly reported by Repacholi and Gopnik (1997).

Results

For both the food and object request tasks, the first food or object offered to the experimenter on the desire and stereotyping trials served as the dependent measure. In those infrequent cases in which the child offered the experimenter an item then withdrew it (less than 2% of the trials), the final item given to the experimenter was used. In the case of missing data (i.e., trials in which the child failed to offer the experimenter any food item or object), the data were averaged across the remaining trials. Overall, less than 4% of the trials were missing data: 3% of the food desire trials, 2% of the object desire trials, and 2% of the object stereotyping trials. As mentioned in the Method section, children missing data from more than 50% of the trials within any of the experimental conditions (male desire, female desire, male stereotyping, female stereotyping) were excluded from the analyses. Analyses were also conducted on the children's mean response latencies for the food and object request tasks (desire and stereotyping trials). These analyses failed to yield any significant main effects or interactions and are not reported in the present paper. For the gender-labelling task and questionnaires, missing data were handled similarly to the data from the food and object request tasks, with the data averaged across the remaining items. An alpha level of .05 (two-tailed) was used for all statistical tests.

Warm-Up Task

Children's performance on the warm-up task was first examined to determine their understanding of the task requirements. Compliance with the experimenter's initial requests for the objects was high, with 56.25% of the children giving the experimenter the

requested items on at least five of the six trials and 92.18% of the children giving the experimenter the requested items on three or more of the trials across the two visits. With modelling of the appropriate behaviour and/or correction, all of the children successfully completed at least five of the six warm-up trials. Compliance on the warm-up task was similar across the two age groups, genders, and experimental groups, as well as across the three experimenters. Thus, the children appeared to have understood the requirements of the task and were compliant in giving the experimenters the requested objects. Performance during the warm-up task was not related to performance on the experimental tasks (see Appendix H, for correlations).

Food Request Task

Children were administered the food request task in an attempt to replicate Repacholi and Gopnik's (1997) findings of desire understanding for food items in children as young as 18 months of age. Furthermore, their performance on this task provided a baseline measure of desire understanding against which their understanding of desires for nonfood items (gender-typed objects) could be compared. Finally, since previous research had shown that 73% of children were successful at this task as early as 18 months of age (Repacholi & Gopnik, 1997), this task also served as a warm-up task for the more difficult object request task that followed.

Baseline food preferences. During the baseline food trials, the children were presented with broccoli flowerets and cheese crackers and were allowed to sample each food to determine their initial food preferences. As expected, the majority of children displayed a preference for the crackers over the broccoli during the baseline trials, with 75.00% of the children displaying a preference for the crackers on Visit 1 and 64.06% on Visit 2. The remaining children displayed a preference for the broccoli (7.81% and 9.38% on Visit 1 and Visit 2, respectively), did not touch or taste either food (6.25% and 3.13%), or displayed no food preference (10.94% and 23.44%). Of those children who displayed a preference for one of the two foods, 90.57% and 87.23% preferred the

crackers over the broccoli on Visit 1 and Visit 2, respectively, $\chi^2(1, n = 53) = 34.89$, $p < .001$, and $\chi^2(1, n = 47) = 26.06$, $p < .001$. Overall, there was a high degree of consistency in children' food preferences across the two visits, with 71.19% of the children maintaining their preference for crackers, broccoli, or neither food across the two visits.

For the desire trials, each child received two preference conditions, one on each visit: (a) a match condition, in which the experimenter expressed a preference for crackers, which was assumed to match the child's own preference; and (b) a mismatch condition, in which the experimenter expressed a preference for broccoli, which was assumed to differ from the child's own preference. As the baseline food trials indicated that some children did not prefer the crackers over the broccoli, these children were reassigned to the appropriate group based on their preference data. For example, a child who demonstrated a preference for broccoli on the baseline trials and observed the experimenter expressing a preference for broccoli on the desire trials was reassigned to the match preference group. The results of the analyses with the reassigned children did not differ in any respect from the analyses conducted on the original match and mismatch groups. Thus, for simplicity, only the analyses from the original preference groups are presented.

Understanding of others' desires for food. To determine children's understanding of desires, the food offered to the experimenter in response to his or her expressions of pleasure or disgust for the food items was examined through a series of chi-square analyses. It was hypothesized that children at both 20 and 24 months of age would correctly infer the experimenter's desires from the affective displays, offering the experimenter crackers in the match condition and broccoli in the mismatch condition.

Of the 60 children who offered a food item to the experimenter on both visits, only 25.00% correctly offered the desired food on both trials, 56.67% correctly offered the desired food on one of the two trials (56.67% on Visit 1 and 50.00% on Visit 2), and

18.33% failed to offer the desired food on either trial. This pattern of performance closely corresponded to the pattern expected by chance (i.e., a 25% chance of making two correct or two incorrect choices, and a 50% chance of making one correct choice across the two trials), $\chi^2(2, n = 60) = 1.60$, *ns*. There were no differences in performance between the two age groups, nor between boys and girls, $\chi^2(2, n = 60) = 0.94$, *ns*, and $\chi^2(2, n = 60) = 3.94$, *ns*, respectively. Examination of the children's performance in the match and mismatch conditions separately revealed that 57.37% of the children correctly inferred that the experimenter desired crackers in the match condition, whereas only 47.62% of the children correctly inferred that the experimenter desired broccoli in the mismatch condition. Neither of these response patterns differed significantly from chance, $\chi^2(1, n = 61) = 1.33$, *ns* and $\chi^2(1, n = 63) = 0.14$, *ns*, respectively. Thus, contrary to prediction, the present results failed to replicate Repacholi and Gopnik's (1997) findings of desire understanding, even in this older sample of children.

In the current study, each child completed two food desire trials, one in which the experimenter expressed a preference for crackers, and one in which the experimenter expressed a preference for broccoli. In the study by Repacholi and Gopnik (1997), however, each child completed only one food desire trial, with children randomly assigned to one of the two food preference conditions. In an attempt to replicate Repacholi and Gopnik's findings, therefore, the data from the first food desire trial was examined separately. Moreover, as task order effects may have interfered with children's performance on the food request task, only the subset of children ($n = 32$) who received the food request task first (i.e., immediately following the warm-up task) were included in the analyses. The results revealed that 64.52% of the children complying with the experimenter's request ($n = 31$) correctly offered the desired food, which was not significantly different from chance (i.e., 50%), $\chi^2 (1, n = 31) = 2.61$, *ns*. There was no significant difference in the proportion of correct responses in the match (64.29%) versus mismatch (64.71%) preference groups, $\chi^2(1, n = 31) = 2.58$, *ns*, nor were there any effects

of age group or sex, $\chi^2(1, n = 31) = 0.59$, *ns*, and $\chi^2(1, n = 31) = 1.59$, *ns*, respectively.

Although children's performance on the food desire trial in Visit 1 was somewhat better than their performance across the two visits, the results failed to reach statistical significance. Only 64.52% of the children were found to correctly offer the experimenter the desired food. In comparison, 73% of the 18-month-olds who complied with the experimenter's request ($n = 55$) in the Repacholi and Gopnik (1997) study were found to correctly offer the desired food, which was significantly above chance. Thus, the present study failed to replicate their results, finding no evidence of desire understanding for food items even in this older sample of 20- and 24-month-old children. The proportion of children found to offer the experimenter the desired food in the present study, however, was similar to the proportion of 18-month-old infants found to offer the desired food in McKoy and Poulin-Dubois's (1999) replication of Repacholi and Gopnik's study.

Emotional and behavioural responses during the food request task. The emotional and behavioural responses displayed by the children during the food desire trial in Visit 1 (for those children who received the food desire trial first) were also examined to determine whether their relatively poor performance might be due to teasing behaviours. Of the 11 children who failed to offer the experimenter the desired food, there was no significant difference between the number of children who smiled or laughed and the number of children who did not smile or laugh during the experimenter's affective displays (5 vs. 6) or following his or her request for food (3 vs. 8), $\chi^2(1, n = 11) = 0.09$, *ns*, and $\chi^2(1, n = 11) = 2.27$, *ns*, respectively. Furthermore, only 2 of the 31 children initially refused to give the experimenter the desired food (i.e., offered and withdrew the desired food, or withheld the desired food following the experimenter's request): one later complied, giving the experimenter the desired food, and one gave the undesired food. Finally, gaze at the experimenter's face while offering a food item was also examined, as it was felt that children might monitor the experimenter's reaction if they were attempting to tease the experimenter. Of those children who failed to offer the

experimenter the desired food, significantly more children gazed at the experimenter's face than looked elsewhere (9 vs. 2) while offering the food item, $\chi^2(1, n = 11) = 4.45$, $p < .05$. In contrast, of those children who correctly offered the experimenter the desired food, roughly equal numbers of children gazed at the experimenter's face or looked elsewhere (9 vs. 11), $\chi^2(1, n = 20) = 0.20$, *ns*. It is possible that some of the children may have been reluctant to share the desired food with the experimenter. Thus, they may have gazed at the experimenter's face while offering the undesired food, in order to monitor the experimenter's reaction to their noncompliance. Given that 6 of the 9 children who gazed at the experimenter's face while offering the undesired food did so with no evidence of humour (i.e., no smiling or laughing), this behaviour is more likely to represent active noncompliance than attempts to tease the experimenter. It is also unlikely that these children were monitoring the experimenter's face because they were unsure of what item to give. If this were the case, the children should have been equally likely to randomly select the desired versus the undesired food to give to the experimenter. Thus, the present results suggest that children's failure to demonstrate an understanding of desires for food items cannot be attributed to their attempts to tease the experimenter by offering the undesired food. However, some children may not have complied with the experimenter's request for the desired food due to their reluctance to share these items.

Children may have been particularly reluctant to share the crackers when these were desired by the experimenter, since the vast majority of children demonstrated a preference for the crackers over the broccoli. Alternatively, if the children perceived the experimenter's expressed desire for broccoli as surprising or incongruous in the mismatch condition, they may have been reluctant to give the experimenter the desired food in this condition. Either of the above responses could have resulted in some children giving the undesired food, and these children may have monitored the experimenter's facial expressions to determine his or her reaction to their failure to give the desired food.

To test these hypotheses, children's emotional and behavioural responses were

compared across the match and mismatch preference conditions. Of those children who offered the experimenter a food item on the desire trial in Visit 1, 14 received the match preference condition, and 17 received the mismatch preference condition. Results revealed no difference in the number of children in the match and mismatch conditions who refused to give the desired food (1 vs. 1), or gazed at the experimenter's face while offering a food item (8 vs. 10). $\chi^2(1, n = 18) = 0.22$, *ns*. Thus, the type of food desired by the experimenter did not seem to influence children's emotional and behavioural responses and cannot account for the pattern observed in some children, in which they looked at the experimenter's face while offering an undesired item. It may be that some children were attempting to test the rules of the experimental situation by offering the experimenter the undesired food, independent of which type of food was desired. They then may have monitored the experimenter's face for a reaction to this active noncompliance. Research on children's interactions within the family suggest that children of this age frequently disobey their parents' requests or break family rules to test the consequences of rule violations (Dunn, 1988).

The above results also suggest that the children did not perceive the experimenter's expressed desire for broccoli as surprising or incongruous in the mismatch condition. If this were the case, children in the mismatch condition would be expected to be more likely to refuse to give the experimenter the desired food, or to monitor the experimenter's gaze to determine whether broccoli was indeed desired. However, as stated above, no differences were found in the number of children who refused to give the desired food, or gazed at the experimenter's face while offering a food, in the match versus mismatch conditions. Moreover, chi-square analyses revealed that similar numbers of children in the match and mismatch conditions smiled or laughed during the experimenter's affective displays (7 vs. 10) or following the experimenter's request for food (6 vs. 7), $\chi^2(1, n = 17) = 0.53$, *ns*, and $\chi^2(1, n = 13) = 0.08$, *ns*, respectively. Thus, the children's emotional and behavioural responses did not suggest that they found the

experimenter's desire for broccoli humourous or surprising.

Object Request Task

The purpose of the object request task was to assess whether young children used information about a person's gender to make inferences about his or her desires for gender-stereotyped objects.

Baseline object preferences. During the baseline trials, the children were presented with one masculine- and one feminine-stereotyped object, and were allowed to play with the items to determine their initial preferences. The percentage of trials in which the child showed a preference for the feminine or masculine objects, out of those trials in which the child touched or played with one or more of the objects, was computed. Children were considered to be gender stereotyped in their preferences if they demonstrated a preference for the gender-appropriate object on more than 50% of the trials.

Across both visits, 45.31% of the children displayed gender-typed preferences, 31.25% displayed cross-gender-typed preferences, and 23.44% displayed no object preference. Of the 49 children who displayed a preference for one of the two objects, similar numbers displayed gender-typed (n = 29) versus cross-gender-typed preferences (n = 20), χ^2 (1, n = 49) = 1.65, *ns*. There were no significant differences between the number of boys and the number of girls showing gender-typed, cross-gender-typed, or no object preferences (see Table 2). χ^2 (2, N = 64) = 3.56, *ns*. However, there was a great deal of inconsistency in children's object preferences across visits, with only 34.38% of the children displaying the same preference in both visits. In fact, a Pearson product-moment correlation computed between the percentage of trials in which a gender-typed preference was shown on Visit 1 compared to Visit 2, showed a nonsignificant correlation between children's preferences across the visits, r(64) = .17, *ns*. It was felt that the lack of correlation between children's object preferences on Visit 1 versus Visit 2 may be due to fatigue or boredom with the objects on Visit 2. Specifically, by the second visit,

Table 2

Number of Children Displaying Gender-Typed, Cross-Gender-Typed, and No Object Preferences Across Both Visits, in Visit 1 and Visit 2

Visit	Gender-Typed Preference	Cross-Gender-Typed Preference	No Preference
Across both visits			
Boys (n = 32)	18 (56.25%)	9 (28.13%)	5 (15.63%)
Girls (n = 32)	11 (34.38%)	11 (34.38%)	10 (31.25%)
Total (n = 64)	29 (45.31%)	20 (31.25%)	15 (23.44%)
Visit 1			
Boys (n = 32)	14 (43.75%)	6 (18.75%)	12 (37.50%)
Girls (n = 32)	11 (34.38%)	7 (21.88%)	14 (43.75%)
Total (n = 64)	25 (39.06%)	13 (20.31%)	26 (40.63%)
Visit 2			
Boys (n = 32)	9 (28.13%)	7 (21.88%)	16 (50.00%)
Girls (n = 32)	10 (31.25%)	13 (40.63%)	9 (28.13%)
Total (n = 64)	19 (29.69%)	20 (31.25%)	25 (39.06%)

the children had acquired a great deal of exposure to the objects, having had the opportunity to play with the items on the baseline trials, and observing two sets of experimental trials (desire and stereotyping) with the same items during Visit 1. Thus, the Visit 1 baseline trials were considered to be a more valid measure of children's preferences than the Visit 2 baseline trials, and further analyses were conducted on the Visit 1 data alone.

On Visit 1, 39.06% of the children displayed gender-typed object preferences, 20.31% displayed cross-gender-typed preferences, and 40.63% displayed no preference. Of the 38 children who displayed a preference for one of the two objects, there were somewhat more children who displayed gender-typed (n = 25) than cross-gender-typed preferences (n = 13). This effect approached statistical significance, χ^2 (1, n = 38) = 3.79, $p < .10$. There was no significant difference between the number of boys and the number of girls displaying gender-typed, cross-gender-typed, or no preferences (see Table 2). χ^2 (2, N = 64) = 0.59, *ns*. The number of children displaying gender-typed preferences versus cross-gender-typed or no preference was also compared for the two experimental groups (gender consistent and gender inconsistent) to ensure the equivalence of these groups. Of interest, there were significantly more children with gender-typed object preferences in the gender-consistent group compared to the gender-inconsistent group, χ^2 (1, N = 64) = 5.32, $p < .05$. Specifically, of the 32 children in the gender-consistent group, 17 children (53.13%) displayed gender-typed preferences and 15 children (46.88%) displayed either cross-gender-typed or no preferences. In contrast, of the 32 children in the gender-inconsistent group, only 8 children (25.00%) displayed gender-typed preferences, with the remaining 24 children (75.00%) displaying either cross-gender-typed or no preferences.

One possible explanation for this discrepancy between the experimental groups is that the children may have been influenced by the preferences displayed by the experimenters. That is, when the children in the gender-inconsistent group observed the

experimenters displaying cross-gender-typed object preferences, they may have been more likely to play with the cross-gender-typed items or with both items equally, compared to the children in the gender-consistent group who observed the experimenters displaying gender-typed object preferences. Although the children received the baseline trial for a particular object pair prior to observing the experimenter express a preference for the same pair, it is possible that the experimenter's preferences influenced the children's behaviour on subsequent trials. To test this hypothesis, the children who received the baseline trials interspersed with the desire trials were examined separately from the children who received the baseline trials interspersed with the stereotyping trials (see Figure 3). On the stereotyping trials, no preferences for the objects were expressed by the experimenters, so the children's object preferences would not be affected by the experimenters' preferences. In this case, children in the gender-consistent and gender-inconsistent groups would be expected to show similar gender-typed preferences. However, on the desire trials, preferences for the objects were expressed by the experimenters, which may have affected children's preferences on subsequent trials. In this case, the prediction is that the children in the gender-inconsistent group may be less gender stereotyped in their preferences than the children in the gender-consistent group.

Consistent with the above hypothesis, similar numbers of children were found to display gender-typed object preferences in the gender-consistent (50.00%) and gender-inconsistent groups (43.75%) for the 32 children who received the baseline trials interspersed with the stereotyping trials, $\chi^2(1, n = 32) = 0.13$, *ns*. In contrast, significantly fewer children were found to display gender-typed object preferences in the gender-inconsistent group (6.25%) relative to the gender-consistent group (56.25%) for the 32 children who received the baseline trials interspersed with the desire trials, $\chi^2(1, n = 32) = 9.31, p < .01$. Thus, the present results support the interpretation that children in the gender-inconsistent group, who observed the experimenter expressing cross-gender-typed object preferences on each trial, were less likely to show gender-typed preferences in

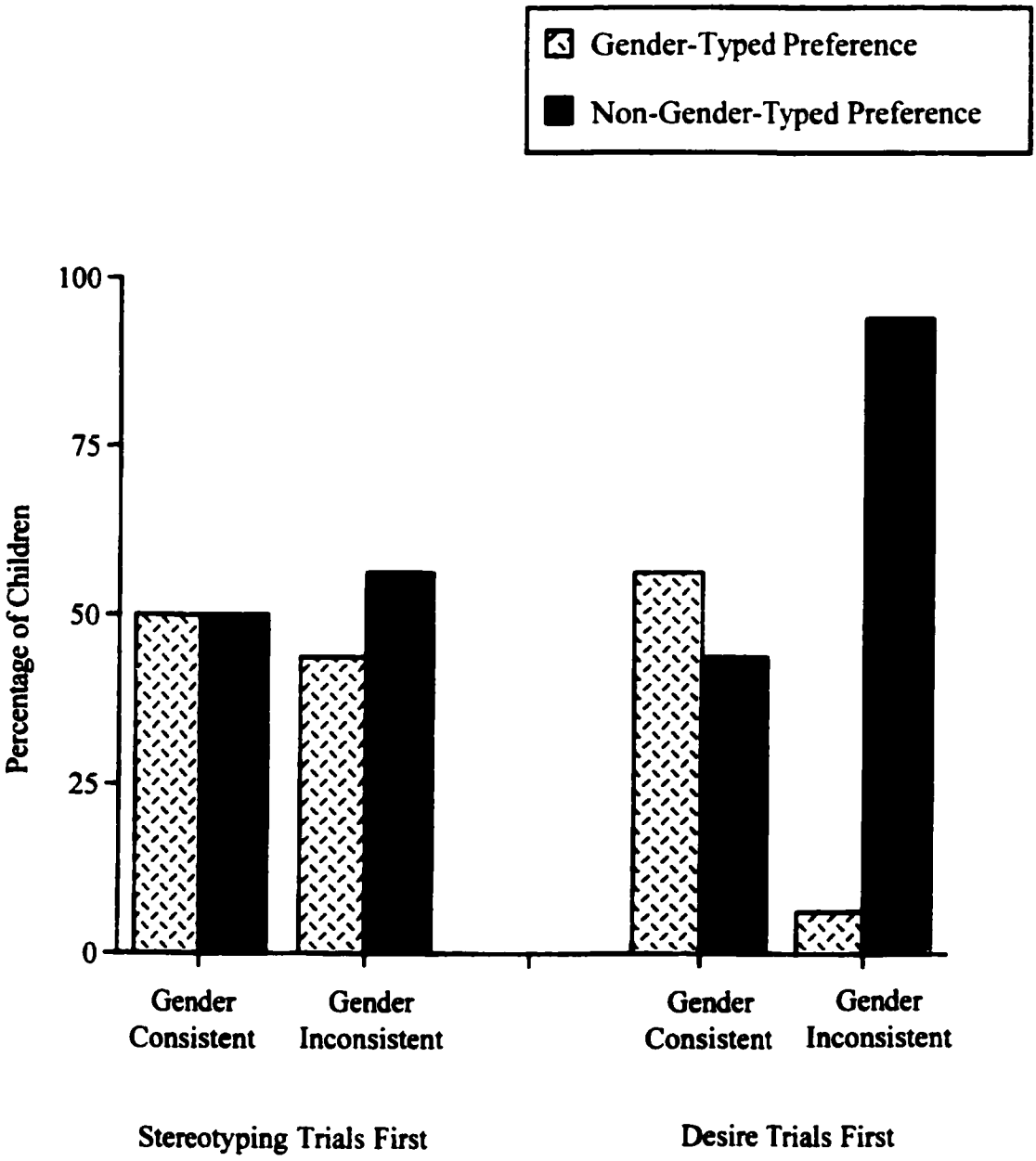


Figure 3. Percentage of children displaying gender-typed object preferences versus non-gender-typed object preferences (i.e., cross-gender-typed or no preferences) during the Visit 1 baseline trials, as a function of experimental group and task order.

subsequent baseline trials. Children in this group appeared to imitate, or pattern their behaviour after, the preferences displayed by the experimenter. Such modelling effects on children's gender-typed behaviours have been previously documented in older samples (see Bussey & Bandura, 1999; Powlishta et al., 2001).

Knowledge of the gender stereotyping of the objects. Children's performance on the stereotyping trials was examined to determine whether they were aware of the gender stereotyping of the objects included in the present study. To assess for visit and order effects, a preliminary 2 (visit) x 2 (task order: stereotyping vs. desire trials first) x 2 (condition order: male vs. female stereotyping first) mixed model ANOVA was first conducted on the percentage of trials in which the appropriate gender-typed object was offered to the experimenter (Appendix I, Tables I1 and I2). Visit was the within groups variable, and task and condition order were the between-groups variables. No significant effects of visit or order were found.

Children's stereotype knowledge was subsequently assessed by conducting a 2 (age group: 20 vs. 24 months) x 2 (sex of child) x 2 (group: gender consistent vs. gender inconsistent) x 2 (condition: male vs. female stereotyping) mixed model ANOVA on the percentage of feminine-typed objects given to the male and female experimenters during the stereotyping trials (Appendix I, Table I3). Age group, sex of child, and group were the between-groups variables, and condition was the within-groups variable. Knowledge of the gender stereotyping of the objects would be demonstrated by a main effect of condition, with children giving a higher proportion of feminine objects to the female experimenter than to the male experimenter. By 24 months of age, children were predicted to show knowledge of gender stereotypes. The mean percentage of feminine-typed objects given to the male and female experimenters during the stereotyping trials and standard deviations are presented in Tables 3 and 4. Only a significant Sex of Child x Age Group interaction emerged, $F(1, 56) = 4.63$, $p < .05$. Simple effects analyses, using the Bonferroni correction, revealed that at 24 months of

Table 3

Mean Percentages of Feminine-Stereotyped Objects (and Standard Deviations) Given to the Male and Female Experimenters During the Stereotyping Trials for the 20-Month-Old Children

Sex of Child	Male Experimenter		Female Experimenter	
	M	*SD*	*M*	*SD*
Boys (n = 16)				
Gender-consistent group	34.38	26.52	50.00	29.88
Gender-inconsistent group	59.38	22.90	53.13	31.16
Across groups	46.88	27.20	51.56	29.54
Girls (n = 16)				
Gender-consistent group	53.13	28.15	55.21	21.33
Gender-inconsistent group	56.25	29.12	56.25	22.16
Across groups	54.69	27.72	55.73	21.02
Total (n = 32)				
Gender-consistent group	43.75	28.14	52.60	25.22
Gender-inconsistent group	57.81	25.36	54.69	26.17
Across groups	50.78	27.30	53.65	25.31

Note. There were equal numbers of girls and boys in the gender-consistent and gender-inconsistent groups.

Table 4

Mean Percentages of Feminine-Stereotyped Objects (and Standard Deviations) Given to the Male and Female Experimenters During the Stereotyping Trials for the 24-Month-Old Children

Sex of Child	Male Experimenter		Female Experimenter	
	M	*SD*	*M*	*SD*
Boys (n = 16)				
Gender-consistent group	43.75	22.16	61.46	18.33
Gender-inconsistent group	61.46	12.55	68.75	22.16
Across groups	52.60	19.65	65.10	20.00
Girls (n = 16)				
Gender-consistent group	37.50	42.26	47.92	19.80
Gender-inconsistent group	43.75	17.68	43.75	25.88
Across groups	40.63	31.46	45.83	22.36
Total (n = 32)				
Gender-consistent group	40.63	32.76	54.69	19.71
Gender-inconsistent group	52.60	17.41	56.25	26.61
Across groups	46.61	26.51	55.47	23.05

Note. There were equal numbers of girls and boys in the gender-consistent and gender-inconsistent groups.

age, the boys gave a significantly higher proportion of feminine objects to both experimenters (M = 58.85%, SD = 20.09%) than did girls (M = 43.23%, SD = 20.09%), $F(1, 56) = 4.84$, $p < .05$. At 20 months of age, however, there was no difference in the proportion of feminine objects given to both experimenters by boys (M = 49.22%, SD = 20.09%) compared with girls (M = 55.21%, SD = 20.09%), $F(1, 56) = 0.71$, *ns*. Contrary to prediction, there was no main effect of condition, nor a Condition x Age Group interaction. The children did not give a significantly higher proportion of feminine items to the female experimenter (M = 54.56%, SD = 24.03%) than to the male experimenter (M = 48.70%, SD = 26.78%). Thus, an understanding of gender stereotypes for objects was not demonstrated in the present sample. Moreover, there was no interaction of condition with sex of child, indicating that girls were not more advanced in their knowledge of stereotypes than boys, contrary to the results of Serbin et al.'s (2001) study.

Understanding of others' desires for gender-typed objects. To determine children's understanding of desires for gender-typed objects, the objects offered to the experimenters on the desire trials were examined.

To assess for visit and order effects, a preliminary 2 (visit) x 2 (task order: stereotyping vs. desire trials first) x 2 (condition order: male vs. female desire first) mixed model ANOVA was first conducted on the percentage of trials in which the desired object was offered to the experimenter (Appendix I, Tables I4 and I5). Task and condition order were the between-groups variables, and visit was the within-groups variable. No visit or order effects were found.

Children's understanding of desires was then assessed by conducting a 2 (age group: 20 vs. 24 months) x 2 (sex of child) x 2 (group: gender consistent vs. gender inconsistent) x 2 (condition: male vs. female desire) mixed model ANOVA on the percentage of feminine-typed objects given to the male and female experimenters during the desire trials (Appendix I, Table I6). Age group, sex of child, and group were the between-groups variables, and condition was the within-groups variable. If the children

correctly inferred the experimenters' desires, a Group x Condition interaction would be expected. In the gender-consistent group, the children should give a greater proportion of feminine-typed objects to the female experimenter than to the male experimenter. Conversely, in the gender-inconsistent desire group, the children should give a greater proportion of feminine-typed objects to the male experimenter than to the female experimenter. A three-way Group x Condition x Age Group interaction was predicted. Specifically, it was hypothesized that the 20-month-old children would follow the experimenters' desires in both the gender-consistent and gender-inconsistent groups, regardless of the item's gender stereotyping, producing the Group x Condition response pattern described above. Conversely, the 24-month-old children were predicted to be more likely to follow the experimenters' desires in the gender-consistent group than in the gender-inconsistent group. In the gender-consistent group, they were hypothesized to perform similarly to the 20-month-olds, giving a greater proportion of feminine-typed objects to the female experimenter than to the male experimenter. In the gender-inconsistent group, in contrast, it was hypothesized that the 24-month-olds' greater knowledge of gender stereotypes might lead them to make more gender-based object selections. Thus, they might consistently select the gender-appropriate object to give to the experimenter, giving a higher proportion of feminine-typed objects to the female experimenter than to the male experimenter. Given previous research suggesting that girls may have more advanced knowledge of gender stereotypes than boys in the first few years of life (Poulin-Dubois et al., 2002; Serbin et al., 2001), a Condition x Age x Sex interaction might also emerge.

Other response patterns were also possible. The 24-month-old children might use both the experimenters' desire expressions and their gender-typed knowledge to determine their item selections, possibly resulting in equal proportions of feminine-typed objects given to the male and female experimenters. Alternatively, they may base their desire inferences primarily on the experimenter's affective displays, considering the

affective information more relevant or salient than the gender of the individual and the gender-typing of the objects. In this latter case, children would be expected to give the experimenter the desired objects in both the gender-consistent and gender-inconsistent conditions, despite showing significant knowledge of gender stereotypes on the stereotyping trials.

Mean percentages of feminine-typed objects given to the male and female experimenters during the desire trials, and standard deviations, are presented in Tables 5 and 6. The analysis revealed no significant main effects or interactions. Contrary to prediction, the children in the gender-consistent group did not give a greater proportion of feminine objects to the female experimenter (M = 49.22%, SD = 28.74%) than to the male experimenter (M = 51.30%, SD = 26.28%). Similarly, there was no significant difference in the proportion of feminine items given to the female experimenter (M = 44.53%, SD = 25.19%) compared to the male experimenter (M = 54.17%, SD = 26.77%) in the gender-inconsistent group. Thus, the children did not follow the experimenters' expressed desires in either age group, nor were there any differences in performance between the boys and the girls. Boys and girls in each age group were equally likely to offer the experimenter the undesired object as the desired object.

To summarize, the present study failed to find evidence of desire understanding for gender-typed objects in 20- and 24-month-old children. Moreover, no gender-stereotype knowledge for the objects was found. These results were surprising as previous research has shown that children as young as 18 months of age are capable of reasoning about desires for food items (Repacholi & Gopnik, 1997) and other studies have demonstrated gender-stereotype knowledge for toys and objects in 18- to 24-month-old children (e.g., Poulin-Dubois et al., 2002; Serbin et al., 2001; Weinraub et al., 1984).

Emotional and behavioural responses during the object desire and stereotyping trials. As for the food request task, children's emotional and behavioural responses to the object desire and stereotyping trials were examined to determine whether their poor

Table 5

Mean Percentages of Feminine-Stereotyped Objects (and Standard Deviations) Given to the Male and Female Experimenters During the Desire Trials for the 20-Month-Old Children

Sex of Child	Male Experimenter		Female Experimenter	
	M	*SD*	*M*	*SD*
Boys (n = 16)				
Gender-consistent group	59.38	15.71	65.63	22.90
Gender-inconsistent group	40.63	35.20	50.00	29.88
Across groups	50.00	28.05	57.81	26.95
Girls (n = 16)				
Gender-consistent group	53 13	36.44	40.63	26.52
Gender-inconsistent group	54.17	23.57	40.63	22.90
Across groups	53.65	29.65	40.63	23.94
Total (n = 32)				
Gender-consistent group	56.25	27.30	53.13	27.20
Gender-inconsistent group	47.40	29.77	45.31	26.17
Across groups	51.82	28.46	49.22	26.55

Note. There were equal numbers of girls and boys in the gender-consistent and gender-inconsistent groups.

Table 6

Mean Percentages of Feminine-Stereotyped Objects (and Standard Deviations) Given to the Male and Female Experimenters During the Desire Trials for the 24-Month-Old Children

Sex of Child	Male Experimenter		Female Experimenter	
	M	*SD*	*M*	*SD*
Boys (n = 16)				
Gender-consistent group	43.75	17.68	53.13	24.78
Gender-inconsistent group	65.63	22.90	40.63	26.52
Across groups	54.69	22.76	46.88	25.62
Girls (n = 16)				
Gender-consistent group	48.96	31.95	37.50	35.36
Gender-inconsistent group	56.25	22.16	46.88	24.78
Across groups	52.60	26.83	42.19	29.89
Total (n = 32)				
Gender-consistent group	46.35	25.09	45.31	30.58
Gender-inconsistent group	60.94	22.30	43.75	25.00
Across groups	53.65	24.50	44.53	27.48

Note. There were equal numbers of girls and boys in the gender-consistent and gender-inconsistent groups.

performance on these tasks may have been due to their attempts to tease the experimenter by offering the incorrect object (see Appendix J for analyses). No differences between the two genders, age groups, or experimental groups were found in the level of "teasing" responses displayed during the desire and stereotyping trials. Moreover, correlation analyses (see Appendix J) indicated that the emotional and behavioural responses displayed by the children during the object desire and stereotyping trials were unrelated to their performance, suggesting that the presence of teasing behaviours cannot account for their poor performance on these tasks.

Relationship between children's understanding of desires for food items and their understanding of desires for nonfood items. To determine whether children's understanding of desires for food items was related to their understanding of desires for nonfood items, children's performance on the food desire trials was correlated with their performance on the object desire trials. For the food request task, children were given a score of 1 if they correctly offered the experimenter the desired food on both trials, and a score of 0 if they offered the experimenter the desired food on one of the two trials or on neither trial. The point-biserial correlation computed between the children's performance on food request task and the percentage of desired objects offered to the experimenter on the object request task approached statistical significance, $r(60) = .22$, $p < .10$. Children who showed a greater understanding of desires on the food request task were also somewhat more likely to correctly infer the experimenters' desires for nonfood items on the object request task. Thus, children's understanding of others' desires for food items and their understanding of others' desires for nonfood items appear to be related, although the relationship is fairly small. This finding also suggests that children treated these two tasks as similar, relying more heavily on the affective cues provided by the experimenter than on gender cues to make their desire inferences, even when gender-typed objects were used.

Relationship between children's object preferences and their ability to infer the experimenter's desires. Previous research has suggested that children may have difficulty reasoning about other people's desires when they themselves have a strong and conflicting desire (Moore et al., 1995). Thus, it is possible that children in the present study who displayed a strong preference for the masculine objects, for example, may have had difficulty understanding that the experimenter could desire feminine objects. To test this possibility, a 2 (object preference: masculine vs. feminine objects) x 2 (desire condition: masculine vs. feminine objects) mixed model ANOVA was conducted on the percentage of desired objects given to the experimenters on the object desire trials, for both experimental groups (gender consistent and gender inconsistent) combined (Appendix K, Table K1). Object preference served as the between-groups variable and condition as the within-groups variable. Children's preferences for the masculine and feminine objects during the Visit 1 baseline trial served as the preference measure. As the experimenters' desire expressions were found to influence the children's preferences when the desire trials were interspersed with the baseline trials, only the children who received the stereotyping trials first were included in the present analyses. Of these 32 children, those who displayed a preference for the feminine items on at least three of the four trials on Visit 1 were classified as having a feminine preference (n = 12; 4 boys, 8 girls), whereas those children who displayed a preference for the masculine toys on at least three of the four trials were classified as having a masculine preference (n = 9; 7 boys, 2 girls).

There was no significant condition main effect, nor Condition x Object Preference interaction, $F(1, 19) = 0.48$, *ns*, and $F(1, 19) = 2.38$, *ns*, respectively. This indicates that children's object preferences did not influence the percentage of desired feminine or masculine items given to the experimenters. Mean percentages of desired items given to the experimenters when feminine items were desired were 45.37% (SD = 26.72%) for children who preferred masculine objects versus 54.17% (SD = 35.09%) for those who

preferred feminine objects. Mean percentages of desired items given to the experimenters when masculine items were desired were 52.78% (*SD* = 26.35%) for children who preferred masculine objects and 34.72% (*SD* = 24.32%) for those who preferred feminine objects. Thus, children's object preferences did not interfere with their ability to understand the object preferences of the experimenters.

Gender-Concept Knowledge and Exposure to Gender Stereotyping in the Home

It was initially hypothesized that children with greater knowledge of, and exposure to, gender-typed roles, activities, and interests would be more likely to make gender-based inferences of others' desires, selecting the desired item to give to the experimenter when it was gender consistent, but not when it was gender inconsistent. In the gender-inconsistent condition, they were predicted to give the gender-appropriate, but undesired item. Thus, the 24-month-old children, who were expected to have greater knowledge of gender stereotypes than the 20-month-olds, were predicted to be more likely to make gender-based inferences of others' desires. However, no evidence of gender-stereotype knowledge nor gender-based desire reasoning was found for this age group.

To test whether gender knowledge and exposure to gender stereotyping were related to the ability to make gender-based desire inferences, correlations were computed between the children's scores on various gender-concept and gender-typing measures and the percentage of desired objects given to the experimenters in the gender-consistent and gender-inconsistent desire conditions. The gender-concept and gender-typing measures included (a) the gender-labelling task, which assessed the children's knowledge of gender labels; (b) the Sex-Typed Child Care and Household Tasks Questionnaire, which assessed their exposure to gender-typed activities in the home; (c) the Item Familiarity Questionnaire, which assessed their exposure to, and preference for, gender-typed toys and objects; and (d) the percentage of correct selections made on the object stereotyping trials, which provided a measure of their gender-stereotype knowledge. Descriptive

statistics and preliminary analyses are first presented for the gender-labelling task and the two questionnaires, following which correlations between each of the measures and the percentage of desired objects given to the experimenters in the gender-consistent and gender-inconsistent conditions are presented. In addition, correlations were computed between the number of correct item selections made on the stereotyping trials and the other gender-concept and gender-typing measures to determine the relationship between these variables.

Knowledge of gender labels. Children's performance on the gender-labelling task was examined to determine their level of gender-category knowledge. The percentage of correct targets selected was computed for the child and adult pictures separately, and for all pictures combined, averaging across both the child and adult pictures. An initial 2 (sex of child) x 2 (age group: 20 vs. 24 months) x 2 (picture type: child vs. adult pictures) mixed model ANOVA was conducted on the percentage of correct targets selected for the child and adult pictures to assess for age and sex effects (Appendix L, Table L1). Picture type was the within-groups variable, and sex of child and age group were the between-groups variables. Mean percentages of correct targets selected for the child and adult pictures and standard deviations are presented in Table 7. Significant main effects for age group and picture type emerged, which were qualified by a significant Age Group x Picture Type interaction, $F(1, 53) = 7.73$, $p < .01$. Simple effects analyses using the Bonferroni correction revealed that the 20- and 24-month-old children performed similarly on the child pictures, with both age groups correctly selecting 63% of the targets ($M = 63.27\%$, $SD = 18.73\%$; $M = 63.22\%$, $SD = 19.10\%$, respectively), $F(1, 53) = 0.00$, *ns*. In contrast, for the adult pictures, an age difference emerged, with the 24-month-olds (M = 87.93%, SD = 16.00%) outperforming the 20-month-olds (M = 70.00%, SD = 24.28%) in the percentage of correct selections made, $F(1, 53) = 10.61$, $p < .01$. No difference in boys' and girls' knowledge of gender labels was found.

Table 7

Mean Percentages of Correct Targets Selected (and Standard Deviations) for the Child and Adult Pictures on the Gender-Labelling Task, as a Function of Age Group and Sex of Child

Age Group	Child Targets		Adult Targets	
	M	*SD*	*M*	*SD*
20-month-olds				
Boys (*n* = 13)	65.51	16.21	70.77	25.93
Girls (*n* = 15)	61.33	21.04	69.33	23.64
Total (n = 28)	63.27	18.73	70.00	24.28
24-month-olds				
Boys (*n* = 14)	60.71	16.80	90.48	15.63
Girls (*n* = 15)	65.56	21.33	85.56	16.51
Total (n = 29)	63.22	19.10	87.93	16.00

To determine if the children's performance on the gender-labelling task was above chance, *t* tests were conducted comparing the mean percentage of correct selections made on the child, adult, and total scales against chance (50%) for the 20- and 24-month-old children separately. Both age groups performed significantly above chance on the child, adult, and total gender-labelling scales. Mean percentage scores, standard deviations, and *t* values for the gender-labelling task are presented in Table L2 of Appendix L. Finally, in order to compare the present results with previous research using a similar gender-labelling task (see Leinbach & Fagot, 1986), the percentages of children obtaining passing scores on the child versus adult gender-labelling scales were computed. As the probability of correctly selecting the target on each picture pair by chance alone was 50%, children had to correctly select the target picture on all six trials of the child- and adult-labelling tasks to demonstrate above chance responding at the .05 level of significance. Thus, for this latter analysis, only those children who gave responses for all of the child and adult pictures were included ($n = 49$). Using these criteria, 8.16% of the children obtained a passing score on the child-labelling task and 46.94% of the children obtained a passing score on the adult-labelling task. There were no significant differences between the number of 20- and 24-month-olds passing the child- (9.52% vs. 7.14%) and adult-labelling tasks (38.10% vs. 53.57%), $\chi^2(1, n = 49) = 0.09$, *ns*, and $\chi^2(1, n = 49) = 1.16$, *ns*, respectively.

The present results are consistent with data reported by Leinbach and Fagot (1986), using a similar task. In their study, fewer than 8% of the children younger than 26 months successfully passed the child-labelling task, whereas 55% passed the adult-labelling task. This offers some support that the current gender-labelling task is a valid measure of children's gender-category knowledge and suggests that the present sample has age-appropriate gender knowledge.

Exposure to gender stereotyping in the home. The Sex-Typed Child Care and Household Tasks Questionnaire was administered to the parents to assess to what extent

the children were exposed to gender stereotyping in the home. On this measure, lower scores (i.e., close to 1) represented high levels of gender-typed role division in the home and higher scores (i.e., close to 5) represented high levels of nontraditional or cross-gender-typed role division in the home. Parents' mean gender-typing scores on this questionnaire were 2.24 (range = 1.45 to 3.10, SD = 0.34) for the feminine tasks and 2.50 (range = 1.76 to 3.19, SD = 0.30) for the masculine tasks. Thus, traditionally feminine tasks were more likely to be performed in a gender-typed manner at home than were traditionally masculine tasks, as reported by the parents, $t(61) = -4.07$, $p < .001$. However, parents' mean gender-typing scores for both the feminine and masculine tasks were significantly below the midpoint, 3, which represented equal division of labour in the home on the traditionally gender-typed tasks, $t(62) = -17.86$, $p < .001$, and $t(61) = -12.90$, $p < .001$. Furthermore, these means were similar to those obtained by Serbin et al. (1993), using the same questionnaire, in their sample of 245 parents of school-aged children, (M = 2.03, SD = 0.46; and M = 2.51, SD = 0.40, for the feminine and masculine tasks, respectively). They were also consistent with a previous study conducted with 24-month-old children, using the same questionnaire, in which mean gender-typing scores of 2.19 and 2.78 were obtained for the feminine and masculine tasks, respectively (Beissel, 1998). These data confirm that the children in the present study were exposed to similar levels of gender stereotyping in their homes as has been reported for other samples of children.

Mean gender-typing scores were also averaged across the masculine and feminine tasks to create an overall gender-typing score. The overall mean gender-typing score was 2.39 (range = 1.93 to 3.00, SD = 0.20), which was significantly below the midpoint 3, representing an equal division of labour at home, $t(62) = -23.86$, $p < .001$. A 2 (sex of child) x 2 (age group: 20 vs. 24 months) between-groups ANOVA conducted on the overall gender-typing scores revealed no age or sex differences (see Appendix L, Table L3). Overall mean gender-typing scores were 2.33 (SD = 0.18) and 2.43 (SD = 0.26) for

the 20-month-old boys and girls, respectively. Corresponding mean gender-typing scores for the 24-month-old boys and girls were 2.40 (SD = 0.22) and 2.41 (SD = 0.13), respectively. Thus, the results indicated that the present sample of 20- and 24-month-old children were exposed to mild gender stereotyping of adult household and child-care activities in their homes.

Exposure to gender-typed objects. The Item Familiarity Questionnaire assessed to what extent children played with, or were exposed to, the toys and other objects used in the present study. Most of the children were familiar with the items used. Over 70% of the children had dolls, toy cars, purses, and necklaces at home (either their own, siblings', or parents') or at daycare, and 60.94% either had a tie at home or observed their father wearing a tie. The children were less familiar with baseball gloves, however, with only 39.06% of the children having baseball gloves at home or at daycare. For the blue and pink stuffed bears, children's exposure to blue and pink colours was assessed (i.e., exposure to blue or pink clothing, toys, or bedroom decor), since few children were expected to have blue and pink bears specifically. Almost all of the children (over 90%) had some blue clothing, toys, or bedroom decor, whereas half of the children (53.13%) had some pink clothing, toys, or bedroom decor. As expected, girls were significantly more likely to play with feminine toys, such as dolls and purses, and to have pink clothing, toys, and/or bedroom decor, whereas boys were significantly more likely to play with cars and baseball gloves, and to have blue toys, clothing, and/or bedroom decor (see Appendix L, Table L4, for means, standard deviations, and t values for parents' ratings of the frequencies with which the children played with, or were exposed to, each of these toys or objects). Furthermore, mothers or other females in the household were significantly more likely to be observed wearing necklaces (M = 3.28, SD = 1.57, on a scale from 1 to 5, with 1 = *never* and 5 = *always*), than were fathers or other males (M = 2.06, SD = 1.71), $t(63) = 5.69$, $p < .001$. Finally, 53.13% of the children observed their fathers wearing ties at least sometimes. Thus, the present results suggest that the

children were generally familiar with the items used in the study, with the possible exception of the baseball glove. Furthermore, these data confirm that these items are gender stereotyped, with females more likely to have, play with, or wear the feminine items, and males, to have, play with, or wear the masculine items.

A total gender-typing score was calculated for each child based on the parent's ratings of the frequency with which the child played with gender-appropriate toys, or observed gender stereotyping of the items used in the study at home. For each item pair, children were given a score of 1 (*gender-typed*) if they played with, or wore, the gender-appropriate item more frequently than the gender-inappropriate item. Conversely, they were given a score of 0 (*non-gender-typed*) if they played with, or wore, the gender-inappropriate item more than the gender-appropriate item, or if they played with, or wore, both items equally. For the doll/car and purse/baseball glove pairs, children's mean gender-typing scores were .63 (SD = .49) and .55 (SD = .50), respectively. For the pink and blue bears, the child's exposure to each colour was determined by averaging the parent's ratings of the frequency with which the child wore pink or blue clothing, their ratings of the amount of pink or blue toys the child had at home, and their ratings for the child's bedroom decor (1 = *does not have pink blue room*; 5 = *has pink/blue room*). Children's mean gender-typing score for pink/blue was .69 (SD = .47). For the necklace and tie pair, children were given a score of 1 (*gender-typed*) for each of the following: (a) if they played with or wore the gender-appropriate item more than the gender-inappropriate item, (b) if they observed their mothers or another female in the household wearing necklaces more than they observed their fathers or another male wearing necklaces, and (c) if they observed their fathers wearing ties at least sometimes. These three gender-typing scores were then averaged to form a single gender-typing score for the necklace/tie pair (M = .60, SD = .35). The gender-typing scores for each of the item pairs were averaged to create a total gender-typing score for each child. This score ranged from 0 to 1, with a score of 0 indicating *no gender-typed preferences or gender-typed*

exposure to the items, and a score of 1 indicating a *high level of gender-typed preferences or gender-typed exposure* to the items. Children's mean gender-typing score across all four item pairs was .62 (*SD* = .21), indicating a moderate degree of exposure to, or preference for, gender-typed objects. A 2 (sex of child) x 2 (age group: 20 vs. 24 months) between-groups ANOVA conducted on children's mean gender-typing scores across all four pairs revealed no age or sex differences (see Appendix L, Table L5). At 20 months of age, mean gender-typing scores were .58 (*SD* = .15) for the boys and .60 (*SD* = .26) for the girls. At 24 months of age, similar values were obtained, with mean gender-typing scores of .65 (*SD* = .21) for the boys and .63 (*SD* = .24) for the girls. Thus, children of both ages and sexes showed similar levels of exposure to, or preference for, gender-typed objects.

Relationship between children's knowledge of, or exposure to, gender stereotypes and their understanding of others' desires for gender-typed objects. Children's total scores on the gender-labelling task, parents' total gender-typing score on the Sex-Typed Child Care and Household Tasks Questionnaire, and children's total gender-typing score on the Item Familiarity Questionnaire were correlated with the percentage of desired objects given to the experimenters for the gender-consistent and gender-inconsistent groups separately (see Table 8). For the gender-inconsistent group, a significant positive correlation was found between children's total gender-labelling scores and the percentage of desired objects given to the experimenters, $r(30) = .42$, $p < .05$. Contrary to prediction, children with greater knowledge of gender labels were more likely to give the experimenters the desired gender-inconsistent object than were children with less knowledge of gender labels. That is, children in the gender-inconsistent group with greater gender knowledge were less likely to make gender-based inferences of the experimenter's desires: instead, these children followed the experimenter's expressed desires even though the experimenter desired gender-inconsistent objects. Children's knowledge of gender labels was not found to be related to their performance on the desire

Table 8

Pearson-Product Moment Correlations Computed Between Children's Scores on the Gender-Concept and Gender-Typing Measures and the Percentage of Desired Objects Given to the Experimenters on the Object Desire Trials

Measure	*n*	*r*
Gender-Consistent Group		
Gender-Labelling Task, total score	27	- .05
SCCHT Questionnaire, total gender-typing score	32	- .21
Item Familiarity Questionnaire, total gender-typing score	32	.00
Stereotyping Trials, % correct given	32	.27
Gender-Inconsistent Group		
Gender-Labelling Task, total score	30	.42 *
SCCHT Questionnaire, total gender-typing score	31	.06
Item Familiarity Questionnaire, total gender-typing score	32	.03
Stereotyping Trials, % correct given	32	.22

Note. SCCHT Questionnaire = Sex-Typed Child Care and Household Tasks Questionnaire.

* $p < .05$.

task for the gender-consistent group, however, $r(27) = -.05$, *ns*. Thus, children with greater gender-category knowledge were not more likely to offer the experimenter the desired object when it was gender consistent. No relationships were found between the measures of children's exposure to gender-stereotyped activities (i.e., parent's total gender-typing score on the Sex-Typed Child Care and Household Tasks Questionnaire) or objects (i.e., children's total gender-typing score on the Item Familiarity Questionnaire) and their performance on the object desire task for either experimental group. Finally, children's knowledge of gender stereotypes, as assessed by the percentage of correct selections made on the stereotyping trials, was correlated with their performance on the desire trials for the two experimental groups separately. Neither of the correlations were significant, indicating that children with greater gender-stereotype knowledge were not more likely to make gender-based inferences of the experimenter's desires, nor were they more likely to follow the experimenter's expressed desires for the objects.

Relationship between children's knowledge of the gender stereotyping of familiar objects and their knowledge of gender labels and exposure to gender stereotyping in the home. To assess whether children's knowledge of gender labels and their exposure to gender stereotyping in the home were related to their knowledge of gender stereotypes for objects, children's total gender-labelling score and their total gender-typing scores on the two questionnaires were correlated with the percentage of correct item selections made during the stereotyping trials. These correlations were computed for the whole sample ($N = 64$), across the two experimental groups. A significant positive correlation was found between children's total gender-labelling score and the number of correct selections made on the stereotyping trials, $r(57) = .27$, $p < .05$. Children with greater knowledge of gender labels demonstrated greater knowledge of object gender stereotypes. This offers some validation that the stereotyping trials were measuring gender-concept knowledge. No relationship was found, however, between children's knowledge of object gender stereotypes and their exposure to, or preference for, gender-typed objects, $r(64) = -.08$, *ns*.

nor between their object stereotype knowledge and their exposure to gender-typed activities in the home. $r(63) = -.04$, *ns*.

Understanding of Others' Desires for Gender-Typed Objects in Those Children Who Displayed a Moderate Level of Understanding of Others' Desires for Food Items and Adequate Gender-Category Knowledge.

The results of the present study failed to demonstrate evidence of desire understanding for food items or objects in 20- and 24-month-old children. This contrasts with Repacholi and Gopnik's (1997) findings of desire understanding for food items in children as young as 18 months of age. It was expected that the food request task would be easier than the object request task as the children needed only to infer the experimenter's food desires from his or her affective displays. In the object request task, however, the gender stereotyping of the objects, the gender of the experimenter, and the experimenter's affective displays could all serve as potential sources of information about the experimenter's preferences. Given the relatively poor performance of the present sample on the food request task, a subset of the children who met minimum performance requirements on the food request task were selected for further analyses on the object request task. Specifically, those children who correctly offered the experimenter the desired food on a minimum of one of the two trials were retained for further analyses. Moreover, given that the experimental manipulation on the object request task required that the children be able to differentiate the male and female experimenters, only children who demonstrated adequate gender-category knowledge on the gender-labelling task were included. Adequate performance on the gender-labelling task was defined as obtaining a total score (on the child and adult pictures combined) above 60%. The selection criteria used were fairly lenient, but this was necessary to ensure an adequate sample size for analyses. A total of 31 children (11 20-month-olds, 20 24-month-olds; 14 girls, 17 boys) were selected based on the above criteria. There were 16 children in the gender-consistent condition, and 15 children in the gender-inconsistent condition. Due to the relatively

small sample size, only the primary variables of interest were included in the analyses. Consequently, the effects of task and condition order on children's performance during the desire and stereotyping trials were not assessed. However, these variables were not found to influence children's performance on the experimental tasks for the whole sample, and therefore were not expected to significantly affect children's performance in the subsample.

Knowledge of the gender stereotyping of familiar objects in the subsample. The percentage of gender-stereotyped objects correctly given to the male and female experimenters during the stereotyping trials was subjected to a 2 (visit) x 2 (age group: 20 vs. 24 months) x 2 (group: gender consistent vs. gender inconsistent) mixed model ANOVA to assess for visit, age, and group effects (Appendix M, Tables M1 and M2). Visit was the within-groups variable, and age group and group were the between-groups variables. A significant effect of visit was found, with children making a higher percentage of correct object selections on their first visit (M = 63.71%, SD = 24.87%) compared to their second (M = 47.58%, SD = 23.09%), $F(1, 27) = 5.32$, $p < .05$. Thus, subsequent analyses were conducted for each visit separately. There were no other significant main effects or interactions.

Children's gender-stereotype knowledge was assessed by conducting a 2 (sex of child) x 2 (condition: male vs. female stereotyping) between-groups ANOVA on the percentage of feminine items given to the male and female experimenter on the stereotyping trials for Visit 1 and Visit 2 separately (see Figure 4). Means and standard deviations are presented in Appendix M, Table M3. Knowledge of gender stereotypes would be demonstrated by a significant main effect of condition. For the Visit 1 stereotyping trials (Appendix M, Table M4), the ANOVA yielded a significant condition main effect, $F(1, 27) = 8.36$, $p < .01$. As predicted, children gave a significantly greater proportion of feminine-typed items to the female experimenter (M = 66.07%, SD = 18.62%) than to the male experimenter (M = 38.24%, SD = 29.47%). Therefore, the

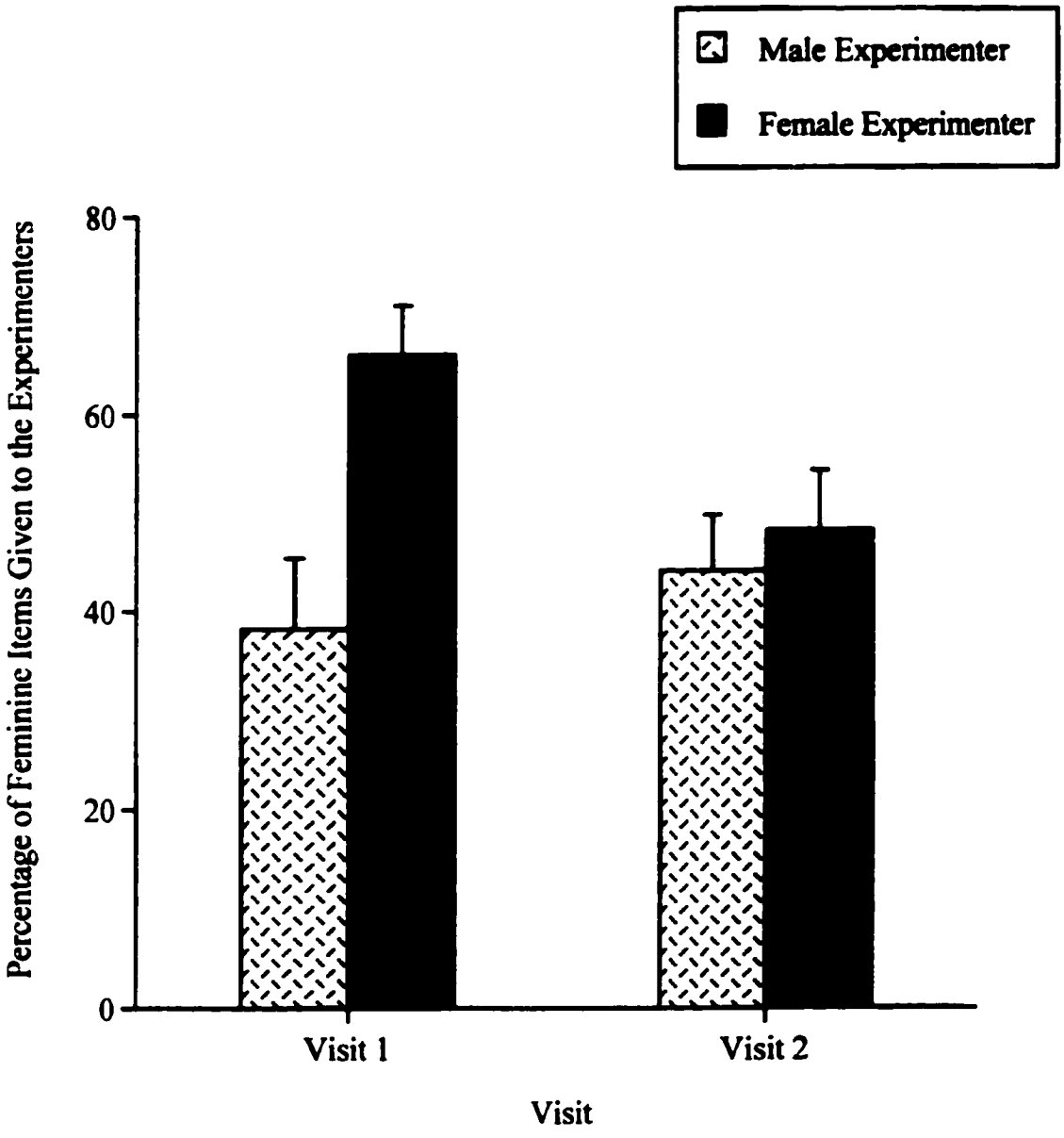


Figure 4. Percentage of feminine items given to the male and female experimenters during the stereotyping trials, as a function of visit, for the $n = 31$ subsample.

results suggest that these children had an understanding of gender stereotypes for objects, as they correctly offered the female experimenter a higher proportion of feminine-typed objects, and the male experimenter a higher proportion of masculine-typed objects. The condition main effect had an associated effect size of $\eta^2 = .24$, indicating that this effect was quite large and robust (Cohen, 1988). For the Visit 2 stereotyping trials (Appendix M, Table M5), no significant main effects or interactions emerged. On their second visit, the children gave similar proportions of feminine-typed objects to the female ($M = 48.21\%$, $SD = 22.92\%$) and male experimenters ($M = 44.12\%$, $SD = 22.34\%$), $F(1, 27) = 0.11$, *ns*.

Understanding of others' desires for gender-typed objects in the subsample. A preliminary 2 (visit) x 2 (age group: 20 vs. 24 months) mixed model ANOVA was first conducted on the percentage of desired objects given to the experimenters to assess for possible visit and age effects (Appendix M, Tables M6 and M7). Age group served as the between-groups variable, and visit as the within-groups variable. No significant main effects or interactions were found.

Children's desire understanding was then assessed by conducting a 2 (sex of child) x 2 (group: gender consistent vs. gender inconsistent) x 2 (condition: male vs. female desire) mixed model ANOVA on the percentage of feminine-typed objects given to the male and female experimenters during the object desire trials (Appendix M, Tables M8 and M9). An understanding of desires would be demonstrated by a significant Condition x Group interaction. This analysis revealed a significant condition main effect, $F(1, 27) = 8.47$, $p < .01$, which was qualified by a significant Condition x Group interaction, $F(1, 27) = 6.03$, $p < .05$, and a significant Sex of Child x Condition x Group interaction, $F(1, 27) = 5.54$, $p < .05$. To determine the source of the Sex of Child x Condition x Group interaction, separate 2 (condition: male vs. female desire) x 2 (group: gender consistent vs. gender inconsistent) ANOVAs were conducted for each sex (see Figure 5).

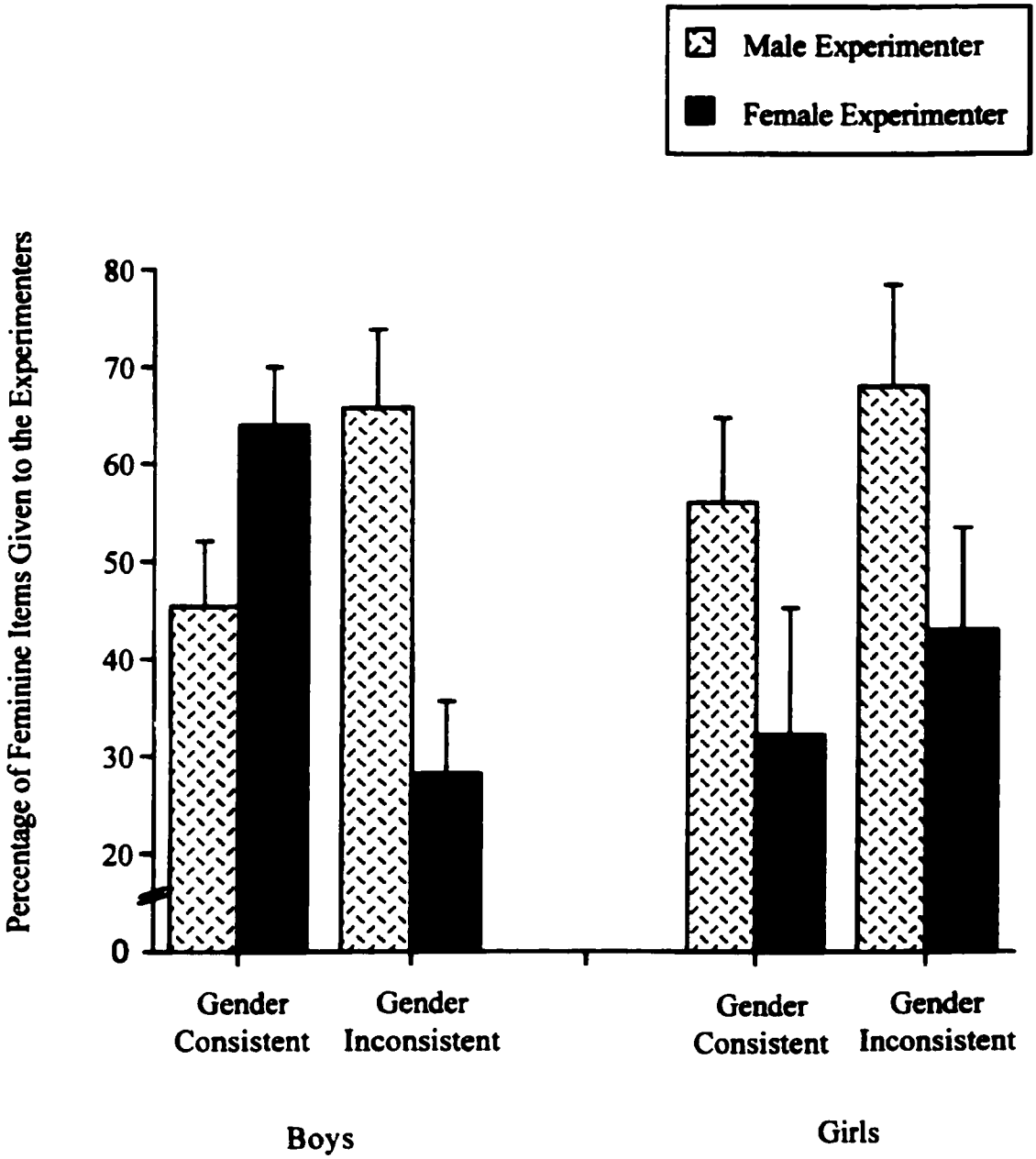


Figure 5. Percentage of feminine items given to the male and female experimenters during the desire trials, as a function of sex of child and experimental group, for the $n = 31$ subsample.

For the boys (Appendix M, Table M10), the ANOVA revealed a significant Condition x Group interaction, $F(1, 15) = 23.62$, $p < .001$. Simple effects analyses, using the Bonferroni correction, revealed that boys in both experimental groups followed the experimenters' expressed desires, giving the experimenters the desired gender-appropriate objects in the gender-consistent condition and the desired gender-inappropriate objects in the gender-inconsistent condition. Specifically, in the gender-consistent condition, the boys gave a significantly higher proportion of feminine objects to the female experimenter ($M = 63.89\%$, $SD = 18.16\%$) than to the male experimenter ($M = 45.37\%$, $SD = 20.03\%$), $F(1, 15) = 5.49$, $p < .05$. Conversely, in the gender-inconsistent condition, the boys gave a significantly higher proportion of feminine objects to the male experimenter ($M = 65.63\%$, $SD = 22.90\%$) than to the female experimenter ($M = 28.13\%$, $SD = 20.86\%$), $F(1, 15) = 19.99$, $p < .001$. The associated effect sizes were $\eta^2 = .27$ and $\eta^2 = .57$, respectively, indicating that the differences between the male and female conditions for both experimental groups were quite large and robust (Cohen, 1988).

For the girls (Appendix M, Table M11), only a significant condition main effect emerged, $F(1, 12) = 5.09$, $p < .05$. There was no significant Condition x Group interaction, $F(1, 12) = 0.00$, *ns*, indicating that the girls did not follow the experimenters' expressed desires for the objects. Interestingly, girls in both experimental groups gave the experimenters a higher proportion of the gender-inappropriate objects. That is, the girls gave a significantly higher proportion of feminine objects to the male experimenter ($M = 61.91\%$, $SD = 25.26\%$) than to the female experimenter ($M = 37.50\%$, $SD = 30.62\%$), and conversely, a significantly higher proportion of masculine objects to the female experimenter ($M = 62.50\%$, $SD = 30.62\%$) than to the male experimenter ($M = 38.09\%$, $SD = 25.26\%$) in both desire conditions. The associated effect size was $\eta^2 = .30$, indicating that this effect was quite large (Cohen, 1988).

Understanding of other people's desires for food items. To determine if the observed sex difference in the children's performance on the object desire trials could be due to differences in boys' and girls' basic desire understanding, their performance was compared on the food request task. A chi-square analysis revealed no significant difference between the proportion of boys and the proportion of girls that offered the experimenter the desired food across the two trials, $\chi^2(1, n = 30) = 0.74$, *ns*. For the boys, 56.25% offered the experimenter the desired food on one of the two trials, and 43.75% offered the desired food on both trials. For the girls, 71.43% offered the experimenter the desired food on one of the two trials, and 28.57% offered the desired food on both trials. Thus, no sex differences in the ability to make inferences about others' desires based on affective cues were found. This suggests that differences in basic desire understanding cannot account for the observed sex difference in children's performance on the object desire trials.

Gender knowledge and exposure to gender stereotyping in the home. To determine whether the observed sex difference in the children's performance on the desire trials could be due to differences between the boys and the girls in gender knowledge or exposure to gender stereotyping in the home, a 2 (sex of child) x 2 (group: gender consistent vs. gender inconsistent) between-groups MANOVA was performed (see Appendix N). No significant main effects or interactions emerged from this analysis (see Appendix N, Table N1), indicating that the boys and the girls in both experimental groups had similar levels of gender knowledge and exposure to gender stereotyping in the home. Thus, differences in gender knowledge or exposure to gender stereotyping could not account for the observed differences in the children's performance on the object desire trials.

Emotional and behavioural responses to the object desire trials for subsample. Children's emotional and behavioural responses to the object desire trials were also examined to determine if sex or group differences on these measures could account for

the observed pattern of results. A difference between the two experimental groups was found for the amount of amusement (smiling or laughing) displayed by the children. Because the experimenters expressed desires for the gender-inappropriate objects in the gender-inconsistent condition, it was initially hypothesized that some children might respond to this perceived incongruity or absurdity by smiling or laughing. In the gender-consistent condition, however, the experimenters expressed desires for the gender-appropriate objects. As these desires were not expected to be perceived as surprising or absurd, few children were predicted to show signs of amusement in this condition. Unexpectedly, children of both sexes were found to smile and laugh more frequently during the gender-consistent condition than during the gender-inconsistent condition. Five of the 7 girls in the gender-consistent condition (71.43%) showed amusement when responding to the experimenters' request for an object on the majority of the trials (75% or more), compared to only 2 of the 7 girls (28.57%) in the gender-inconsistent group. Similarly, 7 of the 9 boys in the gender-consistent condition (77.78%) laughed or smiled when responding to the experimenters' request for an object on the majority of trials, compared to 2 of the 9 boys (22.22%) in the gender-inconsistent condition. Moreover, for both sexes, the presence of laughing or smiling following the experimenters' request for an object was negatively correlated with the percentage of desired objects given to the experimenters across the experimental conditions, $r(14) = -.56$, $p < .05$, and $r(17) = -.49$, $p < .05$, for the girls and boys respectively. That is, children who displayed high levels of smiling or laughter during the object desire trials were more likely to offer the experimenters the undesired objects. These results suggest that the presence of smiling or laughing when responding to the experimenters' request may be indicative of teasing behaviours, and indicates that these responses were more frequent in the gender-consistent condition. However, examination of the children's emotional responses to the task failed to find any differences between the boys and the girls that could account for their different performance on the object desire trials.

To further assess whether teasing behaviours could account for the observed gender difference, an informal examination of the children's performance on the desire and stereotyping trials was carried out. This analysis revealed that two girls in the gender-consistent group consistently selected the incorrect object to give to the experimenters across conditions. Specifically, one girl gave the incorrect object on the majority of trials (75% or more) for both the male desire and female desire conditions, and for both the male and female stereotyping conditions. A second girl gave the incorrect object on the majority of trials (75% or more) for both the male desire and female desire conditions. This pattern of consistently giving the incorrect object was not observed for any of the girls in the gender-inconsistent group, nor for any of the boys in either the gender-consistent or gender-inconsistent groups. Thus, some girls in the gender-consistent condition may have attempted to tease the experimenters by offering them the undesired or incorrect object. However, further research is needed to test this hypothesis.

Summary of children's performance on the object request task desire trials for subsample. To summarize, boys in the above subsample were found to demonstrate an understanding of desires for gender-typed objects, correctly offering the experimenters the desired objects in both the gender-consistent and gender-inconsistent conditions. The girls, in contrast, did not appear to follow the experimenters' desires, nor did they demonstrate gender-based desire reasoning. Instead, they gave the male and female experimenters the gender-inconsistent objects in both of the experimental conditions, despite showing an adequate understanding of the gender stereotyping of the objects on the stereotyping trials and normal gender-category knowledge on the gender-labelling task. One possible explanation for this unexpected finding was that some of the girls in the gender-consistent condition might have attempted to tease the experimenter by offering the undesired, gender-inappropriate objects.

Discussion

The results of the present study confirm that an understanding of other people's desires is emerging during the second year of life. Previous research has suggested that children as young as 18 months of age demonstrate an understanding of others' desires for food items (Repacholi & Gopnik, 1997). Using an adaption of Repacholi and Gopnik's food request task, the current study extends these findings by suggesting that an understanding of desires for gender-typed objects appears to develop in some children during the same period or shortly thereafter. In this paradigm, the experimenter facially and verbally expressed pleasure in response to one of two gender-typed objects, and disgust in response to the other. The experimenter then requested that the child give him or her one object. A subset of the children in the current sample, those who demonstrated adequate gender-category knowledge and a basic understanding of other people's desires for food, successfully offered the experimenter the desired object on the majority of trials. Furthermore, these children showed knowledge of the gender stereotyping of the objects. During the Visit 1 stereotyping trials, they gave the male experimenter a higher proportion of masculine-typed objects and the female experimenter, a higher proportion of feminine-typed objects. Although significant knowledge of gender stereotypes was not found during Visit 2, this was likely due to task order effects.

Use of Gender Information in Making Inferences about Other People's Desires

A second aim of the present research was to determine whether children use information about a person's gender to make inferences about his or her desires for gender-typed objects. Thus, two experimental conditions were included. In the gender-consistent condition, the experimenter expressed a preference for the gender-appropriate objects on all trials (e.g., the female experimenter expressed a preference for the feminine-typed objects). In the gender-inconsistent condition, the experimenter expressed a preference for the gender-inappropriate objects on all trials (e.g, the female experimenter expressed a preference for the masculine-typed objects). If the children used

information about the experimenter's gender to make inferences about his or her desires, they were expected to be more likely to offer him or her the desired object when it was gender consistent rather than gender inconsistent. Although the children in the subsample demonstrated an understanding of the gender stereotyping of the objects, they did not select the gender-appropriate objects to give to the experimenter. Rather, boys in both experimental conditions gave the experimenter the object he or she desired regardless of the object's gender stereotyping. Thus, for the boys, the individuating information provided by the experimenter's affective displays was more salient, or considered more relevant, than the more abstract category-driven gender information.

The present results were less clear regarding the girls' understanding of desires. The girls in the subsample appeared to follow the experimenter's desires in the gender-inconsistent condition, giving the experimenter the desired gender-inappropriate object on the majority of trials. However, during the gender-consistent condition, they also gave the experimenter the gender-inappropriate objects, which were not desired by the experimenter. Thus, the girls appeared to give the experimenters the gender-inappropriate objects in both conditions, regardless of which objects were desired. It may be that the girls found it humourous to violate gender stereotypes by offering the experimenters cross-gender-typed objects. Alternatively, as suggested previously, the girls may have followed the experimenters' desires in the gender-inconsistent condition. However, they may have "teased" the experimenters in the gender-consistent condition, by offering the undesired, gender-inappropriate objects. The informal observation that two girls in the gender-consistent group consistently gave the experimenter the incorrect item across the majority of trials lends some support to this interpretation.

The observed gender difference in children's responses during the object desire trials cannot be attributed to differences in boys' and girls' understanding of desires. Boys and girls performed similarly on the food request task in the current study. Moreover, gender differences in children's desire understanding have not generally been reported in

the literature (Hadwin & Perner, 1991; Repacholi & Gopnik, 1997; Wellman & Banerjee, 1991; Wellman & Woolley, 1990). An alternative possibility is that the different response patterns observed for boys versus girls may be the result of differences in gender knowledge. Some research suggests that girls have more advanced gender knowledge than boys, with girls demonstrating earlier knowledge of gender labels (Poulin-Dubois et al., 1998; Thompson, 1975; Weinraub et al., 1984) and toy gender stereotypes (Serbin et al., 2001) than boys. If girls were more advanced in their gender knowledge, they may have found it more humourous to violate their gender stereotypes by offering the experimenter the gender-inappropriate objects. No gender differences in children's gender knowledge or exposure to gender stereotyping were found in the current study. Thus, differences in boys' and girls' gender knowledge cannot account for the present findings. Further research is necessary to clarify the present gender difference in children's reasoning about males' and females' desires for gender-typed objects.

Contrary to prediction, no age differences in children's tendency to make gender-based desire inferences were found, with children performing similarly on the desire tasks at both 20- and 24-months of age. Moreover, both age groups were found to be comparable in their level of gender knowledge. The 20- and 24-month-old children showed similar knowledge of the gender stereotyping of the objects on the stereotyping trials. Additionally, few or no age differences were found in their knowledge of gender labels, exposure to gender-typed activities in their homes, and exposure to, and preference for, gender-typed objects. Perhaps, gender-based inferences of other people's desires could have been demonstrated in an older sample of children (e.g., 36 months), as gender-related associations may increase in strength and saliency with age.

Finally, no effect of the gender of the experimenter was found. It was initially hypothesized that a male experimenter expressing desires for feminine objects may be perceived as more incongruous than a female experimenter expressing desires for masculine objects. This is because previous research has suggested that the male gender

role tends to be more rigid and traditional (Archer, 1984; Hort et al., 1990). However, in the present study, the children were equally likely to give the male experimenter the desired gender-inappropriate objects as the female experimenter. Rather than basing their item selections on the gender of the experimenter, children in the gender-inconsistent group appeared to base their desire inferences on the affective cues provided.

Implicit Understanding of Desires and Gender Stereotypes

An implicit understanding of other people's desires was also demonstrated by a subset of the children in the overall sample who received their object baseline trials interspersed with the desire trials. Children in the gender-inconsistent group, who witnessed the experimenters expressing desires for the cross-gender-typed items, showed little gender stereotyping in their object preferences on the baseline trials. These children were more likely to display equal preferences for the masculine and feminine objects, or to prefer the cross-gender-typed objects, compared to the children who observed the experimenters expressing desires for the gender-appropriate items (i.e., in the gender-consistent condition). Thus, these children appeared to pattern their own object play after the preferences displayed by the experimenters. This suggests that these children were able to infer the experimenters' desires from their affective displays. In contrast, children in the gender-inconsistent group who did not observe the experimenter expressing desires for the objects until after they had received all four baseline trials did not show a decrease in their gender-typed object play. These children were similar to those in the gender-consistent group in terms of the amount of gender-typed play displayed. This lends further support to the notion that children's play preferences were influenced by the experimenter's desire expressions for the gender-typed objects, and were not the result of artifactual sampling differences between the two experimental groups. Children who observed the experimenters expressing gender-typed desires did not become more gender stereotyped in their object preferences. This may be because the experimenters' desire

expressions for the gender-typed objects were familiar and predictable and therefore, not as salient as the experimenters' desire expressions for the gender-inconsistent objects.

The experimenters' desire expressions were not expected to influence the children's preferences during the baseline trials. This was because the baseline trial for a particular item pair preceded the experimental trial in which the experimenter expressed desires for the items. Thus, the experimenter's desire expressions could only influence children's preferences on subsequent trials, involving different pairs of items. This means that the children had to infer the experimenter's desires for one of the two objects based on his or her affective display. They then had to generalize the preferences they observed for that item pair to the set of masculine- or feminine-typed objects, and use this information to pattern their own play. For example, children had to infer that the female experimenter desired cross-gender-typed objects when she displayed a preference for the car over the doll. This knowledge apparently led the children to also play more with the cross-gender-typed objects (i.e., on a subsequent trial, girls played more with the baseball glove and boys with the purse). The above process is a complicated one and suggests that children are capable of quite sophisticated reasoning both about others' desires and about gender stereotypes.

The finding that the experimenters' expressed preferences for the objects influenced the children's object preferences was also unexpected as no such effect was found for food preferences in the Repacholi and Gopnik (1997) study. In their study, the infants' preferences for the broccoli and crackers were measured both before, and shortly after, they observed the experimenter expressing a desire for one of the two foods. No change in food preferences was observed for the majority of the infants (96%) over this time period. Furthermore, social referencing studies have indicated that children's responses to familiar, unambiguous items, are not influenced by other people's emotional reactions to them (e.g., Gunnar & Stone, 1984). In contrast, the influence of child or adult models on children's play behaviours has been reported in the gender literature. Most of

these studies have focused on the effect of labelling novel or gender-neutral toys or activities as "for boys" or "for girls" on children's subsequent preferences for, and engagement with, the toys or activities. These studies have found that labelling novel or gender-neutral toys or activities as gender appropriate, or observing same-sex children engaged with the activity or toy, encourages children's play with the toy or activity. Conversely, labelling the toy or activity as cross-gender-typed, or observing opposite-sex children engaged with the activity or toy, inhibits children's play with the toy or activity (Shell & Eisenberg, 1990; Thompson, 1975). However, it has generally been thought that once children have acquired knowledge of gender stereotypes, modelling of gender-inconsistent play behaviours by peers or other models would be unlikely to produce change (Martin & Halverson, 1981; Perry & Bussey, 1979; Ross & Ross, 1972). Katz and Walsh (1991), for example, found that observing peers engage in gender-inconsistent behaviours produced limited change in 7- and 10-year-old children's behaviour, unless the peer's behaviour had been reinforced. In contrast, a study by Serbin, Connor, and Iler (1979) found a decrease in gender-typed toy play in 3- and 4-year-old children when a set of dolls and a set of cars were introduced in a nonstereotyped manner in a classroom setting, with both boys and girls called on to demonstrate each set of toys. When those same toys were introduced to a different group of children with references to the traditional stereotyping of the toys, and "gender-appropriate" children were called on to demonstrate each set, the children were found to make gender-typed toy selections. Thus, in some cases, modelling of gender-inconsistent behaviours can produce a decrease in gender-typed play, as was observed in the present study, even if familiar, gender-stereotyped items or activities are used. It is possible that gender-typed knowledge and behaviours may develop independently (Blakemore et al., 1979; Perry et al., 1984; Weinraub et al., 1984), with factors other than gender-typed knowledge also influencing the development of play behaviours and preferences.

Limitations of the Present Research and Methodological Issues

One limitation of the present research was the failure to demonstrate a clear understanding of desires for both food and gender-typed objects in the overall sample. The relatively poor performance exhibited by both the 20- and 24-month-old children on the food request task was unexpected. In a previous study by Repacholi and Gopnik (1997), the vast majority of 18-month-olds (i.e., 73%) were found to correctly offer the experimenter the desired food. This contrasts with the current study. Using the same paradigm, the children were at chance levels in correctly selecting the desired food to offer the experimenter across the two visits. The children's poor performance may have been partially explained by order and task effects. When only the subset of children who received the food request task first were considered, their performance indicated some understanding of the experimenters' desires for the food items, with 65% of the children offering the experimenter the desired food. This result was similar to the proportion of 18-month-old children found to offer the experimenter the desired food in the McKoy and Poulin-Dubois (1999) study. However, it failed to reach statistical significance, likely due in part to the small sample size. Similarly, on the object request task, analyses on the overall sample failed to demonstrate a significant understanding of others' desires for gender-typed objects. Moreover, no understanding of the gender stereotyping of the objects was evident.

One possible interpretation of these null results is that desire understanding may be only emerging in children in the second year of life and therefore, may not be robust in this period of development. This interpretation is unlikely, however, given previous research which has documented an understanding of desires in children as early as 18 months of age. Repacholi and Gopnik (1997) found that the vast majority of 18-month-olds were capable of reasoning about others' desires for food items. Similarly, Tilden and Poulin-Dubois (Poulin-Dubois, 1999; Tilden, 2000) found that children as young as 18 months of age understood that people act to fulfill their desires. These children looked

longer at actions that were incongruent with a protagonist's expressed desire than at actions that were congruent with his or her desire. This suggests that the performed actions violated the children's expectations that people act in order to fulfill their desires. Moreover, studies of children's natural language have indicated that children begin to use desire terms such as "want" to explicitly refer to internal mental states as early as 18 months of age (Bartsch & Wellman, 1995). References to desires are well-established by 2 years of age, with these children demonstrating an understanding of the link between desires and actions, and desires and emotions in their conversations with others, as well as an understanding of the subjectivity of desires. An understanding of desire-dependent actions and emotions has also been demonstrated in children as young as 2½ by Wellman and Woolley (1990) using a story-based paradigm.

A more probable explanation may be that the food and object request tasks used in the present study are not as appropriate for assessing an understanding of desires in children 20 months of age or older. The paradigm was originally used with 14- and 18-month-old children in the Repacholi and Gopnik (1997) study. The paradigm was not found to be appropriate for 14-month-olds in their study, with the vast majority of these infants (68%) failing to respond to the experimenter's initial or subsequent requests for food. Furthermore, although an understanding of desires was successfully demonstrated in the 18-month-olds in this study, a substantial number of these children (30%) also failed to comply with the experimenter's requests. Thus, even in Repacholi and Gopnik's study, there were fairly high levels of noncompliance observed in the sample. In the present study, high levels of noncompliance were similarly observed, with 20% of the initial sample tested (18 of the 88 children tested) eliminated from the final analyses due to noncompliance or fussiness. The fact that a large proportion of the sample was unable to complete the study suggests that demands associated with this task may have been difficult for these young children. These task demands may also have interfered with the performance of the children who were retained in the study. Specifically, the children

were given very attractive objects and foods they were known to enjoy (i.e., the cheese crackers), and were required to give these items to the experimenter. This may be a difficult task particularly for children 20 to 24 months of age. In this age range, children begin to develop a sense of self as separate and distinct from others (Lewis, 1998). This new found self-awareness is often expressed through a desire for greater autonomy and a concern with ownership (e.g., the familiar catchword of 2-year-olds: "Mine!"). Thus, sharing attractive items and toys with an experimenter may be a demanding requirement for children in their "terrible twos", precisely because of their developing self-awareness.

Assertive and resistant behaviours are also predominant in toddlers, with naturalistic research documenting a two-fold increase in children's noncompliant responses to parental requests between the ages of 18 and 24 months (Dunn, 1988). This noncompliance may be attributed to children's growing awareness of rules and their desire to test rule violations (Dunn, 1988). It is possible that some children in the current study refused to give the experimenter the food or object he or she desired in order to test the experimenter's reaction to their noncompliance. The observation that the children who gave the experimenter the undesired item during the food request task were more likely to look at the experimenter's face than at the experimenter's hands or at the item suggests some noncompliance. Since these behaviours occurred predominately in the absence of smiling or laughing, it suggests that the children were not offering the undesired food in an attempt to tease the experimenter. Rather, the children may have been reluctant or unwilling to give the experimenter the food he or she desired and may have looked at the experimenter's face to monitor his or her reaction to their noncompliance. Although the children in the current sample were compliant in the warm-up task, offering the experimenter the requested objects, these objects were not as attractive as the test items. Moreover, children received social praise when they offered the experimenter the requested object during the warm-up task, providing the children with an incentive to comply with the experimenter's request. In contrast, during the

experimental tasks, the children received social praise if they offered the experimenter either of the test items, whether the items were desired or not. Therefore, the children may have been more willing, and/or encouraged, to share the requested objects with the experimenter during the warm-up task. Indeed, no relationship was found between children's performance on the warm-up task and their performance on the experimental tasks. Finally, it is possible that the children viewed the experimenter's affective displays during the experimental tasks as exaggerated and therefore, signalling "play". This may have invited them to test the consequences of rule violations in this game-like context.

The literature shows wide variability in children's performance on theory of mind tasks, depending on the specific methods used. For example, most studies using standard false-belief tasks have reported that an understanding of false beliefs does not emerge until 4 or 5 years of age (e.g., Gopnik & Slaughter, 1991; Hogrefe et al., 1986; Perner, Leekham, & Wimmer, 1987; Wellman et al., 2001). However, Chandler and Hala (1994) found evidence of false-belief understanding in 3-year-olds when deception tasks were used that provided the children with a motive for considering the other person's mental state. Specifically, the children in this study were actively engaged in the deception, being assigned the task of hiding a treat from a second experimenter, or replacing the contents of a familiar container with something unexpected to trick a second experimenter. Other studies have simplified the language of the false-belief questions asked of the children, have modified the task to increase the salience of the protagonists' mental states, or have reduced the salience of the contrasting real-world state of affairs. These procedural changes have also improved young children's performance on false-belief tasks (for a review, see Chandler & Hala, 1994; also see meta-analysis by Wellman et al., 2001). Finally, Bartsch and Wellman (1989) found that 3-year-old children were able to reason about false beliefs when they were asked to explain a story character's completed actions, rather than having to predict the character's future actions. In this study, the children were presented with stories in which the character performed some simple action, such as

looking for a kitten under the piano, when the kitten was in fact under the chair. The children were then asked to explain the character's anomalous action. These researchers found that 74% of the 3-year-olds tested offered an explanation for the character's behaviour that made reference to his or her false beliefs. Therefore, there is substantial evidence that methodological factors can influence children's performance on theory of mind tasks (see meta-analysis by Wellman et al., 2001).

Methodological factors have also been found to influence children's performance on gender-stereotyping tasks. Many studies investigating early knowledge of gender stereotypes have required children to demonstrate their understanding by pointing to, labelling, or sorting various depictions of gender-typed objects or activities (Leinbach, Hort, & Fagot, 1997; Weinraub et al., 1984). Recent research using the preferential looking paradigm, however, has demonstrated an understanding of gender stereotypes for toys (Serbin et al., 2001), metaphors (Eichstedt et al., in press), and activities (Serbin et al., in press) much earlier than previously documented. It may be that the object request task used in the current study was a less sensitive measure of gender understanding than paradigms based on visual attention methodology.

Thus, it is possible that methodological factors may have contributed to children's relatively poor performance in the present study, rather than a true lack of understanding of others' desires or of gender stereotypes. One possibility is that the affective displays used in the desire tasks may have confused the children. The facial and vocal displays were directed toward both the desired and undesired objects. Although the affect differed, it is possible that children were confused by the fact that attention (i.e., gaze and gesture) was directed toward both items. Research by Tilden (2000) suggests that children may have more difficulty inferring other people's desires when equal attention is shown to both objects. Order and task effects were also evident in the current study, and may have negatively impacted the children's performance.

With regard to motivational factors, the children were not provided with an

incentive or motive to infer the experimenter's desires. In fact, one could argue that the children may have been motivated to ignore the experimenter's desires in the object request task, since this would potentially allow the children to keep both items. Perhaps, an understanding of desires for food and gender-typed objects could have been demonstrated in the overall sample if the children were provided with an incentive or motive for doing so (e.g., if the children could comply with the request by doing a fun activity with the food or object, such as pushing the item down a slide to the experimenter). Alternatively, the children may have been aware of the experimenters' desires and/or the gender stereotyping of the objects, but may not have been motivated to give the experimenter the desired food items or objects. This explanation has some support. As discussed previously, some children in the overall sample were observed to have an implicit understanding of desires and gender stereotypes, as indicated by their low levels of gender-typed object preferences after they had observed the experimenters' expressing desires for cross-gender-typed objects. Although these children were aware of the experimenters' desires, they failed to demonstrate this knowledge explicitly by offering the experimenters the objects they desired. Of interest, children who performed more successfully on the food request task also performed more successfully on the object request task. Although this may suggest that an understanding of other people's desires for food and nonfood items are related, it may also suggest that these children were more compliant. Knowledge of gender stereotypes was also not explicitly demonstrated in the overall sample, with the children failing to offer the experimenter the gender-appropriate objects on the stereotyping trials. Perhaps, a clearer understanding of desires and gender stereotypes would have been demonstrated if implicit measures of understanding were used (e.g., anticipatory looking toward the desired item relative to the undesired item, or toward the gender-appropriate vs. -inappropriate object), rather than the object or food item offered to the experimenter. Anticipatory looking was found to be a sensitive measure of understanding in a task investigating children's reasoning about

beliefs (Clements & Perner, 1994). In this study, children as young as 2 years, 11 months of age were found to take into account another person's beliefs when predicting his or her future actions, much earlier than that demonstrated by previous research.

Response latencies have also been used as a more implicit measure of understanding. McCoy and Poulin-Dubois (1999), for example, used response latencies to determine whether 18-month-old children understood that desires are subjective. Two experimental conditions were employed. In the same-person condition, the same experimenter facially and vocally expressed a desire for one of two foods, then requested that the child give her some food. In the different-person condition, one experimenter expressed a desire for one of the two foods, and a different experimenter requested some food. The children were found to take significantly longer to give the experimenter a food in the different-person condition compared to the same-person condition. This increase in response latency for the different-person condition indicates that the children did not simply generalize the preferences of one experimenter to the second experimenter. Rather, they were unsure of which food the second experimenter desired, and therefore had more difficulty deciding which of the two foods to give to the experimenter. Response latencies were also collected and analysed in the current study to assess for differences between the gender-consistent and -inconsistent groups. Response latencies measure, in part, the extent of information processing required to make a decision. Thus, it was hypothesized that response latencies might be longer in the gender-inconsistent group, in which children were provided with affective cues to the experimenter's desires that conflicted with the information provided by the gender-based cues. In the gender-consistent group, both affective and gender cues were expected to lead to the same desire inference, resulting in shorter response latencies. Contrary to prediction, no differences in response latency were found between the two groups. Thus, there was no evidence that children in the gender-inconsistent group attempted to reconcile the conflicting information provided by the gender and affective cues in making their desire inferences.

Although response latencies can be useful in detecting differences between conditions in which information processing loads vary, such as between the gender-consistent and gender-inconsistent conditions, it is not useful as an implicit measure of desire understanding within conditions, where information processing demands are uniform.

Individual differences may also have contributed to the lack of significant findings for the food and object request tasks in the overall sample. For example, children's propensity to "tease" the experimenter by offering the undesired object or food item may have negatively impacted the results. Although the coding scheme used to detect emotional and behavioural responses to the task failed to demonstrate clear evidence of teasing behaviours, this may have been due to the poor sensitivity of the coding measure. Informally, several of the children were observed to display "teasing-like" behaviours, such as holding up a desired item out of reach of the experimenter, while smiling. Some of these children complied with the experimenter's request on certain trials, and displayed teasing responses on others. In the original study by Repacholi and Gopnik (1997), a number of participants were observed to offer and withdraw a piece of food. However, this behaviour appeared to be due to the child's reluctance to share the food item with the experimenter, rather than an attempt to tease the experimenter. The absence of teasing behaviours in the Repacholi and Gopnik study may be related to the age of the children tested. Naturalistic studies of children's interactions with family members have suggested that teasing increases in frequency and elaboration in the second year of life. For example, one study reported teasing behaviours in only 43% of children by 18 months of age, and 90% of children by 24 months of age (Dunn, 1988). This suggests that teasing behaviours may have been more prevalent in the current sample of 20- and 24-month-olds, than in Repacholi and Gopnik's sample of 14- and 18-month-olds. Moreover, in preschool children, teasing behaviours frequently revolve around issues of "possession" (see review by Keltner, Capps, Kring, Young, & Heerey, 2001), which would have impacted the give-and-take tasks used in the present experiment.

Another individual difference that may have impacted the results was the strength of the children's gender associations. The gender literature clearly delineates a gradual development of gender-stereotype knowledge, with very young children acquiring gender associations linking certain objects and activities with the gender categories before others (Kuhn et al., 1978; Leinbach et al., 1997; Poulin-Dubois et al., 2002; Weinraub et al., 1984). Thus, it is possible that individual children may have been more familiar with the gender stereotyping of some of the objects used in the present study than with others. The strength of the gender association linking a particular object with males or females may have influenced the child's willingness to give the experimenter the cross-gender-typed object in that pair. This was informally observed during some of the testing sessions. For example, one girl offered the female experimenter the desired masculine objects on the majority of trials during one of the visits. However, when the experimenter expressed disgust for the doll and a desire for the car, the girl stroked the doll saying, "Big baby, cute baby" and offered the experimenter the doll. In this instance, the child was aware that the experimenter expressed dislike for the doll, but this strongly violated her expectation that the female experimenter should like dolls. As a result, she offered the experimenter the doll on this trial, even though she willingly gave the experimenter cross-gender-typed objects on the other trials. The gender association linking dolls with girls, and vehicles with boys may be particularly strong, as it is among the first to be observed in girls, appearing as early as 18 months of age (Serbin et al., 2001). This variability in the strength of children's gender associations may account for some of the inconsistencies observed in their responses during the desire and gender-stereotyping trials. It may also have contributed to the lack of significant effects.

Understanding of Desires

To summarize, the present study suggests that a rudimentary understanding of desires for gender-typed objects is emerging in some children during the second year of life. This is consistent with previous research, which has documented an understanding of

desires for food items (Repacholi & Gopnik, 1997), knowledge of desire-dependent actions and emotions (Poulin-Dubois, 1999; Tilden, 2000), and explicit references to desires in children's naturalistic speech (Bartsch & Wellman, 1995) by 18 months of age. Of interest, children who showed a greater understanding of other people's desires for food items on the food request task were also somewhat more likely to correctly infer the experimenters' desires for gender-typed objects on the object request task. Although this correlation failed to reach statistical significance, this result suggests that a general understanding of desires may be emerging during this period, rather than item-specific knowledge. In a subset of children, an understanding of other people's desires was demonstrated explicitly during the object request task, and implicitly in the imitation of the experimenters' object preferences during the baseline trials. Implicit knowledge of desires was also observed informally in the teasing responses displayed by some of the children. It has been argued that teasing implies a knowledge of desires, as the child must know what is desired in order to withhold it and must also know that distress follows when a desire is not satisfied (Dunn, 1988; Reddy, 1991).

The current study, along with Repacholi and Gopnik's (1997) research, suggests that some children are able to infer desires on the basis of emotional or affective cues (facial and vocal) during the second year of life, even when attentional cues are held constant. However, this may be a relatively difficult task for children (see Tilden, 2000), with desire inferences more often based on multiple cues, including direction of gaze, physical orientation or movement toward an object, affective, and vocal cues.

Another factor that has been hypothesized to affect children's ability to reason about other people's desires is the similarity between the child's own desires and the expressed desires of the target individual. Specifically, previous research has suggested that children's reasoning about others' desires may be negatively impacted if they have strong and conflicting desires. Moore et al. (1995), for example, found that 3-year-old children had difficulty inferring another child's desire when they, themselves, strongly

desired a different object. These researchers attributed 3-year-olds' difficulties with desire inferences in these situations to their inability to inhibit responses based on their own, more "cognitively salient" mental states. However, in the present study, no relationship was found between children's object preferences and their ability to infer others' desires for these objects. This suggests that the children were able to override their own preferences for the objects in making inferences about the experimenter's desires. It also suggests that children understand that desires are subjective - that other people's desires may conflict with their own.

Thus, the present research demonstrates that by the second year of life, some children show an awareness that people have subjective, internal mental states that connect them to objects in the world. Specifically, they demonstrate an understanding that people have desires for objects, that desires are object-specific, and that desires and emotions are intimately related (i.e., a desired food or object is associated with happiness and an undesired food or object is associated with disgust).

Understanding of Gender Concepts

The present research suggests that an awareness of the gender stereotyping of familiar toys and objects is also emerging in 20- and 24-month-old children. Specifically, those children who performed more successfully on the food request task and demonstrated an adequate understanding of gender labels, were aware of the gender stereotyping of the objects used in the present study. Both boys and girls gave a higher proportion of feminine-typed items (i.e., doll, purse, necklace, and pink bear) to the female experimenter and a higher proportion of masculine-typed items (i.e., car, baseball glove, tie, and blue bear) to the male experimenter during the Visit 1 stereotyping trials. Moreover, an implicit awareness of the gender stereotyping of the objects was demonstrated by a subset of the children in the overall sample who observed the experimenter expressing desires for cross-gender-typed objects. These children used their newly acquired gender knowledge to pattern their subsequent play. Specifically, after

observing the experimenter display a preference for a cross-gender-typed object on a initial trial, they displayed preferences for other cross-gender-typed objects on subsequent trials. To accomplish this, the children had to be aware of the gender stereotyping of the four pairs of objects.

The finding of early gender knowledge in the present study is consistent with previous research reporting knowledge of gender-typed toys, activities, and metaphors in children in the second year of life (Eichstedt et al., in press; Poulin-Dubois et al., 2002; Serbin et al., 2001; Serbin et al., in press). The current study also extends previous research on toy stereotypes in infancy and toddlerhood. Serbin et al. (2001) found that 18-month-old girls were aware of the stereotyping of dolls and vehicles, associating dolls with girls and vehicles with boys. No such knowledge was found for boys in this age group, however. The current study expands this finding, suggesting that gender-typed toy knowledge is emerging in a subset of both boys and girls between the ages of 20 and 24 months. Furthermore, children in the present study demonstrated knowledge of the gender stereotyping of toys and objects other than vehicles and dolls, indicating a more comprehensive understanding of gender stereotypes.

The current study also suggested a significant positive relationship between children's knowledge of gender stereotypes for objects and their knowledge of gender labels. This relationship is consistent with gender-schema theory, which suggests that children first develop the categories of male and female and then begin to form associations linking certain objects, activities, and traits with each gender (Fagot & Leinbach, 1993; Martin, 1991; Martin & Halverson, 1981; Ruble & Martin, 1998). Moreover, it is similar to the results obtained by Fagot, Leinbach, and O'Boyle (1992) in their study of gender concepts in 24- to 36-month-old children. Although knowledge of stereotypes was found to be minimal, these researchers observed that children who labelled the sexes correctly showed greater stereotype knowledge than children who did not label the sexes correctly (see also Fagot & Leinbach, 1989; Martin & Little, 1990).

No relationship was found between children's knowledge of gender stereotypes for objects and their exposure to, or preference for, gender-typed objects in the present study, nor between their stereotype knowledge and their exposure to gender-typed activities in the home. The lack of a significant relationship between children's stereotype knowledge and their preference for gender-typed objects is not surprising given the weak relationship generally found between these two components in the literature (Hort et al., 1991; Ruble & Martin, 1998; Serbin et al., 1993; Weinraub et al., 1984). With respect to exposure to gender-typed activities in the home, several studies have documented less advanced knowledge of gender stereotypes in children from families in which there was a more egalitarian or nontraditional division of labour (Serbin et al., 1993; Turner & Gervai, 1995; Weinraub et al., 1984). However, exposure to gender-typed activities in the home may have a greater impact on children's knowledge of adult activities or possessions. In the Weinraub et al. (1984) study, for example, fathers' participation in feminine role activities was found to be related to children's knowledge of the gender stereotyping of adult possessions, but not to their knowledge of the gender stereotyping of toys.

The current study also suggests that young children's emerging gender schemas may be fairly flexible. Both the boys and the girls in the subsample demonstrated a willingness to offer the experimenter cross-gender-typed objects, despite having adequate knowledge of object gender stereotypes, as indicated by their performance on the Visit 1 stereotyping trials. The boys in the subsample offered the experimenters the desired cross-gender-typed objects in the gender-inconsistent condition. Specifically, although they were aware that a car, for example, is typically for boys, they were willing to violate that stereotype and give the car to a female experimenter if she expressed a desire for that item. Of interest, the girls offered the experimenters the cross-gender-typed objects in both the gender-inconsistent and gender-consistent conditions, despite the desires expressed by the experimenters. One possible interpretation for this response pattern is that the girls may have found it humourous to violate gender stereotypes. Thus, the

children appeared to be aware that gender stereotypes do not have to be adhered to rigidly. This is of interest as traditional accounts of the development of gender schemas suggest that stereotypes are likely to be held quite rigidly until approximately 7 or 8 years of age (Huston, 1983; Ruble & Martin, 1998; Signorella et al., 1993). Younger children have generally been thought to be unaware of the individual variation that can exist in masculinity and femininity among males and females (Ruble & Martin, 1998). The flexibility in young children's gender schemas was also evident in the changes in play preferences (i.e., a decrease in gender-typed play) that was observed for those children who witnessed the experimenter expressing a desire for cross-gender-typed objects.

The flexibility observed in the children's gender schemas was not due to limited knowledge of gender concepts. As mentioned previously, the children in the subsample showed an adequate knowledge of gender stereotypes for objects on the Visit 1 stereotyping trials. In fact, greater gender-category knowledge appeared to be associated with greater flexibility in children's gender schemas. Children in the overall sample with greater knowledge of gender labels were found to be more likely to give the experimenters the desired gender-inconsistent objects than were children with less gender-category knowledge. Perhaps, children with greater gender-category knowledge were more aware of the discrepancy between the experimenter's gender identity and the gender stereotyping of the objects desired. Because the experimenter's expressed desire strongly conflicted with their own gender knowledge, these children may have paid more attention to the experimenter's stated desire, which was reflected in their item selections.

The fact that the boys in the subsample used the information from the experimenters' affective displays, rather than gender information, as a basis for their desire inferences is also of interest. Previous research has consistently shown that children make stereotypic inferences based on knowledge of a person's sex. When told about unfamiliar boys and girls, for example, children will guess that these unknown children will prefer toys or display traits that are gender appropriate (Berndt & Heller,

1986; Haugh, Hoffman, & Cowan, 1980; Martin, 1989). The present research suggests that affective information is given more weight in determining another person's desires than gender information in boys in this age range.

Conclusions and Directions for Future Research

The present study suggests that an understanding of desires for gender-typed objects is emerging in some children by the second year of life. This extends the findings of previous research, which has demonstrated an understanding of others' desires for food items in 18-month-old infants (Repacholi & Gopnik, 1997). An understanding of the gender stereotyping of familiar objects also appears to be emerging during this period. However, the children did not use their budding knowledge of gender roles to make inferences about others' desires for gender-typed objects in this age range. Rather, they tended to rely on the individuating information provided by the experimenter's affective displays to make their desire inferences, rather than basing their inferences on category-driven information. This finding also implies flexibility in children's early gender schemas, with children demonstrating some understanding that one's desires may conflict with traditional gender stereotypes. The results of the present study are limited by the small sample size, and the finding that some, but not all, of the children were able to demonstrate knowledge of desires and gender stereotypes in the food and object request paradigms.

Although the food request paradigm was shown to be an effective measure of desire understanding among 18-month-old children, this paradigm may not be appropriate for children 20 months of age or older. Perhaps the use of implicit measures of desire reasoning (e.g., anticipatory looking) may have provided stronger evidence of desire understanding than the object or food item offered to the experimenter. Future research is needed to develop an effective and reliable paradigm for assessing desire understanding across infancy, toddlerhood, and the early preschool years. This will allow researchers to trace the development of desire understanding, and to examine the links between desire

understanding and other theory of mind developments (e.g., knowledge of beliefs and false beliefs).

Previous research has generally focused on isolated social processes, including gender reasoning, emotional understanding, the development of empathy, and theory of mind. However, these social processes do not operate in a vacuum. Rather, they combine and interact in a myriad of ways. Thus, it will be important for future research to address the interplay of processes in children's social reasoning. The present study attempted to examine the interaction of gender knowledge and children's desire reasoning. This interaction could also be addressed by examining whether children make generalizations about others' desires based on their awareness of gender concepts. That is, if a female experimenter expresses a desire for a toy car and dislike for a doll, will the child then assume that the experimenter will also prefer a baseball glove over a purse? Investigation of the links between gender understanding and desire reasoning across development is also needed. The current study assessed children's use of gender information in reasoning about others' desires for gender-typed objects in 20- and 24-month-old children. These two age groups were selected as they were hypothesized to differ in their level of gender knowledge, allowing desire reasoning to be compared in children with and without gender-typed knowledge. However, contrary to prediction, both age groups were found to be highly similar in their knowledge of gender labels, gender-typed objects, and exposure to gender stereotyping in their homes. Thus, a larger age span would need to be sampled in future research (e.g., 18-, 24-, 30, and 36-month-olds) to fully explore the effect of gender knowledge on desire reasoning.

The present study represents an initial attempt to integrate different areas of social cognition. The results suggest that children's social understanding is quite sophisticated and complex, even at this early age. Much remains to be discovered, however. Future research is needed to assess how children's developing knowledge of social categories (e.g., gender categories, racial groups, social classes, etc.), their social experiences (e.g.,

family interactions, friendships), and their understanding of mental states interact to allow them to predict and understand other people's behaviour. How these factors emerge and interact across development, as well as how they contribute to the child's emerging self-concept and formation of interpersonal relationships is of crucial importance. Only through such research can we begin to have a full understanding of social development across the lifespan, as well as how deviations occur in normal social development.

References

Archer, J. (1984). Gender roles as developmental pathways. *British Journal of Social Psychology, 23,* 245-256.

Astington, J., Harris, P., & Olson, D. (1988). *Developing theories of mind.* New York: Cambridge University Press.

Baron-Cohen, S., & Ring, H. (1994). A model of the mindreading system: Neuropsychological and neurobiological perspectives. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 183-207). Hillsdale, NJ: Erlbaum.

Bartsch, K., & Wellman, H. M. (1989). Young children's attribution of action to beliefs and desires. *Child Development, 60,* 946-964.

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind.* New York: Oxford University Press.

Bauer, P. J., Liebl, M., & Stennes, L. (1998). PRETTY is to DRESS as BRAVE is to SUITCOAT: Gender-based property-to-property inferences by 4-1/2-year-old children. *Merrill-Palmer Quarterly, 44,* 355-377.

Beissel, C. (1998). *Infants' knowledge of gender-stereotyped adult activities: Evidence from an inductive paradigm.* Unpublished undergraduate honour's thesis, Concordia University, Montreal, Québec, Canada.

Berndt, T. J., & Heller, K. A. (1986). Gender stereotypes and social inferences: A developmental study. *Journal of Personality and Social Psychology, 50,* 889-898.

Blakemore, J. E., LaRue, A. A., & Olejnik, A. B. (1979). Sex-appropriate toy preferences and the ability to conceptualize toys as sex-role related. *Developmental Psychology, 15,* 339-340.

Bradbard, M. R., Martin, C. L., Endsley, R. C., & Halverson, C. F. (1986). Influence of sex stereotypes on children's exploration and memory: A competence versus performance distinction. *Developmental Psychology, 22,* 481-486.

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology, 18,* 906-921.

Bretherton, I., McNew, S., & Beeghly-Smith, M. (1981). Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a "theory of mind"? In M. E. Lamb & L. R. Sherrod (Eds.), *Infant social cognition: Empirical and theoretical considerations* (pp. 333-373). Hillsdale, NJ: Erlbaum.

Brown, J. R., & Dunn, J. (1991). 'You can cry, mum': The social and developmental implications of talk about internal states. *British Journal of Developmental Psychology, 9,* 237-256.

Bushnell, I. W., Sai, F., & Mullin, J. T. (1989). Neonatal recognition of the mother's face. *British Journal of Developmental Psychology, 7,* 3-15.

Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review, 106,* 676-713.

Chandler, M., & Hala, S. (1994). The role of personal involvement in the assessment of early false belief skills. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 403-425). Hillsdale, NJ: Erlbaum.

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development, 9,* 377-395.

Cobb, N. J., Stevens-Long, J., & Goldstein, S. (1982). The influence of televised models on toy preferences in children. *Sex Roles, 8,* 1075-1080.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cordua, G. D., McGraw, K. O., & Drabman, R. S. (1979). Doctor or nurse: Children's perception of sex typed occupations. *Child Development, 50,* 590-593.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208,* 1174-1176.

Dunn, J. (1988). *The beginnings of social understanding*. Oxford, England: Basil Blackwell.

Dunn, J. (1991). Young children's understanding of other people: Evidence from observations within the family. In D. Frye & C. Moore (Eds.), *Children's theories of mind: Mental states and social understanding* (pp. 97-114). Hillsdale, NJ: Erlbaum.

Dunn, J., & Munn, P. (1985). Becoming a family member: Family conflict and the development of social understanding in the second year. *Child Development, 56*, 480-492.

Eichstedt, J. A., Serbin, L. A., Poulin-Dubois, D., & Sen, M. G. (in press). Of bears and men: Infants' knowledge of conventional and metaphorical gender stereotypes. *Infant Behavior and Development*.

Ekman, P., & Friesen, W. (1975). *Unmasking the face: A guide to recognizing emotions from facial cues*. Englewood Cliffs, NJ: Prentice-Hall.

Fagot, B. I., & Leinbach, M. D. (1989). The young child's gender schema: Environmental input, internal organization. *Child Development, 60*, 663-672.

Fagot, B. I., & Leinbach, M. D. (1993). Gender-role development in young children: From discrimination to labeling. *Developmental Review, 13*, 205-224.

Fagot, B. I., Leinbach, M. D., O'Boyle, C. (1992). Gender labeling, gender stereotyping, and parenting behaviors. *Developmental Psychology, 28*, 225-230.

Fein, G., Johnson, D., Kosson, N., Stork, L., & Wasserman, L. (1975). Sex stereotypes and preferences in the toy choices of 20-month-old boys and girls. *Developmental Psychology, 11*, 527-528.

Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology, 50*, 21-45.

Frye, D., & Moore, C. (1991). *Children's theories of mind: Mental states and social understanding*. Hillsdale, NJ: Erlbaum.

Gopnik, A., & Slaughter, V. (1991). Young children's understanding of changes in their mental states. *Child Development, 62,* 98-110.

Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics, 56,* 544-549.

Gunnar, M. R., & Stone, C. (1984). The effects of positive maternal affect on infant responses to pleasant, ambiguous, and fear-provoking toys. *Child Development, 55,* 1231-1236.

Hadwin, J., & Perner, J. (1991). Pleased and surprised: Children's cognitive theory of emotion. *British Journal of Developmental Psychology, 9,* 215-234.

Haugh, S. S., Hoffman, C. D., & Cowan, G. (1980). The eye of the very young beholder: Sex typing of infants by young children. *Child Development, 51,* 598-600.

Hogrefe, G. J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development, 57,* 567-582.

Hort, B. E., Fagot, B. I., & Leinbach, M. D. (1990). Are people's notions of maleness more stereotypically framed than their notions of femaleness? *Sex Roles, 23,* 197-212.

Hort, B. E., Leinbach, M. D., & Fagot, B. I. (1991). Is there coherence among the cognitive components of gender acquisition? *Sex Roles, 24,* 195-207.

Huston, A. C. (1983). Sex-typing. In P. H. Mussen (Series Ed.) & E. M. Hetherington (Vol. Ed.), *Handbook of child psychology: Vol. 4. Socialization, personality, and social development* (4th ed., pp. 387-467). New York: Wiley.

Jennings, S. A. (1975). Effects of sex typing in children's stories on preference and recall. *Child Development, 46,* 220-223.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40,* 1-19.

Johnston, K. E., Bittinger, K., Smith, A., & Madole, K. L. (2001). Developmental changes in infants' and toddlers' attention to gender categories. *Merrill-Palmer Quarterly, 47*, 563-584.

Katz, P. A., & Walsh, V. (1991). Modification of children's gender-stereotyped behavior. *Child Development, 62*, 338-351.

Keltner, D., Capps, L., Kring, A. M., Young, R. C., & Heerey, E. A. (2001). Just teasing: A conceptual analysis and empirical review. *Psychological Bulletin, 127*, 229-248.

Koblinsky, S. G., Cruse, D. F., & Sugawara, A. I. (1978). Sex role stereotypes and children's memory for story content. *Child Development, 49*, 452-458.

Kropp, J. J., & Halverson, C. F. (1983). Preschool children's preferences and recall for stereotyped versus nonstereotyped stories. *Sex Roles, 9*, 261-272.

Kuhn, D., Nash, S. C., & Brucken, L. (1978). Sex role concepts of two- and three-year-olds. *Child Development, 49*, 445-451.

Leinbach, M. D., & Fagot, B. I. (1986). Acquisition of gender labels: A test for toddlers. *Sex Roles, 15*, 655-666.

Leinbach, M. D., & Fagot, B. I. (1993). Categorical habituation to male and female faces: Gender schematic processing in infancy. *Infant Behavior and Development, 16*, 317-332.

Leinbach, M. D., Hort, B. E., & Fagot, B. I. (1997). Bears are for boys: Metaphorical associations in young children's gender stereotypes. *Cognitive Development, 12*, 107-130.

Lewis, M. (1998). Emotional competence and development. In D. Pushkar, W. M. Bukowski, A. E. Schwartzman, D. M. Stack, & D. R. White (Eds.), *Improving competence across the lifespan: Building interventions based on theory and research* (pp. 27-36). New York: Plenum Press.

Liben, L. S., & Signorella, M. L. (1993). Gender-schematic processing in children: The role of initial interpretations of stimuli. *Developmental Psychology, 29,* 141-149.

Martin, C. L. (1989). Children's use of gender-related information in making social judgements. *Developmental Psychology, 25,* 80-88.

Martin, C. L. (1991). The role of cognition in understanding gender effects. *Advances in Child Development and Behaviour, 23,* 113-149.

Martin, C. L., Eisenbud, L., & Rose, H. (1995). Children's gender-based reasoning about toys. *Child Development, 66,* 1453-1471.

Martin, C. L., & Halverson, C. F. (1981). A schematic processing model of sex typing and stereotyping in children. *Child Development, 52,* 1119-1134.

Martin, C. L., & Halverson, C. F. (1983). The effects of sex-typing schemas on young children's memory. *Child Development, 54,* 563-574.

Martin, C. L., & Little, J. K. (1990). The relation of gender understanding to children's sex-typed preferences and gender stereotypes. *Child Development, 61,* 1427-1439.

Martin, C. L., Wood, C. H., & Little, J. K. (1990). The development of gender stereotype components. *Child Development, 61,* 1891-1904.

McKoy, K., & Poulin-Dubois, D. (1999, April). *Understanding of the subjectivity of desires in 18-month-olds*. Poster session presented at the biennial meeting of the Society for Research in Child Development, Albuquerque, New Mexico.

Meltzoff, A. N., Gopnik, A., & Repacholi, B. M. (1999). Toddlers' understanding of intentions, desires, and emotions: Explorations of the dark ages. In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 17-41). Mahwah, NJ: Erlbaum.

Miller, C. L., Younger, B. A., & Morse, P. A. (1982). The categorization of male and female voices in infancy. *Infant Behavior and Development, 5,* 143-159.

Moore, C., Jarrold, C., Russell, J., Lumb, A., Sapp, F., & MacCallum, F. (1995). Conflicting desire and the child's theory of mind. *Cognitive Development, 10,* 467-482.

Moses, L. J., Coon, J. A., & Wusinich, N. (2000). Young children's understanding of desire formation. *Developmental Psychology, 36,* 77-90.

O'Brien, M. (1992). Gender identity and sex roles. In V. B. Hanhasselt & M. Hersen (Eds.). *Handbook of social development: A lifespan perspective* (pp. 325-345). New York: Plenum Press.

O'Brien, M. & Huston, A. C. (1985). Development of sex-typed play behavior in toddlers. *Developmental Psychology, 21,* 866-871.

Orlofsky, J. L. (1981). Relationship between sex-role attitudes and personality traits and the Sex Role Behavior Scale-1: A new measure of masculine and feminine role behaviors and interests. *Journal of Personality and Social Psychology, 40,* 927-940.

Perner, J., Leekam, S., & Wimmer, H. (1987). Three-year-old's difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5,* 125-137.

Perry, D. G., & Bussey, K. (1979). The social learning theory of sex differences: Imitation is alive and well. *Journal of Personality and Social Psychology, 37,* 1699-1712.

Perry, D. G., White, A. J., & Perry, L. C. (1984). Does early sex typing result from children's attempts to match their behavior to sex role stereotypes? *Child Development, 55,* 2114-2121.

Picariello, M. L., Greenberg, D. N., & Pillemer, D. B. (1990). Children's sex-related stereotyping of colors. *Child Development, 61,* 1453-1460.

Poulin-Dubois, D. (1999). Infants' distinction between animate and inanimate objects: The origins of naive psychology. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first months of life* (pp. 256-280). Mahwah, NJ: Erlbaum.

Poulin-Dubois, D., Serbin, L. A., & Derbyshire, A. (1998). Toddlers' intermodal and verbal knowledge about gender. *Merrill-Palmer Quarterly, 44*, 338-354.

Poulin-Dubois, D., Serbin, L. A., Eichstedt, J. A., Sen, M. G., & Beissel, C. F. (2002). Men don't put on make-up: Toddlers' knowledge of the gender stereotyping of household activities. *Social Development, 11*, 166-181.

Poulin-Dubois, D., Serbin, L. A., Kenyon, B., & Derbyshire, A. (1994). Infants' intermodal knowledge about gender. *Developmental Psychology, 30*, 436-442.

Powlishta, K. K., Sen, M. G., Serbin, L. A., Poulin-Dubois, D., & Eichstedt, J. A. (2001). From infancy through middle childhood: The role of cognitive and social factors in becoming gendered. In R. K. Unger (Ed.), *Handbook of the psychology of women and gender* (pp. 116-132). New York: Wiley.

Reddy, V. (1991). Playing with others' expectations: Teasing and mucking about in the first year. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading* (pp. 143-158). Oxford, England: Basil Blackwell.

Reis, H. T., & Wright, S. (1982). Knowledge of sex-role stereotypes in children aged 3 to 5. *Sex Roles, 8*, 1049-1056.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33*, 12-21.

Ross, D. M., & Ross, S. A. (1972). Resistance by preschool boys to sex-inappropriate behavior. *Journal of Educational Psychology, 63*, 342-346.

Ruble, D. N., & Martin, C. L. (1998). Gender development. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 933-1016). New York: Wiley.

Ruble, D. N., & Stangor, C. (1986). Stalking the elusive schema: Insights from developmental and social-psychological analyses of gender schemas. *Social Cognition, 4*, 227-261.

Serbin, L. A., Connor, J. M., & Iler, I. (1979). Sex-stereotyped and nonstereotyped introductions of new toys in the preschool classroom: An observational study of teacher behavior and its effects. *Psychology of Women Quarterly, 4*, 261-265.

Serbin, L. A., Poulin-Dubois, D., Colburne, K. A., Sen, M. G., & Eichstedt, J. A. (2001). Gender stereotyping in infancy: Visual preferences for and knowledge of gender-stereotyped toys in the second year. *International Journal of Behavioral Development, 25*, 7-15.

Serbin, L. A., Poulin-Dubois, D., & Eichstedt, J. A. (in press). Infants' responses to gender-inconsistent events. *Infancy*.

Serbin, L. A., Powlishta, K. K., & Gulko, J. (1993). The development of sex typing in middle childhood. *Monographs of the Society for Research in Child Development, 58*(2, Serial No. 232).

Shell, R., & Eisenberg, N. (1990). The role of peers' gender in children's naturally occuring interest in toys. *International Journal of Behavioral Development, 13*, 373-388.

Signorella, M. L., Bigler, R. S., & Liben, L. S. (1993). Developmental differences in children's gender schemata about others: A meta-analytic review. *Developmental Review, 13*, 147-183.

Spelke, E., Phillips, A., & Woodward, A. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 44-78). New York: Clarendon.

Stein, N. L., & Levine, L. J. (1989). The causal organization of emotional knowledge: A developmental study. *Cognition and Emotion, 3,* 343-378.

Thompson, S. K. (1975). Gender labels and early sex role development. *Child Development, 46,* 339-347.

Tilden, J. (2000). Toddlers' reasoning about the origins of human actions, emotions, and knowledge (Doctoral dissertation, Concordia University, 1999). *Dissertation Abstracts International, 60 (12B),* 6400.

Turner, P. J., & Gervai, J. (1995). A multidimensional study of gender typing in preschool children and their parents: Personality, attitudes, preferences, behavior, and cultural differences. *Developmental Psychology, 31,* 759-772.

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin, 121,* 437-456.

Walker-Andrews, A. S., Bahrick, L. E., Raglioni, S. S., & Diaz, I. (1991). Infants' bimodal perception of gender. *Ecological Psychology, 3,* 55-75.

Weinraub, M., Clemens, L. P., Sockloff, A., Ethridge, T., Gracely, E., & Myers, B. (1984). The development of sex role stereotypes in the third year: Relationships to gender labelling, gender identity, sex-typed toy preference, and family characteristics. *Child Development, 55,* 1493-1503.

Wellman, H. M. (1990). *The child's theory of mind.* Cambridge, MA: M.I.T. Press.

Wellman, H. M. (1993). Early understanding of mind: The normal case. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from autism* (pp. 10-39). Oxford, England: Oxford University Press.

Wellman, H. M., & Banerjee, M. (1991). Mind and emotion: Children's understanding of the emotional consequences of beliefs and desires. *British Journal of Developmental Psychology, 9,* 191-214.

Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition, 30,* 239-277.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false-belief. *Child Development, 72,* 655-684.

Wellman, H. M., & Wooley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition, 35,* 245-275.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69,* 1-34.

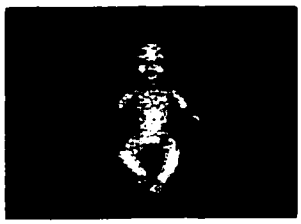
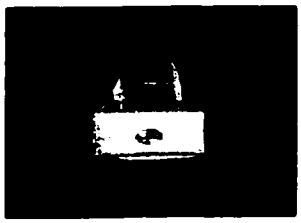
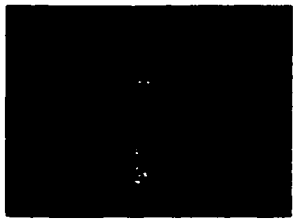
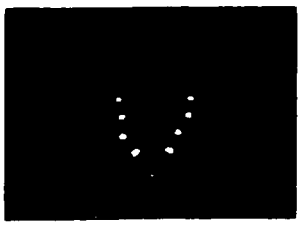
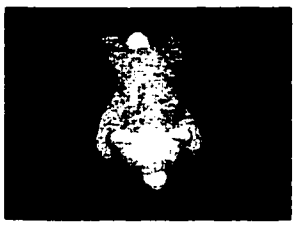
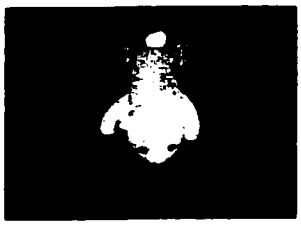
Yuill, N. (1984). Young children's coordination of motive and outcome in judgements of satisfaction and morality. *British Journal of Developmental Psychology, 2,* 73-81.

Yuill, N., & Pearson, A. (1998). The development of bases for trait attribution: Children's understanding of traits as causal mechanisms based on desire. *Developmental Psychology, 34,* 574-586.

Appendix A

Item Pairs Used in the Object Request Task

Item Pairs Used in the Object Request Task

Masculine Stereotyped	*Feminine Stereotyped*
	
	
	
	

Appendix B

Judges' Ratings of the Masculine- and Feminine-Typed Items Used in the Object Request Task

Table B1

Means, Standard Deviations, and t Tests for the Judges' Ratings of the Masculine- and Feminine-Typed Items Used in the Object Request Task

Item	*M*	*SD*	*df*	*t*
Masculine-Typed Items				
Car	1.80	0.63	9	-6.00 ***
Glove	1.20	0.42	9	-13.50 ***
Tie	2.00	0.82	9	-3.87 **
Blue Bear	2.90	0.57	9	-0.56
Feminine-Typed Items				
Doll	4.30	0.48	9	8.51 ***
Purse	5.00	0.00	9	***
Necklace	4.90	0.32	9	19.00 ***
Pink Bear	4.60	0.52	9	9.80 ***

Note. The mean ratings for each item were compared to a gender-neutral score of 3 using *t* tests.

** $p < .01$. *** $p < .001$.

Appendix C

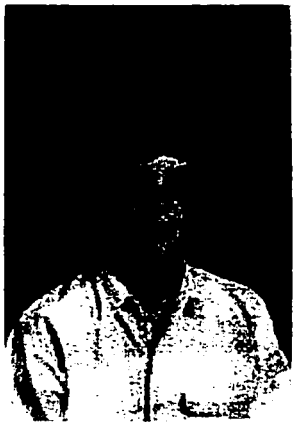
Examples of the Child and Adult Photograph Pairs Used in the Gender-Labelling Task

(European-American and Visible Minorities Versions)

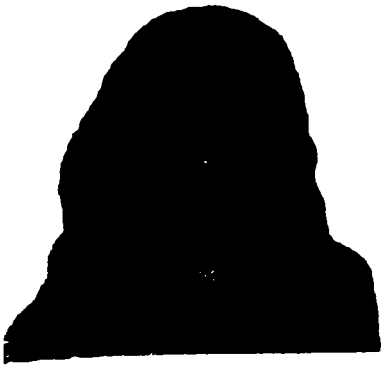
Examples of the Child Photograph Pairs Used in the Gender-Labelling Task

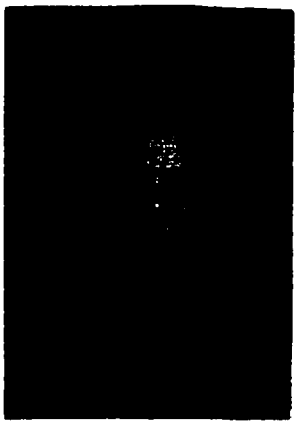
(European-American and Visible Minorities Versions)

Photographs of Boys



Photographs of Girls

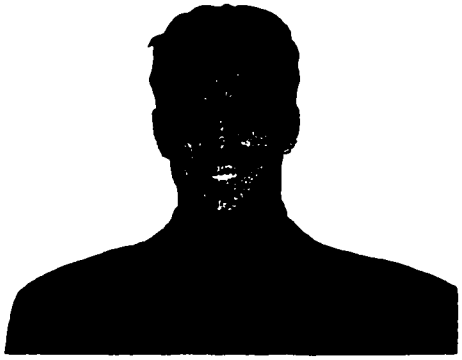


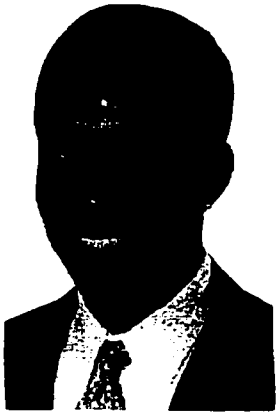


Examples of the Adult Photograph Pairs Used in the Gender-Labelling Task

(European-American and Visible Minorities Versions)

Photographs of Men





Photographs of Women

Appendix D

Judges' Ratings of the Child and Adult Photographs Used in the Gender-Labelling Task

Table D1

Means, Standard Deviations, and t Tests for the Judges' Ratings of the Child Photographs Used in the Gender-Labelling Task, European-American Version

Photograph	*M*	*SD*	*df*	*t*
Photographs of Boys				
Photo 1	1.40	0.52	9	-9.80 ***
Photo 2	2.10	0.74	9	-3.86 **
Photo 3	..90	0.32	9	-11.00 ***
Photo 4	1.60	0.52	9	-8.57 ***
Photo 5	1.20	0.42	9	-13.50 ***
Photo 6	1.60	0.52	9	-8.57 ***
Photographs of Girls				
Photo 1	4.60	0.52	9	9.80 ***
Photo 2	4.30	1.25	9	3.28 **
Photo 3	4.30	0.48	9	8.51 ***
Photo 4	4.90	0.32	9	19.00 ***
Photo 5	4.40	0.52	9	8.57 ***
Photo 6	4.90	0.32	9	19.00 ***

Note. The mean ratings for each photo were compared to a gender-neutral score of 3 using *t* tests.

** $p < .01$. *** $p < .001$.

Table D2

Means, Standard Deviations, and t Tests for the Judges' Ratings of the Adult Photographs Used in the Gender-Labelling Task, European-American Version

Photograph	*M*	*SD*	*df*	*t*
Photographs of Men				
Photo 1	1.10	0.32	9	-19.00***
Photo 2	1.20	0.42	9	-13.50***
Photo 3	1.00	0.00	9	***
Photo 4	1.00	0.00	9	***
Photo 5	1.10	0.32	9	-19.00***
Photo 6	1.10	0.32	9	-19.00***
Photographs of Women				
Photo 1	5.00	0.00	9	***
Photo 2	5.00	0.00	9	***
Photo 3	5.00	0.00	9	***
Photo 4	4.80	0.42	9	13.50***
Photo 5	5.00	0.00	9	***
Photo 6	5.00	0.00	9	***

Note. The mean ratings for each photo were compared to a gender-neutral score of 3 using *t* tests.

*** $p < .001$.

Table D3

Means, Standard Deviations, and t Tests for the Judges' Ratings of the Child Photographs Used in the Gender-Labelling Task. Visible Minorities Version

Photograph	*M*	*SD*	*df*	*t*
Photographs of Boys				
Photo 1	1.10	0.32	9	-19.00***
Photo 2	2.20	1.03	9	-2.45*
Photo 3	1.70	0.67	9	-6.09***
Photo 4	1.10	0.32	9	-19.00***
Photo 5	1.30	0.48	9	-11.13***
Photo 6	1.60	0.52	9	-8.57***
Photographs of Girls				
Photo 1	4.90	0.32	9	19.00***
Photo 2	5.00	0.00	9	***
Photo 3	5.00	0.00	9	***
Photo 4	4.70	0.48	9	11.13***
Photo 5	4.50	0.53	9	9.00***
Photo 6	4.90	0.32	9	19.00***

Note. The mean ratings for each photo were compared to a gender-neutral score of 3 using *t* tests.

* $p < .05$ ** $p < .01$. *** $p < .001$.

Table D4

Means, Standard Deviations, and t Tests for the Judges' Ratings of the Adult Photographs Used in the Gender-Labelling Task. Visible Minorities Version

Photograph	*M*	*SD*	*df*	*t*
Photographs of Men				
Photo 1	1.00	0.00	9	•••
Photo 2	1.00	0.00	9	•••
Photo 3	1.10	0.32	9	-19.00•••
Photo 4	1.10	0.32	9	-19.00•••
Photo 5	1.10	0.32	9	-19.00•••
Photo 6	1.10	0.32	9	-19.00•••
Photographs of Women				
Photo 1	4.80	0.42	9	13.50•••
Photo 2	4.90	0.32	9	19.00•••
Photo 3	4.80	0.42	9	13.50•••
Photo 4	4.60	0.52	9	9.80•••
Photo 5	5.00	0.00	9	•••
Photo 6	5.00	0.00	9	•••

Note. The mean ratings for each photo were compared to a gender-neutral score of 3 using *t* tests.

••• $p < .001$.

Appendix E

Sex-Typed Child Care and Household Tasks Questionnaire

Parents' Activity Questionnaire

Filled out by: _____ Code: _____

Please use the following scale to rate the list of activities according to whether you or your spouse is more likely to do each of the tasks described below. Place a check mark in the box corresponding to your answer. If any of the tasks do not apply, please rate them according to whom you imagine would be more likely to do the chore.

A = much more likely to be done by my spouse
B =somewhat more likely to be done by my spouse
C = equally likely to be done by me or my spouse
D =somewhat more likely to be done by me
E = much more likely to be done by me

Household responsibilities

	much more by my spouse		equally by me or my spouse		much more by me
	A	**B**	**C**	**D**	**E**
housework					
washing the car					
yard work					
opening the wine					
vacuuming					
dusting					
cooking					
emptying the garbage					
doing the barbecuing					
cleaning the bathroom					
mowing the lawn					
adjusting the thermostat					
shovelling the snow					
painting the inside of the house					
repairing household appliances					

	much more by my spouse		equally by me or my spouse		much more by me
	A	B	C	D	E
steam cleaning the carpets					
doing the laundry					
writing items needed on the shopping list					
making travel arrangements for vacation					
changing fuses or resetting circuit breakers					
taking pictures at family household events					
buying groceries					

Child care responsibilities:

	A	B	C	D	E
taking the children along when going on errands					
explaining the facts of life to girls					
infant care					
teaching one's children to drive					
playing catch with children					
buying clothes for children					

	much more by my spouse		equally by me or my spouse		much more by me
	A	B	C	D	E
giving the child a hug					
making sure the children get medical and dental care					
taking the children to a ball game					
getting the children breakfast					
reading a bed-time story to children					
giving the children allowance					
talking about the good old days with children					
talking over things with one's son when he's having problems					
disciplining the girls					
disciplining the boys					
letting the children learn by their mistakes					
dealing with the neighbour whose window the child has broken					
playing with boys					
playing with girls					
explaining the facts of life to boys					
going fishing with children					
helping the children put together unassembled toys					

	much more by my spouse		equally by me or my spouse		much more by me
	A	B	C	D	E
spending a good deal of time with the children					
dealing with a child's teacher when there is a problem at school					
taking the children to church or synagogue					
talking over things with one's daughter when she's having problems					
dealing with the police when the child breaks the law					
going to parent/teacher meetings					

Appendix F

Item Familiarity Questionnaire

FU
54677

IDNO

Filled out by: _____

Toy Questionnaire

Please indicate which of the following toys your child has at home or at daycare by placing a check mark in the appropriate box. If your child does have this toy, please indicate on the rating scale how often your child plays with this toy by circling the appropriate word.

Dolls

O Does not have dolls

O Has at home: own

O Has at home: sibling(s)

O Has at daycare

Exact doll as in study? Y / N

If applicable, please indicate what types of dolls your child has either at home or at daycare?

O baby dolls
O Barbie or similar
O Ken or similar
O action figures (e.g., G. I. Joe)
O soft, cloth dolls
O small plastic dolls (e.g., for dollhouse)
O other, please specify: _____

Frequency with which your child plays with dolls:

Never Rarely Sometimes Often Always

Cars

O Does not have cars

O Has at home: own

O Has at home: sibling(s)

O Has at daycare

Exact car as in study? Y / N

Frequency with which your child plays with cars:

Never Rarely Sometimes Often Always

54677

Purses

O Does not have purses

O Has at home: own

O Has at home: sibling(s)

O Has at daycare

Exact purse as in study? Y / N

Frequency with which your child plays with purses:

Never Rarely Sometimes Often Always

Baseball Gloves

O Does not have baseball gloves

O Has at home: own

O Has at home: sibling(s)

O Has at daycare

Exact baseball glove as in study? Y/ N

Frequency with which your child plays with baseball gloves:

Never Rarely Sometimes Often Always

54677

Necklaces

O Does not have necklaces

O Has at home: own

O Has at home: sibling(s)

O Has at daycare

Exact necklace as in study? Y/N

Does your child see any member of the household wearing necklaces? Y/N

Please indicate sex of family member: O female O male

How frequently does this person wear necklaces:

Never Rarely Sometimes Often Always

Frequency with which your child plays with or wears necklaces:

Never Rarely Sometimes Often Always

Ties

O Does not have ties

O Has own tie(s)

O Sibling has tie(s)

Exact tie as in study? Y/N

Is there a male in the household that wears ties? Y/N

How frequently does this person wear ties:

Never Rarely Sometimes Often Always

Frequency with which your child wears ties:

Never Rarely Sometimes Often Always

54677

Pink Clothes or Toys

O Does not have any pink clothes
O Has own pink clothes
O Sibling has pink clothes

O Does not have any pink toys
O Has at home: own
O Has at home: sibling(s)'
O Has at daycare

Does your child have the exact pink bear as in study? Y/N

What is the dominant colour of your child's room? _____

How often does your child wear pink clothing?

Never Rarely Sometimes Often Always

Amount of pink toys that your child has at home or at daycare?

None A few Several Many Most

Blue Clothes or Toys

O Does not have any blue clothes
O Has own blue clothes
O Sibling has blue clothes

O Does not have any blue toys
O Has at home: own
O Has at home: sibling(s)'
O Has at daycare

Does your child have the exact blue bear as in study? Y/N

How often does your child wear blue clothing?

Never Rarely Sometimes Often Always

Amount of blue toys that your child has at home or at daycare?

None A few Several Many Most

Appendix G

Coding Scheme for Teasing Behaviours and Emotional Reactions to the Task

Coding Scheme for Teasing Behaviours and Emotional Reactions to the Task

For each trial, code whether any of the following behaviours or emotional reactions are present. For the desire trials, if the behaviour or emotional reaction occurs during the experimenter's affective display, prior to the request for the object or food item, place a checkmark in the *Before Request* box. If the behaviour or emotional reaction occurs after the experimenter's request, place a checkmark in the *After Request* box.

Incongruity

Code if any of the following behaviours or affects are present:

- surprise: eyebrows raised, eyes widened; jaw "dropped" with teeth and lips parted
- puzzled expression: child frowns. That is, brows are slanted down and drawn in slightly or furrowed; closed, neutral mouth; lower eyelid raised
- repeats phrase(s) in a questioning tone
- in serious or questioning tone, verbally corrects experimenter's desire (e.g., labels desired object as "yucky" or undesired object as "nice" etc.). Specify label and which object is labelled.

Verbal Tease

- in a teasing tone, child states that experimenter wants a specific object or food item, or labels a particular object or food item as desirable or undesirable. Statement must be accompanied by a smile, smirk, or laugh, as well as direct eye contact with the experimenter. Specify the object or food item the child refers to.

Humour

- Child smiles or laughs. Code *Low (L)* or *High (H)* intensity.
- *For High Intensity:* Smile must be pronounced. That is, a pronounced wrinkle (nasal labial fold) runs down from the nose to the outer edge beyond the lip corners; lower eyelid shows wrinkles below it; the corner of the lips are drawn back and up, and the mouth is open; cheeks are raised. If the child laughs, this should also be coded as *High Intensity (H)*.
- *Note:* If the child's affect stays the same both before and after the experimenter's request for a object, code humour in both *Before Request* and *After Request* categories.

Monitoring Gaze/Gaze Alternation

- While holding an item, the child looks directly at the experimenter's eyes and holds eye contact for several seconds. Alternatively, while holding an item, the child slowly and deliberately looks at the item, then at the experimenter's eyes, and back to the item (or vice versa). If the child is in the process of giving the object or food item to the experimenter, the child must pause before giving the item to the experimenter. Specify if the act occurs (a) *with smile,* if accompanied by a smile, smirk, or laugh; or (b) *without smile.*

Eye Contact with Experimenter when Object or Food Item Given

- Child looks directly at the experimenter's face as he/she gives the object or food item, or immediately prior to giving object or food item. The child gives the item in a fluid motion, with no pause or hesitation evident. Specify if the act occurs (a) *with smile*, if accompanied by a smile, smirk, or laugh; or (b) *without smile*.

Offer/Withdrawal

- Child offers one item to the experimenter (i.e., puts item in the experimenter's hands or holds item up in front of, or above, the experimenter's hands), then withdraws and gives the other item. Record which object or food item is withdrawn. Specify if the act occurs (a) *with smile*, if accompanied by smile, smirk, or laugh; or (b) *without smile*.

Refusal to Give Object or Food Item

- Child holds one item or one bowl (not both) close to body when the experimenter requests an object or food. Alternatively, child grabs one item or bowl quickly and holds it close to body, while looking directly at the experimenter.
- While holding one object or food item, the child tries to convince the experimenter to take the other object or food by labelling the latter as desirable.
- Child offers one item, then withdraws, but later gives the *same item*, or no item, to the experimenter (i.e., action does not qualify for "Offer/Withdrawal").
- Child holds one object or bowl when the experimenter requests an object or food and labels it as "mine".
- Specify item that is "refused", as well as if child later complies. Specify if the act occurs (a) *with smile*, if accompanied by a smile, smirk, or laugh; or (b) *without smile*.

Plays with One or Both Objects, or Eats One or Both Food Items

- Following the experimenter's request for one object, the child plays with one or both objects. Specify the object(s) played with (i.e., masculine, feminine, or both). If the child plays with both objects, also specify which object the child was engaged with longer.
- Following the experimenter's request for one food item, the child eats one or both food items. Specify which food item(s) the child eats (i.e., broccoli, crackers, or both). If the child eats both, indicate which food the child eats more of.

Gives One Object or Food Item Immediately

- Following the experimenter's request for one object or food item, the child immediately selects one object or food item to give to the experimenter.

Gives Both Objects or Food Items

- Following the experimenter's request for one object or food item, the child gives both objects or food items simultaneously to the experimenter.

Gives Second Item

- Following the experimenter's request for one object or food item, the child gives one object or food item to the experimenter first, then gives the second object or food item to the experimenter as well.

Touches Both Objects or Food Items

- Following the experimenter's request for one object or food item, the child touches one object or food item, then gives the other object or food item to the experimenter. Specify object or food item touched. If child touches both items before selecting one, indicate *Both*.

Task Refusal

- Following the experimenter's request for one object or food item, the child pushes away, throws, or otherwise rejects the object(s) or food item(s). Alternatively, the child does not give a object or food item, verbally states "no", or shakes his/her head. Specify if the child later complies.

Note. Some of the individual coding categories represented behaviours or emotional reactions that were found to be either too low in frequency or too high in frequency to discriminate among children. Thus, some low frequency categories were combined to form higher frequency categories for the purposes of the analyses. Other categories were excluded from the analyses. The variables included in the final analyses were: (a) humour during the affective display (i.e., before request, for desire trials only); (b) humour following request; (c) gaze at experimenter. This was coded if the child showed "monitoring gaze/gaze alternation" *or* "eye contact with the experimenter when object or food given" (these could be present with or without smile); and (d) refusal to give desired/gender-typed item. This was coded if any of the following behavioural responses was observed: "offer/withdrawal" of the item, "refusal to give object or food item", and "plays with one or both objects, or eats one or both food items" (these could be present with or without smile).

Appendix H

Correlations Between Children's Performance on the Warm-Up Task and Their Performance on the Food and Object Request Tasks

Correlations between Children's Performance on the Warm-Up Task and Their Performance on the Food and Object Request Tasks

For the food desire trials, children were given a score of 1 if they gave the desired food item on both trials and a score of 0 if they failed to give the desired food item on one or both trials. A nonsignificant point-biserial correlation was obtained between children's food request score and the percentage of warm-up trials successfully completed, $r(60) = .11$, *ns*. Moreover, no significant correlation was found between the percentage of warm-up trials successfully completed and the percentage of desired or gender-appropriate objects given to the experimenters on the desire and stereotyping trials, respectively, both $rs(64) = .05$, *ns*.

Appendix I

Means and Source Tables for the Object Request Task:

Analyses on the Whole Sample (N = 64)

Table 11

Mean Percentages of Appropriate Gender-Typed Objects Offered to the Experimenters (and Standard Deviations) During the Stereotyping Trials as a Function of Visit, Task Order, and Condition Order (N = 64)

Task Order	Visit 1		Visit 2	
	M	*SD*	*M*	*SD*
Desire trials first (n = 32)				
Female stereotyping first	57.81	19.83	48.44	21.35
Male stereotyping first	54.69	27.72	47.40	19.65
Across condition order	56.25	23.76	47.92	20.19
Stereotyping trials first (n = 32)				
Female stereotyping first	65.63	27.20	53.13	28.69
Male stereotyping first	48.96	30.56	47.40	25.59
Across condition order	57.29	29.69	50.26	26.90
Across task order (n = 64)				
Female stereotyping first	61.72	23.75	50.78	24.99
Male stereotyping first	51.82	28.85	47.40	22.44
Across condition order	56.77	26.68	49.09	23.62

Note. There were equal numbers of children in the two condition orders: female versus male stereotyping trials first.

Table 12

ANOVA Source Table for the Object Request Task: Preliminary Analysis to Assess for Order and Visit Effects on the Stereotyping Trials (N = 64)

Source	*df*	*F*	η^2
	Between subjects		
Task Order (T)	1	0.21	.00
Condition Order (C)	1	3.16 [1]	.05
T x C	1	1.49	.02
S within-group error	60	(446.15)	
	Within subjects		
Visit (V)	1	2.24	.04
V x T	1	0.02	.00
V x C	1	0.40	.01
V x T x C	1	0.19	.00
V x *S* within-group error	60	(842.56)	

Note. Values enclosed in parentheses represent mean square errors.

[1] $p < .10$

Table I3

ANOVA Source Table for the Object Request Task: Object Offered During the Stereotyping Trials

Source	*df*	*F*	η^2
	Between subjects		
Sex of Child (S)	1	0.92	.02
Age Group (A)	1	0.05	.00
Group (G)	1	2.18	.04
S x A	1	4.63*	.08
S x G	1	1.36	.02
A x G	1	0.02	.00
S x A x G	1	0.00	.00
S within-group error	56	(807.37)	
	Within subjects		
Condition (C)	1	2.28	.04
C x S	1	0.50	.01
C x A	1	0.60	.01
C x G	1	2.08	.04
C x S x A	1	0.06	.00
C x S x G	1	0.41	.01
C x A x G	1	0.01	.00
C x S x A x G	1	0.41	.01
C x *S* within-group error	56	(481.85)	

Note. Values enclosed in parentheses represent mean square errors. $^*p < .05$

Table I4

Mean Percentages of Desired Gender-Typed Objects Offered to the Experimenters (and Standard Deviations) During the Desire Trials as a Function of Visit, Task Order, and Condition Order (N = 64)

Task Order	Visit 1		Visit 2	
	M	*SD*	*M*	*SD*
Desire trials first (*n* = 32)				
Female desire first	43.75	26.61	60.94	25.77
Male desire first	51.04	27.70	53.13	22.13
Across condition order	47.40	26.98	57.03	23.96
Stereotyping trials first (*n* = 32)				
Female desire first	56.25	26.61	50.00	18.26
Male desire first	43.75	31.84	56.25	32.27
Across condition order	50.00	29.56	53.13	25.99
Across task order (*n* = 64)				
Female desire first	50.00	26.94	55.47	22.66
Male desire first	47.40	29.59	54.69	27.27
Across condition order	48.70	28.10	55.08	24.87

Note. There were equal numbers of children in the two condition orders: female versus male desire trials first.

Table I5

ANOVA Source Table for the Object Request Task: Preliminary Analysis to Assess for Order and Visit Effects on the Desire Trials

Source	*df*	*F*	η^2
	Between subjects		
Task Order (T)	1	0.02	.00
Condition Order (C)	1	0.12	.00
T x C	1	0.09	.00
S within-group error	60	(743.60)	
	Within subjects		
Visit (V)	1	1.89	.03
V x T	1	0.49	.01
V x C	1	0.04	.00
V x T x C	1	3.33 [t]	.05
V x *S* within-group error	60	(688.05)	

Note. Values enclosed in parentheses represent mean square errors.

[t] $p < .10$

Table I6

ANOVA Source Table for the Object Request Task: Object Offered During the Desire Trials

Source	*df*	*F*	η^2
	Between subjects		
Sex of Child (S)	1	1.12	.02
Age Group (A)	1	0.09	.00
Group (G)	1	0.04	.00
S x A	1	0.13	.00
S x G	1	1.24	.02
A x G	1	2.40	.04
S x A x G	1	0.54	.01
S within-group error	56	(734.99)	
	Within subjects		
Condition (C)	1	1.55	.03
C x S	1	1.55	.03
C x A	1	0.48	.01
C x G	1	0.64	.01
C x S x A	1	0.94	.02
C x S x G	1	0.74	.01
C x A x G	1	0.83	.02
C x S x A x G	1	1.17	.02
C x *S* within-group error	56	(708.01)	

Note. Values enclosed in parentheses represent mean square errors.

Appendix J

Analyses on the Emotional and Behavioural Responses During the Object Desire and Stereotyping Trials

Analyses on the Emotional and Behavioural Responses During the Object Desire and Stereotyping Trials.

For the object desire trials, a 2 (sex of child) x 2 (age group: 20 vs. 24 months) x 2 (group: gender consistent vs. gender inconsistent) between-groups MANOVA was first conducted on four emotional and behavioural responses to assess for sex, age, and group effects (Appendix J, Table J1). The mean percentage of trials in which smiling or laughing was displayed by the children during the experimenter's affective displays and following the experimenter's request for an item served as two of the dependent measures. The mean percentage of trials in which the child gazed at the experimenter while offering him or her an object and the proportion of trials in which the child refused to give the experimenter the desired object (e.g., offered and withdrew the desired object; played with the desired object) served as the remaining two dependent measures. A similar MANOVA was conducted on children's emotional and behavioural responses during the stereotyping trials (Appendix J, Table J2). However, the level of humour exhibited during the affective displays was not included as a dependent variable, as the experimenters did not express desires for the objects during the stereotyping trials. The three other dependent variables were the same. No significant main effects or interactions emerged from either of the MANOVAs performed.

To determine if children's emotional and behavioural responses to the object desire trials were related to their performance on this task, Pearson product-moment correlations were computed between each of the above dependent measures and the percentage of desired objects given to the experimenters during the desire trials. None of

the correlations were statistically significant (see Appendix J. Table J3, for *r* values). Correlations were also computed between the three emotional and behavioural responses used for the stereotyping trials above and the percentage of correct item selections made on the stereotyping trials. One child emerged as a significant outlier (greater than 4 *SD*s from the mean), refusing to give the experimenter the gender-appropriate object far more frequently than the other children in the sample ($M = 50.00\%$). To prevent this outlying score from unduly biasing the correlation performed, this child was assigned a value equal to 3 *SD*s from the mean. No significant correlations were found (see Appendix J. Table J3).

Table J1

Multivariate and Univariate Analyses of Children's Emotional and Behavioural Responses to the Object Desire Trials

Dependent Variable	Boys (n = 32)		Girls (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour during affective display	64.84	37.36	64.06	32.80	0.01	.00
Humour following request	55.08	37.55	63.28	31.74	0.92	.02
Gaze at experimenter	47.27	28.35	52.73	30.08	0.55	.01
Refusal to give desired item	7.03	10.01	10.55	15.26	1.13	.02
$F(4, 53) = 0.81$, $\eta^2 = .06$						

Dependent Variable	20-month-olds (n = 32)		24-month-olds (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour during affective display	56.64	36.34	72.27	32.02	3.19 ^{t}	.05
Humour following request	52.73	35.88	65.63	32.84	2.26	.04
Gaze at experimenter	46.88	23.76	53.13	33.75	0.72	.01
Refusal to give desired item	8.59	12.87	8.98	13.18	0.01	.00
$F(4, 53) = 0.95$, $\eta^2 = .07$						

Note. Table continued on following page.

$^t p < .10$

Dependent Variable	GC (n = 32)		GI (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour during affective display	66.41	35.70	62.50	34.49	0.20	.00
Humour following request	66.02	34.08	52.34	34.55	2.54	.04
Gaze at experimenter	47.66	30.85	52.34	27.58	0.40	.01
Refusal to give desired item	8.98	13.18	8.59	12.87	0.01	.00
$F(4, 53) = 1.00$, $\eta^2 = .07$						

Dependent Variable	20-month-olds (n = 32)		24-month-olds (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour during affective display					0.80	.01
Boys (n = 32)	53.13	38.32	76.56	33.50		
Girls (n = 32)	60.16	35.12	67.97	30.95		
Humour following request					0.05	.00
Boys (n = 32)	47.66	41.39	62.50	32.91		
Girls (n = 32)	57.81	29.89	68.75	33.54		
Gaze at experimenter					0.91	.02
Boys (n = 32)	47.66	20.01	46.88	35.50		
Girls (n = 32)	46.09	27.66	59.38	31.79		
Refusal to give desired item					2.35	.04
Boys (n = 32)	9.38	12.50	4.69	6.25		
Girls (n = 32)	7.81	13.60	13.28	16.75		
$F(4, 53) = 0.93$, $\eta^2 = .07$						

Note. Table continued on following page. GC = gender-consistent group; GI = gender-inconsistent group.

Dependent Variable	GC (n = 32)		GI (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour during affective display					1.79	.03
Boys (n = 32)	60.94	42.05	68.75	32.91		
Girls (n = 32)	71.88	28.32	56.25	35.94		
Humour following request					0.00	.00
Boys (n = 32)	61.72	39.91	48.44	35.02		
Girls (n = 32)	70.31	27.72	56.25	34.76		
Gaze at experimenter					0.91	.02
Boys (n = 32)	48.44	28.82	46.09	28.77		
Girls (n = 32)	46.88	33.70	58.59	25.71		
Refusal to give desired item					0.13	.00
Boys (n = 32)	7.81	10.08	6.25	10.21		
Girls (n = 32)	10.16	15.95	10.94	15.05		
$F(4, 53) = 0.89$, $\eta^2 = .06$						

Note. Table continued on following page. GC = gender-consistent group; GI = gender-inconsistent group.

Dependent Variable	GC (n = 32)		GI (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour during affective display					0.51	.01
20-month-olds (n = 32)	55.47	38.18	57.81	35.61		
24-month-olds (n = 32)	77.34	30.35	67.19	33.81		
Humour following request					1.30	.02
20-month-olds (n = 32)	54.69	37.05	50.78	35.79		
24-month-olds (n = 32)	77.34	27.47	53.91	34.38		
Gaze at experimenter					0.40	.01
20-month-olds (n = 32)	42.19	23.66	51.56	23.66		
24-month-olds (n = 32)	53.13	36.66	53.13	31.79		
Refusal to give desired item					0.35	.01
20-month-olds (n = 32)	7.81	12.81	9.38	13.31		
24-month-olds (n = 32)	10.16	13.86	7.81	12.81		
$F(4, 53) = 0.46$, $\eta^2 = .03$						

Note. Table continued on following page. GC = gender-consistent group; GI = gender-inconsistent group.

Dependent Variable	GC (n = 32)		GI (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour during affective display					0.03	.00
20-month-olds: Boys (n = 16)	45.31	44.79	60.94	31.65		
Girls (n = 16)	65.63	29.69	54.69	41.15		
24-month-olds: Boys (n = 16)	76.56	35.00	76.56	34.35		
Girls (n = 16)	78.13	27.35	57.81	32.69		
Humour following request					1.51	.03
20-month-olds: Boys (n = 16)	54.69	45.78	40.63	38.24		
Girls (n = 16)	54.69	29.08	60.94	32.35		
24-month-olds: Boys (n = 16)	68.75	34.72	56.25	32.04		
Girls (n = 16)	85.94	15.58	51.56	38.64		
Gaze at experimenter					1.35	.02
20-month-olds: Boys (n = 16)	42.19	16.28	53.13	22.90		
Girls (n = 16)	42.19	30.57	50.00	25.88		
24-month-olds: Boys (n =16)	54.69	37.76	39.06	33.70		
Girls (n = 16)	51.56	38.05	67.19	24.03		
Refusal to give desired item					0.01	.00
20-month-olds: Boys (n =16)	9.38	12.94	9.38	12.94		
Girls (n = 16)	6.25	13.36	9.38	14.56		
24-month-olds: Boys (n =16)	6.25	6.68	3.13	5.79		
Girls (n = 16)	14.06	18.22	12.50	16.37		
$F(4, 53) = 1.34$, $\eta^2 = .09$						

Note. GC = gender-consistent group: GI = gender-inconsistent group.

Table J2

Multivariate and Univariate Analyses of Children's Emotional and Behavioural Responses to the Object Stereotyping Trials

Dependent Variable	Boys (n = 32)		Girls (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour following request	51.17	34.69	57.42	38.48	0.49	.01
Gaze at experimenter	36.72	24.17	44.53	31.10	1.21	.02
Refusal to give gender-typed item	8.20	11.71	3.91	7.40	3.06 [1]	.05
$F(3, 54) = 1.62$, $\eta^2 = .08$						

Dependent Variable	20-month-olds (n = 32)		24-month-olds (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour following request	49.61	36.13	58.98	36.79	1.10	.02
Gaze at experimenter	38.28	26.17	42.97	29.77	0.44	.01
Refusal to give gender-typed item	7.42	10.93	4.69	8.84	1.24	.02
$F(3, 54) = 0.94$, $\eta^2 = .05$						

Dependent Variable	GC (n = 32)		GI (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour following request	59.38	35.64	49.22	37.15	1.29	.02
Gaze at experimenter	38.67	27.74	42.58	28.37	0.30	.01
Refusal to give gender-typed item	7.03	11.42	5.08	8.32	0.63	.01
$F(3, 54) = 0.73$, $\eta^2 = .04$						

Note. Table continued on following page. GC = gender-consistent group; GI = gender-inconsistent group. [1] $p < .10$

Dependent Variable	20-month-olds (n = 32)		24-month-olds (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour following request					0.38	.01
Boys (n = 32)	49.22	35.20	53.13	35.21		
Girls (n = 32)	50.00	38.19	64.84	38.52		
Gaze at experimenter					1.21	.02
Boys (n = 32)	38.28	21.64	35.16	27.09		
Girls (n = 32)	38.28	30.78	50.78	31.11		
Refusal to give gender-typed item					3.06 [t]	.05
Boys (n = 32)	11.72	13.28	4.69	8.98		
Girls (n = 32)	3.13	5.59	4.69	8.98		

$F(3, 54) = 1.39$, $\eta^2 = .07$

Dependent Variable	GC (n = 32)		GI (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour following request					3.06 [t]	.05
Boys (n = 32)	48.44	34.72	53.91	35.57		
Girls (n = 32)	70.31	34.12	44.53	39.26		
Gaze at experimenter					0.01	.00
Boys (n = 32)	34.38	22.59	39.06	26.17		
Girls (n = 32)	42.97	32.26	46.09	30.86		
Refusal to give gender-typed item					0.63	.01
Boys (n = 32)	10.16	13.86	6.25	9.13		
Girls (n = 32)	3.91	7.53	3.91	7.53		

$F(3, 54) = 1.23$, $\eta^2 = .06$

Note. Table continued on following page. GC = gender-consistent group; GI = gender-inconsistent group. [t] $p < .10$

Dependent Variable	GC (n = 32)		GI (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour following request					0.00	.00
20-month-olds (n = 32)	54.69	38.15	44.53	34.45		
24-month-olds (n = 32)	64.06	33.50	53.91	40.24		
Gaze at experimenter					0.59	.01
20-month-olds (n = 32)	33.59	28.03	42.97	24.14		
24-month-olds (n = 32)	43.75	27.39	42.19	32.87		
Refusal to give gender-typed item					0.03	.00
20-month-olds (n = 32)	8.59	13.48	6.25	7.91		
24-month-olds (n = 32)	5.47	9.09	3.91	8.80		
$F(3, 54) = 0.20$, $\eta^2 = .01$						

Note. Table continued on following page. GC = gender-consistent group; GI = gender-inconsistent group.

Dependent Variable	GC (n = 32)		GI (n = 32)			
	M	*SD*	*M*	*SD*	$F(1, 56)$	η^2
Humour following request					3.38 [t]	.06
20-month-olds: Boys (n = 16)	54.69	38.35	43.75	33.41		
Girls (n = 16)	54.69	40.61	45.31	37.76		
24-month-olds: Boys (n = 16)	42.19	32.00	64.06	36.86		
Girls (n = 16)	85.94	16.95	43.75	43.30		
Gaze at experimenter					0.98	.02
20-month-olds: Boys (n = 16)	29.69	11.45	46.88	26.52		
Girls (n = 16)	37.50	38.96	39.06	22.60		
24-month-olds: Boys (n = 16)	39.06	30.21	31.25	25.00		
Girls (n = 16)	48.44	25.39	53.13	37.65		
Refusal to give gender-typed item					0.03	.00
20-month-olds: Boys (n = 16)	14.06	16.95	9.38	8.84		
Girls (n = 16)	3.13	5.79	3.13	5.79		
24-month-olds: Boys (n = 16)	6.25	9.45	3.13	8.84		
Girls (n = 16)	4.69	9.30	4.69	9.30		
$F(3, 54) = 1.45$, $\eta^2 = .08$						

Note. GC = gender-consistent group: GI = gender-inconsistent group.

[t] $p < .10$

Table J3

Pearson-Product Moment Correlations Computed Between Children's Emotional and Behavioural Responses and the Percentage of Desired/Correct Objects Given to the Experimenters on the Desire and Stereotyping Trials

Trial Type	*n*	*r*
Desire Trials		
Humour during affective display	64	-.23 [t]
Humour following request	64	-.22 [t]
Gaze at experimenter	64	.09
Refusal to give desired object	64	.13
Stereotyping Trials		
Humour following request	64	-.14
Gaze at experimenter	64	-.09
Refusal to give gender-typed item	64	.04

[t] $p < .10$.

Appendix K

Source Table for the Object Request Task, Baseline Preference Trials: Analysis of Children's Performance on the Object Desire Trials as a Function of their Object Preferences

Table K1

ANOVA Source Table for the Object Request Task, Baseline Trials (n = 21): Analysis of Children's Performance on the Desire Trials as a Function of their Object Preferences

Source	*df*	*F*	η^2
	Between subjects		
Object Preference (T)	1	0.25	.01
S within-group error	19	(870.28)	
	Within subjects		
Condition (C)	1	0.48	.03
C x T	1	2.38	.11
C x *S* within-group error	19	(777.71)	

Note. Values enclosed in parentheses represent mean square errors.

Appendix L

Summary Tables for the Gender-Concept and Gender-Typing Measures

Table L1

ANOVA Source Table for the Gender-Labelling Task: Percentage of Correct Targets Selected as a Function of Picture Type, Age Group, and Sex of Child.

Source	*df*	*F*	η^2
	Between subjects		
Sex of Child (S)	1	0.12	.00
Age Group (A)	1	4.51*	.08
S x A	1	0.11	.00
S within-group error	53	(491.51)	
	Within subjects		
Picture Type (P)	1	23.03***	.30
P x S	1	0.29	.01
P x A	1	7.73**	.13
P x S x A	1	0.91	.02
P x *S* within-group error	53	(306.13)	

Note. Values enclosed in parentheses represent mean square errors.

*$p < .05$, **$p < .01$, ***$p < .001$

Table L2

Mean Percentage of Correct Targets Selected, Standard Deviations, and t Tests Against Chance for the Gender-Labelling Task, Child, Adult, and Total Scales for the 20- and 24-Month-Old Children Separately

Scale	*M*	*SD*	*df*	*t*
20-month-olds				
Child Scale	63.27	18.73	27	3.75 **
Adult Scale	70.00	24.28	27	4.36 ***
Total Scale	66.78	17.65	27	5.03 ***
24-month-olds				
Child Scale	63.22	19.10	28	3.73 **
Adult Scale	87.93	16.00	28	12.77 ***
Total Scale	75.52	12.75	28	10.78 ***

** $p < .01$. *** $p < .001$

Table L3

ANOVA Source Table for the Sex-Typed Child Care and Household Tasks Questionnaire: Children's Overall Gender-Typing Scores as a Function of Sex of Child and Age Group.

Source	*df*	*F*	η^2
	Between subjects		
Sex of Child (S)	1	1.09	.02
Age Group (A)	1	0.31	.01
S x A	1	0.78	.01
S within-group error	59	(0.04)	

Note. Values enclosed in parentheses represent mean square errors.

Table L4

Parents' Ratings of the Mean Frequencies With Which Their Child Played With, or Were Exposed to, Each of the Feminine- and Masculine-Typed Items Used in the Study on the Item Familiarity Questionnaire

	Mean frequency rating					
	Girls (n = 32)		Boys (n = 32)			
Item	*M*	*SD*	*M*	*SD*	*df*	*t*
Feminine-typed items						
Doll	3.53	1.16	2.78	1.21	62	2.53 *
Purse	3.00	1.24	1.88	0.94	62	4.08 ***
Necklace	2.69	1.03	2.16	1.17	62	1.93 †
Pink	2.77	0.70	1.58	0.54	62	7.63 ***
Masculine-typed items						
Car	3.03	0.91	3.94	0.72	61	- 4.39 ***
Baseball glove	1.19	0.54	1.66	0.94	62	- 2.46 *
Tie	1.50	0.95	1.66	1.04	62	- 0.63
Blue	3.00	0.49	3.66	0.60	62	- 4.78 ***

Note. Frequency ratings ranged from a score of 1 to a score of 5 (1 = *never*, 3 = *sometimes*, 5 = *always*). The child's exposure to pink and blue was determined by averaging the parent's ratings of the frequency with which the child wore pink or blue clothing, their ratings of the amount of pink or blue toys the child had at home, and their ratings for the child's bedroom decor (1 = *does not have pink/blue room*; 5 = *has pink/blue room*).

† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table L5

ANOVA Source Table for the Item Familiarity Questionnaire: Analysis of Children's Mean Gender-Typing Scores as a Function of Sex of Child and Age Group

Source	*df*	*F*	η^2
	Between subjects		
Sex of Child (S)	1	0.00	.00
Age Group (A)	1	0.66	.01
S x A	1	0.15	.00
S within-group error	60	(0.05)	

Note. Values enclosed in parentheses represent mean square errors.

Appendix M

Source Tables For the Object Request Task: Analyses on the Subsample of Children Using the Food Request Task and the Gender-Labelling Task as Screening Measures (n = 31)

Table M1

Mean Percentages of Gender-Stereotyped Objects Correctly Offered to the Experimenters (and Standard Deviations) During the Stereotyping Trials as a Function of Visit, Age Group, and Experimental Group (n = 31)

Age Group	Visit 1		Visit 2	
	M	*SD*	*M*	*SD*
20-month-olds				
Gender-consistent group (*n* = 6)	62.50	30.62	37.50	26.22
Gender-inconsistent group (*n* = 5)	60.00	28.50	45.00	20.92
Across groups (*n* = 11)	61.36	28.20	40.91	23.11
24-month-olds				
Gender-consistent group (*n* = 10)	67.50	28.99	50.00	24.85
Gender-inconsistent group (*n* = 10)	62.50	17.68	52.50	21.89
Across groups (*n* = 20)	65.00	23.51	51.25	22.83
Across age groups				
Gender-consistent group (*n* = 16)	65.63	28.69	45.31	25.27
Gender-inconsistent group (*n* = 15)	61.67	20.85	50.00	21.13
Across groups (*n* = 31)	63.71	24.87	47.58	23.09

Table M2

ANOVA Source Table for the Object Request Task: Preliminary Analysis to Assess for Visit, Group, and Age Effects on the Stereotyping Trials

Source	*df*	*F*	η^2
	Between subjects		
Age Group (A)	1	1.39	.05
Group (G)	1	0.01	.00
A x G	1	0.10	.00
S within-group error	27	(480.58)	
	Within subjects		
Visit (V)	1	5.32 *	.17
V x A	1	0.18	.01
V x G	1	0.36	.01
V x A x G	1	0.01	.00
V x *S* within-group error	27	(755.28)	

Note. Values enclosed in parentheses represent mean square errors.

*$p < .05$

Table M3

Mean Percentages of Feminine-Stereotyped Objects Offered to the Male and Female Experimenters (and Standard Deviations) During the Stereotyping Trials as a Function of Sex of Child, for Visit 1 and Visit 2 Separately (n = 31)

Condition	Visit 1		Visit 2	
	M	*SD*	*M*	*SD*
Female Experimenter				
Boys (n = 8)	68.75	22.16	56.25	22.16
Girls (n = 6)	62.50	13.69	37.50	20.92
Total (n = 14)	66.07	18.62	48.21	22.92
Male Experimenter				
Boys (n = 9)	36.11	25.34	43.52	29.10
Girls (n = 8)	40.63	35.20	44.79	16.63
Total (n = 17)	38.24	29.47	44.12	23.34
Across Conditions				
Boys (n = 17)	51.47	28.60	49.51	26.10
Girls (n = 14)	50.00	29.42	41.67	18.20
Total (n = 31)	50.81	28.49	45.97	22.86

Table M4

ANOVA Source Table for the Object Request Task: Object Offered During the Visit 1 Stereotyping Trials

Source	*df*	*F*	η^2
	Between subjects		
Sex of Child (S)	1	0.01	.00
Condition (C)	1	8.36 **	.24
S x C	1	0.33	.01
S within-group error	27	(673.55)	

Note. Values enclosed in parentheses represent mean square errors.

** $p < .01$

Table M5

ANOVA Source Table for the Object Request Task: Object Offered During the Visit 2 Stereotyping Trials

Source	*df*	*F*	η^2
	Between subjects		
Sex of Child (S)	1	1.09	.04
Condition (C)	1	0.11	.00
S x C	1	1.43	.05
S within-group error	27	(530.95)	

Note. Values enclosed in parentheses represent mean square errors.

Table M6

Mean Percentages of Desired Objects Offered to the Experimenters (and Standard Deviations) During the Desire Trials as a Function of Visit and Age Group (n = 31)

Age Group	Visit 1		Visit 2	
	M	*SD*	*M*	*SD*
20-month-olds (n = 11)	62.88	34.03	56.82	16.17
24-month-olds (n = 20)	50.42	23.95	62.50	27.51
Total (n = 31)	54.84	28.04	60.48	23.96

Table M7

ANOVA Source Table for the Object Request Task: Preliminary Analyses to Assess for Visit and Age Effects on the Desire Trials

Source	*df*	*F*	η^2
	Between subjects		
Age Group (A)	1	0.22	.01
S within-group error	29	(744.13)	
	Within subjects		
Visit (V)	1	0.21	.01
V x A	1	1.89	.06
V x *S* within-group error	29	(616.78)	

Note. Values enclosed in parentheses represent mean square errors.

Table M8

Mean Percentages of Feminine-Stereotyped Objects Offered to the Male and Female Experimenters (and Standard Deviations) During the Desire Trials as a Function of Sex of Child and Experimental Group (n = 31)

Experimental Group	Female Experimenter		Male Experimenter	
	M	*SD*	*M*	*SD*
Gender-consistent group				
Boys (n = 9)	63.89	18.16	45.37	20.03
Girls (n = 7)	32.14	34.50	55.95	22.93
Total (n = 16)	50.00	30.28	50.00	21.30
Gender-inconsistent group				
Boys (n = 8)	28.13	20.86	65.63	22.90
Girls (n = 7)	42.86	27.82	67.86	27.82
Total (n = 15)	35.00	24.64	66.67	24.40
Across groups				
Boys (n = 17)	47.06	26.34	54.90	23.21
Girls (n = 14)	37.50	30.62	61.91	25.26
Total (n = 31)	42.74	28.28	58.06	24.01

Table M9

ANOVA Source Table for the Object Request Task: Object Offered During the Desire Trials (n = 31)

Source	*df*	*F*	η^2
	Between subjects		
Group (G)	1	0.07	.00
Sex of child (S)	1	0.03	.00
G x S	1	2.08	.07
S within-group error	27	(670.71)	
	Within subjects		
Condition (C)	1	8.47 **	.24
C x G	1	6.03 *	.18
C x S	1	1.64	.06
C x G x S	1	5.54 *	.17
C x *S* within-group error	27	(520.07)	

Note. Values enclosed in parentheses represent mean square errors.

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

Table M10

ANOVA Source Table for the Object Request Task: Object Offered During the Desire Trials. Analysis for Boys Only (n = 17)

Source	*df*	*F*	η^2
	Between subjects		
Group (G)	1	0.92	.06
S within-group error	15	(556.46)	
	Within subjects		
Condition (C)	1	2.71	.15
C x G	1	23.62 ***	.61
C x *S* within-group error	15	(281.38)	

Note. Values enclosed in parentheses represent mean square errors.

*** $p < .001$.

Table M11

ANOVA Source Table for the Object Request Task: Object Offered During the Desire Trials. Analysis for Girls Only (n = 14)

Source	*df*	*F*	η^2
	Between subjects		
Group (G)	1	1.10	.08
S within-group error	12	(813.51)	
	Within subjects		
Condition (C)	1	5.09 *	.30
C x G	1	0.00	.00
C x *S* within-group error	12	(818.44)	

Note. Values enclosed in parentheses represent mean square errors.

* $p < .05$.

Appendix N

MANOVA on Measures of Gender Knowledge and Exposure to Gender Stereotyping in the Home for the Children in the Subsample (n = 31)

Description of MANOVA Conducted on the Measures of Children's Gender Knowledge and Exposure to Gender Stereotyping in the Home

A 2 (sex of child) x 2 (group: gender consistent vs. gender inconsistent) between-groups MANOVA was performed. The percentage of correct selections made on the Visit 1 stereotyping trials and the mean percentage of correct pictures selected on the gender-labelling task (total scale) served as two of the dependent measures. The third dependent measure assessed children's exposure to gender stereotyping in the home. A single gender-typing measure was created by averaging children's mean gender-typing scores on the Item Familiarity Questionnaire, which assessed the extent to which children preferred, or were exposed to, gender-stereotyped toys and objects, and parents' mean gender-typing scores on the Sex-Typed Child Care and Household Tasks Questionnaire (SCCHT). Children's mean gender-typing scores on the Item Familiarity Questionnaire ranged from 0 (*non-gender-typed*) to 1 (*gender-typed*), whereas parents' mean gender-typing scores on the SCCHT ranged from 1 to 5, with 1 representing *highly traditional role-division* in the home, and a score of 5, *highly nontraditional role-division*. To computed an average gender-typing score based on these two measures, parents' mean gender-typing scores on the SCCHT were rescaled to make them equivalent with the Item Familiarity Questionnaire gender-typing scores, such that a score of 0 represented *nontraditional role-division* in the home and a score of 1, *traditional role-division*.

Table N1

MANOVA Summary Table: Multivariate and Univariate Analyses of the Gender-Concept and Gender-Typing Measures for the Object Desire Trials for the Children in the Subsample, in Relation to Sex of Child and Experimental Group (n = 31)

Dependent Variable	Boys (n = 17)		Girls (n = 14)			
	M	*SD*	*M*	*SD*	$F(1, 27)$	η^2
Visit 1 stereotyping trials, % correct given	66.18	23.29	60.71	27.24	0.32	.01
Gender-labelling task, total score	79.90	9.34	79.76	7.81	0.00	.00
Exposure to gender stereotyping	0.65	0.09	0.63	0.11	0.26	.01
$F(3, 25) = 0.20$, $\eta^2 = .02$						

Dependent Variable	GC (n = 16)		GI (n = 15)			
	M	*SD*	*M*	*SD*	$F(1, 27)$	η^2
Visit 1 stereotyping trials, % correct given	65.63	28.69	61.67	20.85	0.14	.01
Gender-labelling task, total score	80.21	6.72	79.44	10.38	0.05	.00
Exposure to gender stereotyping	0.65	0.10	0.64	0.10	0.02	.00
$F(3, 25) = 0.06$, $\eta^2 = .01$						

Note. Table continued on following page. GC = gender-consistent group; GI = gender-inconsistent group.

Dependent Variable	GC (n = 16)		GI (n = 15)			
	M	*SD*	*M*	*SD*	$F(1, 27)$	η^2
Visit 1 stereotyping trials. % correct given					0.14	.01
Boys (n = 17)	69.44	20.83	62.50	26.73		
Girls (n = 14)	60.71	37.80	60.71	13.36		
Gender-labelling task. total score					0.05	.00
Boys (n = 17)	80.56	7.22	79.17	11.79		
Girls (n = 14)	79.76	6.56	79.76	9.45		
Exposure to gender stereotyping					0.01	.00
Boys (n = 17)	0.65	0.08	0.65	0.11		
Girls (n = 14)	0.64	0.13	0.63	0.10		

$F(3, 25) = 0.05$, $\eta^2 = .01$

Note. GC = gender-consistent group: GI = gender-inconsistent group. Of the 17 boys. 9 were in the gender-consistent group and 8 were in the gender-inconsistent group. Of the 14 girls. 7 were in the gender-consistent group and 7 were in the gender-inconsistent group.