ABSTRACT

Title of Document: CHILDREN'S MORAL REASONING ABOUT ATTRIBUTION OF INTENTIONS: THE INFLUENCE OF GENDER STEREOTYPES AND THEORY OF MIND

Megan Clark Kelly, Ph.D., 2011

Directed By: Professor Melanie Killen, Department of Human Development

The present study investigated how holding gender stereotypes and having a false belief theory of mind impacts children's understanding of intentionality when evaluating morally relevant stories. Children 3 – 4, 5 – 6, and 7 – 8-years-of-age (N = 127) were interviewed about the intentions of a potential transgressor in two hypothetical stories. Both stories involved a child accidentally (or on purpose) putting another child's toy into their own backpack. One of the stories utilized the taking of a toy that was gender stereotype consistent (a girl taking a doll) while the other story involved a gender stereotype inconsistent toy (a boy taking a doll or a girl taking a truck). A false belief theory of mind task as well as gender stereotype knowledge, tolerance, and flexibility tasks were administered to each participant.

Results revealed that children over-attributed negative intentions and endorsed more punishment in the story with the counter-stereotypic toy than in the story with the stereotype consistent toy, indicating that stereotypes were impacting the children's decisions concerning intentionality. Additionally, across scenarios, older children as well as children able to pass the false belief theory of mind task, endorsed less punishment and indicated less negative intentions than their counterparts, demonstrating that as children get older and more cognitively advanced they are better able to see the ambiguity of a morally relevant scenario, despite gender stereotypes, in order to attribute less negative intentions. Furthermore, children who were aware of gender stereotypes and children who were tolerant of others playing with any toy regardless of the associated gender stereotype also endorsed less punishment and indicated less negative intentions than their counterparts.

The present study therefore shows how children may erroneously focus on stereotypic knowledge when making attributions of intentionality. This is important as over-attributing negative intentions can lead to peer rejection and exclusion. Understanding when and how contextual variables such as gender stereotypes as well as when and how having a false belief theory of mind impacts attributions of intentions is critical to understanding the ontogeny and development of moral reasoning.

CHILDREN'S MORAL REASONING ABOUT ATTRIBUTION OF INTENTIONS: THE INFLUENCE OF GENDER STEREOTYPES AND THEORY OF MIND

By

Megan Clark Kelly

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:
Professor Melanie Killen, Chair
Professor Natasha Cabrera
Professor Geetha Ramani
Professor Meredith Rowe
Professor Deborah Speece, Dean's Representative

Acknowledgements

There has been a tremendous amount of work that has gone into earning my Ph.D., and I have, by no means, worked alone. I need to start by thanking my family, for they were there from the start and are here at the end, and are the reason I am able to do what I do. I first have to thank Christopher. Finishing a Ph.D. 9 months pregnant with two children is not an easy feat, and would have been an impossible feat without the support of a true partner. Thank you for all the weekends, nights, and days that you stayed with the boys so that I could work. Thank you for adjusting your schedule so that mine could work. Thank you for believing that I could finish when I did not. Thank you for not giving up on me, and knowing that the crazy would only last so long. Thank you for drying my tears and helping me stay motivated. Thank you for your enthusiasm in my every achievement and your interest (even if feigned) in my work all along the way. Thank you! I also need to thank my beautiful children Gavin, Colin, and little Catherine (thank you for waiting until 7 days after I defended to come into the world). Thank you for understanding those days and nights, bath times and dinners that I missed and for knowing that I loved you.

I owe a zealous thank you to parents for instilling in me the value of education, showing me that it was my most important job, and ensuring that I could continue on with my studies and pursue my passion. Thank you for your constant assurance that I could accomplish anything that I desired if I set a goal and worked toward it, and for your undying belief in me. I would especially like to thank my mother for showing me what a strong, independent, intelligent woman with a family looks like, how she conducts herself, and how she sets priorities; without your

example, I never would have had the strength to know I could do this. Thank you also to my sisters Sara and Katie for being my best friends, and for showing by example how one of the Clark girls can be intelligent, beautiful, and graceful in life and love. Thank you to my entire family of aunts, uncles, cousins, and in-laws. There is no other group of people I would rather be associated with, and no other group of people who have ever been more supportive of someone. I love you all.

I also need to thank my University of Maryland family. Thank you Melanie for your endless advice over the years, for your thousands of re-writes that have made all of my work stronger, for the hours you have put into my edification, and for not giving up on me. Thank you also to my committee members (Natasha, Meredith, Geetha, and Dr. Speece), for taking time out of your *very* busy schedules to provide the guidance, direction, and assistance I needed to successfully complete this project. I also want to take this opportunity to thank my lab mates past and present for all the hours of working together that have made me a better scientist and helped me in endless ways throughout my many years here. Thank you especially to Alaina Brenick, who was there at the start and who now has a permanent home in my heart and to Cam Richardson for being such an amazing and inspirational work partner. Thank you also to Nancy Margie and Stefanie Sinno for all the advice, academic and otherwise, that has been endlessly helpful. Thank you to Kelly Lynn and Aline for all your help and collaboration on this dissertation project, and of course to my wonderful undergraduate team of Nancy Paik, Kevin McGrath, Esther Wei, and Madi Moin, who were always willing and capable to do all the work I handed off to you,

you are the best! Thank you everyone! I look forward to more years of all your help and support, and some reciprocal help and support from me as well!

Finally, I would like to thank all of the teachers, administrators, and parents who trusted me to interview and spend time with their children, and to the children for allowing me a window into your thoughts and minds, what an amazing gift you all gave me! Without all of you, there would be no study, and no inspiration for further study, so thank you!

Table of Contents

List of Tables

List of Figures

Chapter 1: Theoretical Rationale

Introduction

Children's moral judgments and reasoning have an undeniable and critical impact on their social development from a very early age. Children form concepts of fairness and equality as early as 2 ½-years-of-age, and are able to apply these conceptions to their social experiences. These concepts of fairness and equality, which are rudimentary to moral reasoning, combine with conceptions of authority, tradition, social norms, and personal preferences to constitute organized systems of a child's overall social knowledge that are cultivated from children's experiences in their social world (Smetana, 2006; Turiel, 1983, 1998). These systems of social knowledge are used by children to reason about or make judgments in their social world. Concurrently, children's organization and use of these systems of social knowledge facilitate our understanding of their social reasoning process.

The social cognitive domain model, a theoretical model of social reasoning, allows for researchers to explore attribution of intentionality and decision making about familiar events, and then evaluate what children are reasoning about when making those decisions. More specifically, it uniquely permits researchers to investigate when children are taking stereotypes into account when making a decision about intentionality. This is critically important in gaining an understanding of the ontogeny, formation and the developmental trajectories of prejudice (the tendency to attribute negative characteristics to outgroup members), stereotyping (making judgments about an individual's traits or behaviors based on group membership), and

discrimination (differential treatment based on biased beliefs about one's group membership) (Killen, Richardson, & Kelly, 2010).

Social Cognitive Domain Model

The Social Cognitive Domain Model is a model of social reasoning that has identified three conceptually distinct domains of social reasoning used by individuals from early childhood to adulthood; moral (fairness, justice, equality), social conventional (authority, tradition, social norms), and psychological (personal). The social domain model proposes that individuals apply different forms of reasoning to a range of situations (Smetana, 2006; Turiel, 1983, 1998). The model distinctively maintains that individuals utilize one of these domains when understanding and making decisions about their social worlds, including decisions made in intergroup situations (Horn, 2003; Killen, Richardson, & Kelly, 2010; Smetana, 2006).

Additionally, the social cognitive domain model contends that individuals may apply reasons from one domain (moral, social conventional, or psychological) or more than one domain (e.g., both moral and personal) to a given situation, and that the judgments made may include interpretations of specific features of the situation (Killen, McGlothlin, & Lee-Kim, 2002a). When individuals apply reasons from one domain alone, it is considered straightforward. For example, nearly all children will say that hitting is wrong, and use moral reasoning concerning harm for others as their justifications (Killen, McGlothlin, & Lee-Kim, 2002b). In contrast, when a situation is complex and individuals combine multiple domains, moral, social conventional, or personal, it is considered multifaceted. Reasoning about issues such as stereotypes typically fall into the multifaceted domain as they can be accepted or rejected for

reasons related to fairness (moral) or accepted because of a need for a group to function optimally (social conventional) or because of a personal choice (psychological). It is proposed that when individuals evaluate complex, multifaceted acts and issues, they weigh different considerations and give priority to one perspective, or form of reasoning, over another.

The social cognitive domain model additionally proposes that social reasoning varies by context and that in a given situation individuals have to assess the multiple dimensions often present in a context in order to make an evaluation. The model therefore stipulates that it is critical for researchers to analyze experimental situations in terms of both the components of the context as well as the predictions about how individuals will analyze it. The model provides guidance with the explication of moral components (e.g., issues of fairness, justice, or rights), societal components (e.g. customs and cultural expectations), and personal components (e.g., personal choice, privacy, intimate relationships), with the addition of informational assumptions (judgments about reality, the nature of learning, etc), which invariably enter into the evaluations of social contexts (Killen, McGlothlin, et al., 2002a).Cultivating a stronger understanding of the ontogeny and development of moral reasoning is essential to understanding children's social development in general, and more specifically it is critical to understanding the judgments and decisions that children are making, and the reasoning behind these decisions.

As early as 3, children can distinguish moral and conventional issues from personal ones in both the home and at preschool (Killen & Smetana, 1999; Nucci & Weber, 1995; Weber, 1993). At the same time, children of the same age consistently

apply moral criteria to events entailing physical harm (hitting or hurting) (Smetana, 1981), and eventually move on to include psychological harm (Smetana, 1993) as well. In middle childhood, children's understanding of the moral domain expands from one of concrete harm and others' welfare to a more complex understanding of fairness defined in terms of equality and equal treatment (Helwig, 1990; Nucci, 2001). During pre-adolescence, conceptions of fairness expand even further to envelop more and more complex social situations concurrently involving both moral issues and non-moral issues in the social conventional and personal domains. Additionally, at this time, adolescents begin to, among other things, prioritize issues of group functioning, over other moral and social concerns (Horn, 2003; Killen, Lee-Kim, McGlothlin, & Stangor, 2002). Knowing more about how children and adolescents are using the innumerable possible contextual considerations when reasoning about social and moral issues will deepen our understanding of moral and social reasoning.

Study Rationale

Morality, Intentionality, and Theory of Mind

Central to, and inextricable from the moral reasoning about an action or behavior, is the intentionality that motivated the action. This is because behaviors alone do not reveal the moral status of an action; it is necessary to know what the actor intended to do, and whether the intentions were positive or negative towards another individual. In fact, the morality of a behavior or action is often judged by the intentions of the actor (Leslie, Knobe, & Cohen, 2006). To fully understand children's moral reasoning and how children bring their moral concepts to bear on

their social interactions and conflicts, it is critical to understand how children's ability to interpret intentionality impacts their moral judgments and justifications (Wainryb & Brehl, 2006). The study of intentionality is central to the study of theory of mind, or the understanding that others have beliefs, intentions, and desires different from one's own (Astington, Harris, & Olson, 1988; Leslie, Knobe, & Cohen, 2006).

There is a tremendous amount of research delving into the myriad intricacies of the field of theory of mind research starting in early infancy going up to and beyond middle childhood (Wellman & Liu, 2004; Woodward & Needham, 2009). Wellman and Liu (2004) define theory of mind as children's understanding of other people's mental states (pp. 523) and consolidated the wealth of theory of mind knowledge into an age related progression of the different types of theory of mind understandings such as the understanding of desires versus beliefs and ignorance versus false beliefs. A central task developed in the 1980s is referred to as the "false belief task" which assesses children's ability to utilize their knowledge of another person's belief states to predict his or her subsequent actions when those beliefs differ from reality and from the child's own knowledge (Wimmer & Perner, 1983). For example, a box that has pictures of chocolates on the outside is shown to a participant. The participant is asked what they believe is on the inside, and nearly all children state chocolates. The box is then opened to show that it in fact has pencils on the inside, not chocolates. The box is then closed and the child is asked what another child, who has never opened nor seen the box, will think is inside. Typically, children younger than 4-years-of-age fail this task as they predict the person's actions based on reality (the other child will believe that pencils are inside the box) rather

than the person's false belief. Children begin to pass this task around 5-years-of-age, distinguishing between what is reality, and the false belief that the child who has not seen the inside of the box will have, thereby displaying an understanding of false belief theory of mind (Wellman, Cross, & Watson, 2001).

Other researchers, such as Karpinski and Scullin (2009), have delved into the relationships between theory of mind and other cognitive capacities such as executive functioning, and their effects on social processes such as the ability to influence a preschooler during an interview. These very different arenas of theory of mind research add greatly to the field in their own ways, and can also add to our general understanding of theory of mind and its effects on social decision making, such as the decisions around intentionality. Despite this extensive surge of research on theory of mind, there is surprisingly little research integrating theory of mind and morality (for exceptions see: Helwig, Zelazo, & Wilson, 2001; Killen, Mulvey, Richardson, Jampol, & Woodward, under review; Leslie, et al., 2006; Wainryb & Brehl, 2006; Zelazo, Helwig, & Lau, 1996).

Intentionality is a sphere where theory of mind and moral judgment can meet (Leslie, et al., 2006). Intentions though, cannot be known by the action alone, unless explicitly stated by the actor. Intentions must be inferred from contextual information such as the time, place, objects, other involved individuals, information about the actor such as beliefs, desires, stereotypes, past behaviors, in addition to the information about the action itself (Malle, Moses, & Baldwin, 2001). The contextual information therefore can play a critical role in the determination of intentionality.

Historically, the founder of the field of developmental psychology, Jean Piaget, a Swiss psychologist, observed that the moral reasoning of the developing child began to reflect an understanding of intentions around 10 years of age (Piaget, 1932/1965), while more recent research indicates that children can do this much earlier (Smetana, 1995; Turiel, 1983). In fact, some researchers have been able to determine that children as young as 14 – 18 months of age can differentially imitate, and thereby differentiate, intentional and accidental actions (Carpenter, Akhtar, & Tomasello, 1998). Recently, researchers studying morality and theory of mind have converged on demonstrating when theory of mind competence is related to moral judgment, and how moral judgment can be influenced by a lack of theory of mind ability (Baird & Astington, 2004; Lagattuta, 2005; Leslie, et al., 2006; Wainryb & Brehl, 2006). Killen and colleagues have been able to determine that while children can discern intentionality in a standard false belief theory of mind task by 4- to 5-years-of-age, children have a harder time discerning intentionality in a morally relevant hypothetical scenario (Killen et al., under review).

Challenges within the Theory of Mind Research

The field of theory of mind research is so vast and varied, that utilizing it to inform research about morality and intentionality is challenging. Primarily, of the few studies that do connect morality and theory of mind research, very few have a standardized methodology of examining morality, and even fewer examine morality and theory of mind within the same task. This makes it difficult for comparisons to be made across the various studies, and allows for only correlational relationships between morality and theory of mind to be examined, and limits the ability to

determine how theory of mind competence bears on moral judgments. Furthermore, the vast majority of theory of mind measures provide either very limited or no contextual information. In the prototypical theory of mind false belief measure, (e.g., Mark is using markers on the art table. Mark goes outside. Another person comes in the room and puts the markers in a cabinet. When Mark comes back inside where will he look for the markers?), no social information is provided. The individual taking the task does not know who the markers belong to, is not expected to have any emotional reaction to the markers, does not know why the markers were moved, and does not know the relationship between the two individuals. The missing social information would likely be utilized by the children taking the theory of mind task in formulating their decisions concerning the outcome of events. In fact, research by Killen and colleagues has shown that contextual information in hypothetical scenarios, such as relationships of the actors in the scenario, the social group that a person belongs to, the intergroup nature of a scenario, or the skin color of an actor in the scenario are all used to make decisions concerning the actors in the scenario (Horn, Killen, & Stangor, 1999; Killen, Kelly, Richardson, & Jampol, in press; McGlothlin & Killen, 2006).

Killen and colleagues addressed many of these concerns when conducting a study that embedded a false belief theory of mind task within a morally relevant and standardized scenario (Killen, et al., under review). The authors interviewed children from 4- to 8-years-of-age about a series of tasks, one prototypical moral transgression, (e.g., a child gets pushed off a swing by another child who wants to use it), one prototypical theory of mind problem (e.g., crackers placed in an empty crayon

box when no one else is in the room), and a problem that embedded a theory of mind belief in a morally relevant hypothetical scenario, a morally relevant theory of mind task.

In the morally relevant theory of mind task, the participants were read a vignette involving a child who is helping the teacher clean up the classroom while the rest of the class is outside at recess, and accidentally throws away another child's special cupcake. This vignette is referred to as the "accidental transgressor paradigm", and provides a great deal of social information. The relationship between the actors is known, the object of interest is a cupcake, which is able to evoke a much stronger emotional reaction and connection then markers, and there is an identified victim, the child whose cupcake is thrown away. Additionally, this scenario is still an extension of the theory of mind paradigm. Killen and colleagues measured children's false belief theory of mind, moral judgments, and morally relevant theory of mind with children, including an attribution of intentions of the potential transgressors' actions ("do you think [the potential transgressor] did something all right or not all right?"), the acceptability of punishing the potential transgressor for his/her actions ("do you think [the potential transgressor] should get in trouble?").

In the prototypical false belief theory of mind task, the participants are asked predict what a child will expect to find inside a crayon box if a child did not see the crackers put into the crayon box. In contrast, in the morally relevant theory of mind task developed by Killen and colleagues (2009), the participants were asked to predict what the actor who is cleaning up the room will expect to find inside the paper bag if the actor did not see the cupcake in the bag. The questions extended past the

questions typically seen in a theory of mind task as the participants were asked to make evaluations, judgments, and justifications of intentionality, punishment acceptability, and attribution of emotions (Killen, Mulvey, et al., in press).

The main findings from this study were that children without false belief theory of mind competency were more likely to attribute negative intentions and found it more acceptable to punish children in the morally relevant hypothetical scenario than children with theory of mind, and that children attributed negative emotions and intentions to the accidental transgressor up until 8 – years-of-age (Killen, Mulvey, et al., in press). This study made it clear for the first time, that integrating moral judgment and false belief theory of mind poses challenges to children. One possible outcome of this error in judgment is over-attributing negative intentions to other children, which has been found to result in peer rejection and exclusion (Crick & Dodge, 1994). Over-attributing negative intentions occurs when an accidental transgressor, for example, is assumed to have negatively motivated intentions.

This study by Killen and colleagues opens a door into what is impacting a child's ability to make an attribution of intentionality. It is clear that while a 4- or 5-year-old can correctly distinguish intentionality in a straightforward scenario involving no emotional valence or moral concerns, the addition of a moral premise adds a layer of complexity to the scenario that makes it more difficult to correctly decipher intentionality. What is not known is what other aspects of a scenario may also contribute to increasing difficulty or greater ease in the ability to discern intentionality. Adding more contextual features, such as a highly gender stereotyped

toy, instead of an object which elicits no emotional reaction, to a scenario aimed at examining a child's ability to attribute intentionality will allow for a more thorough investigation of the impact of those contextual features.

Killen and colleagues found that the fundamental determinations of intentionality, a core aspect of moral and social reasoning, is affected by the introduction of a moral premise to a straightforward false belief theory of mind scenario, which falls into line with previous social domain research indicating that when children and adolescents encounter complex scenarios, they use more reasoning about rules, conventions, and group norms, or social conventional reasoning (Killen, Mulvey, et al., in press). These new data add a wealth of information concerning what children are taking into account when deciding if someone did something on purpose, or when making an attribution of intentionality, and additionally these data inform us that these determinations of intentionality are context dependent. There is a paucity of research evaluating what other aspects of the context, (social relationships, relationships between the objects of interest and the actors, the personal characteristics of the actors), are affecting the decisions surrounding attribution of intentionality. In the current study, morality and theory of mind will be assessed in the context of situations involving gender stereotypic expectations. Before describing the design of the study, the literature on stereotype knowledge and intergroup attitudes will be reviewed.

Morality and Intergroup Attitudes

There has been a great deal of previous social domain research indicating that group identification, intergroup interactions, and stereotypes all impact exclusion

decisions (Horn, 2003; Killen, Sinno, & Margie, 2007; Smetana, 2006). It is important to explore whether this impact of intergroup interactions, attitudes, and stereotypes will impact decisions surrounding attributions of intentionality, which was a central goal of the present study.

Social Domain theorists have been researching the contextual information that impacts moral reasoning for several decades (for reviews see: Killen, et al., 2010; Smetana, 2006), and have found that intergroup attitudes strongly impact an individual's use of moral reasoning (see Killen, Richardson, & Kelly, 2010 for examples). Intergroup research has been ongoing for decades. Sherif (1966) defined intergroup behavior as, "Whenever individuals belonging to one group interact, collectively or individually, with another group or its members in terms of their group identification, we have an instance of intergroup behavior" (pp. 12). In order for there to be groups in this intergroup behavior however, it is necessary that there is an external recognition that the group exists as well as an internal sense of awareness of group membership, an awareness that the group membership has a value associated to it, and some emotional investment in that awareness (Tajfel, 1981). Intergroup attitudes, therefore, are the evaluations and emotional investment that individuals apply to their own group as well as to other groups.

When including intergroup categories in assessments of straightforward moral transgression, research has shown that the vast majority of children and adolescents view it as wrong and unfair (Killen, Margie, & Sinno, 2006). For example, Horn (2003) demonstrated that adolescents evaluated the denial of resources to individuals based on group membership, such as belonging to a clique, as wrong, while Killen

and colleagues additionally found that nearly 95% of children and adolescents rejected gender and racial exclusion in the context of denying educational access to others (Killen, Lee-Kim, et al., 2002). At the same time, Horn and Killen noted that in more complex, multifaceted situations, stereotypical expectations and increased reliance on group functioning, with age, functioned as the basis for exclusion and supplanted moral concerns for fairness (Horn, 2003; Killen, Lee-Kim, et al., 2002).

Intergroup Attitudes and Attribution of Intentions

Knowing that children and adolescent's reasoning about intergroup interactions changes with age and with the complexity of the context, it is expected that intergroup interactions might also impact a child' or adolescent's attribution of intentionality. One way to explore the impact of intergroup attitudes on children's interpretation of intentionality is by creating tasks in which children have to judge the intentionality of another child in a scenario that invokes stereotypic associations or stereotype knowledge. We know that children and adolescents have been shown to have stereotypes that become more entrenched with age (Stangor & Schaller, 1996), and children and younger adolescents have been shown to make accusations based on stereotypic assumptions with less available information than older adolescents, who reserve judgment until more evidence is presented (Horn, et al., 1999; Killen, et al., in press). These children and adolescents may use this stereotypic information to make inaccurate attributions of intentions based on stereotypes. Gaining a better understanding of the ages that this occurs, and the contextual scenarios that encourage this misattribution of intentions will inevitably broaden our understanding of intergroup biases of children, as well as how to help ameliorate theses biases.

Gender is one of the most salient intergroup categories in society today, especially for children, and thus would be a robust intergroup category to use in our examination of the impact of intergroup attitudes on interpretations of intentionality. As has been demonstrated by extensive research on gender development, gender identity and gender labeling is pervasive (Liben & Bigler, 2002). Gender stereotypes, or knowledge representations or beliefs about sex-related behavior and characteristics (Ashmore & Del Boca, 1979), serve to organize thought and guide action (Martin & Halverson, 1981; Weisner & Wilson-Mitchell, 1990). Yet, in the context of situations involving fairness, children and adolescents often view gender stereotypes as wrong because they lead to unfair and unequal treatment of others (Killen & Stangor, 2001), such as when boys are excluded from girl type activities like ballet or when girls are excluded from boy-type activities like football. These are often evaluated as unfair and discriminatory (Killen & Stangor, 2001).

While children and adolescents view gender exclusion in peer contexts as wrong, children as young as 18 months have acquired an understanding of the social differentiation of the genders, such that they associate hearts, softness, and the color pink with females, and bears, roughness, and the color blue with males (Eichstedt, Serbin, Poulin-Dubois, & Sen, 2002). Children as young as 3 have been found to have gender stereotypes and have been also been found to rate masculine items as more desirable than feminine items (Eichstedt, et al., 2002; Killen & Stangor, 2001). Furthermore, children as young as 4 can hold stereotypes and base initial judgments of ambiguous social situations on stereotypic knowledge (Killen, Pisacane, Lee-Kim, & Ardila-Rey, 2001). It serves to reason then, that intergroup stereotypes, such as

gender stereotypes will impact how children discern and reason about intentionality in every day social interactions.

Studying this convergence of morality, theory of mind, and intergroup relations, is an important line of research as it impacts a child's moral reasoning, their understanding of intentionality, and the impact of intergroup stereotypes on the process. It is crucial to have an understanding of how children are making attributions of intentions, and what contextual information is being used to make these attributions, which is the goal of the current study.

Current Study Design and Hypotheses

There are three theories being further explored in the current study, social domain theory, gender stereotype theory, and false belief theory of mind. As illustrated earlier, one of the three components of social cognitive domain theory is moral development, and a central feature of many of the everyday moral questions and dilemmas individuals encounter on regular basis concerns intentionality. When examining the morality of everyday decisions such as whether to attribute blame to someone who has potentially made a transgression, attributions of intentionality are made. Additionally, as explicated earlier, attributions of intentionality are central to the study of false belief theory of mind, as the study of theory of mind explores children's understanding of other's belief states (Wellman & Liu, 2004). More specifically, the false belief task assesses children's ability to utilize their knowledge of another person's belief states to predict his or her subsequent actions when those beliefs differ from reality and from the child's own knowledge (Wimmer & Perner, 1983). This is very complex cognitive task which requires the child to be able to

predict another person's belief state, a skill that is essential to predicting another person's intentions.

Studying children's capacity to pass the false belief theory of mind task, and understanding how and when that capacity impacts their ability to make attributions of intentionality in morally relevant scenarios will facilitate our understanding of moral development, as it will begin to elucidate when a child can tell if someone has done something intentionally or not. Understanding when a child can make that determination will help us to understand the ontogeny of that very central aspect of moral development, as well as what cognitive capacities are necessary for it to occur. It is important to keep in mind though, that intentions cannot be known by the action alone, unless explicitly stated by the actor. Intentions must be inferred from contextual information such as the time, place, objects, other involved individuals, information about the actor such as beliefs, desires, stereotypes, past behaviors, in addition to the information about the action itself (Malle, Moses, & Baldwin, 2001). The contextual information therefore can play a critical role in the determination of intentionality, and has to be considered when attempting to determine the ontogeny of this ability. The child has to be able to not only determine intentionality, but has to be able to take in the contextual information that informs that determination. Varying the contextual information, such as the stereotype consistency of the object of interest in a morally relevant scenario, will help us to understand what aspects of the context are being attended to and are impacting the determination of intentionality. All of these things together will facilitate our ability to determine the ontogeny of the skill of determining intentionality in a morally relevant scenario.

In addition to the current study furthering the field of moral development by providing a means of discerning the ontogeny of attributing intentions, the field of false belief theory of mind will be furthered as well. The current study will embed a false belief theory of mind task into a morally relevant scenario, thereby giving it context and social relevance, as typical everyday scenarios in which a child would be making determinations of other people's belief states would occur. This will allow for a more thorough understanding of the age at which a child can pass the false belief theory of mind task when it is in a more real to life scenario, thereby pushing the field of false belief theory of mind forward.

To date, no research has examined the impact of intergroup attitudes on moral judgments in the context of situations involving false belief theory of mind competencies. It is proposed, however, that this line of research will push the field of moral development forward and reveal new information about the ontogeny of moral reasoning and the emergence of intergroup attitudes. It is also possible that this research will find children's attributing of intentionality in ambiguous situations to be impacted by the stereotypes they hold, thereby laying the foundation for children to assign blame to another because they are consistent with a stereotype, or assign innocence because they are not consistent with a stereotype, rather than withholding judgment due to the ambiguity of the scenario.

As seen through the previously discussed research of Killen and colleagues, the social-cognitive domain model provides a useful heuristic for investigating these issues. It allows for an in depth critical analysis of the evaluations, judgments, and justifications that children and adolescents make concerning everyday situations,

which will allow for a more thorough understanding of children's attribution of intentionality and thus their moral reasoning. This proposal will use the social-cognitive domain model to examine how intergroup attitudes, such as gender stereotypes, interact with a child's false belief theory of mind to impact children's social evaluations, judgments, and reasoning about intentionality so as to further our understanding, as well as the field of moral development.

Drawing on the findings from Killen (Killen et al., in press) the present study was designed to measure when and how children integrate three types of knowledge in an everyday peer exchange at school: 1) moral reasoning (is an act all right or not all right?); 2) intentionality of others (does the act require knowledge of another's mental state?) and 3) gender stereotypes (do stereotypes contribute to judgments about intentionality or morality?). In order to test these different types of judgments, children were presented with a task that was modified from previous research (Killen et al., in press), in which they were asked to evaluate an accidental transgressor that "destroyed" a desired object. In the present study, children evaluated an accidental transgressor that "misplaced" a desired object that was either consistent or inconsistent with gender stereotypic expectations. In this way, intergroup attitudes will be introduced into the "accidental transgressor" paradigm. The goal is to investigate whether knowledge and use of stereotypic expectations regarding ownership of toys (dolls, trucks) is related to judgments of intentionality and moral judgment. For example, do children attribute negative intentions in ambiguous situations when the act fits a gender stereotypic expectation?

Study Design

The emergence of both gender stereotypes and theory of mind arise between the ages of 3 and 8 years of age. It has been found that children develop a theory of mind by the age of 4, and continue to make refinements to the ability throughout early childhood (Wellman & Lui, 2004). Additionally, children have been found to recognize, use, and make decisions using gender stereotypes as early as 3 (Killen et al., 2001). For these reasons, participants will be ages 3-4, 5-6, and 7-8 years of age.

Participants will first be interviewed with two of four possible morality and theory of mind scenarios in an intergroup context, the Intergroup Attitudes Attributions of Intentions Task (see Table 1 for task design). Each participant will be administered a gender consistent scenario, in which both the gender of the agent of the scenario and the gender of the potential victim are consistent with the gender stereotype of the object of interest that is "misplaced" (e.g., two girls and a doll, or two boys and a truck). The gender of the participant will be matched to the gender of the actors in the gender consistent scenarios. Each participant will also be administered a gender inconsistent scenario, in which the gender of the agent of the scenario is different from the gender of the potential victim, and the object of interest that is "misplaced" is inconsistent in gender stereotype with the agent of the scenario. Half of the participants will be administered a gender inconsistent scenario in which the agent of the scenario is male and the object of interest is of low status (doll). The other half of the participants will be administered a gender inconsistent scenario in which the agent of the scenario is female and the object of interest is of high status (truck). This design will allow for an examination of the distinct impact of

ingroup/outgroup identification, participants' identification with the agent, as well as identification with the object (toy).

The scenario used to measure children's judgments for a "gender stereotypic inconsistent act" is the following:

> "This is Tina and this is Mark (show pictures). They are children in this classroom who like playing with dolls and trucks and balls. Tina and Mark have backpacks that look the same, except Tina's has a T on the front and Mark's has an M on the front. This is Tina's (show T backpack) and this is Mark's (show M backpack). They are playing with their toys at school when the teacher asks them to get their backpacks ready to go home before they go outside for recess. Tina picks up her doll (show doll) and puts it next to the pile of backpacks near the door (place next to backpacks), where her backpack and Mark's backpack are sitting. Then, she goes outside for recess (show Tina leaving). Mark stays inside to help the teacher clean up the classroom. Mark sees Tina's doll and puts it into this backpack (show M backpack, and slide doll behind it). Can you show me which backpack the doll is in now? (If incorrect, retell/clarify)."

In this scenario, one child, Mark, is the classroom helper and potential transgressor, and he puts another child's toy into his own backpack. This act does not represent a stereotype as boys do not stereotypically play with dolls, and in fact is considered stereotype inconsistent. In the other stereotype inconsistent scenario, the classroom helper and potential transgressor is a female, and the toy that she puts into her own backpack is a truck. This is stereotype inconsistent, as girls do not stereotypically play with trucks. In the stereotype consistent scenarios, the classroom helper and potential transgressor puts another child of the same gender's stereotypically consistent toy into their own backpack (a girl putting a doll into her backpack, and a boy putting a truck into his backpack). These scenarios are considered gender consistent as the gender of the potential transgressor matches the gender stereotype of the toy taken.

The participants will be asked to make judgments and justifications concerning the intentions of the agent in the scenario, the acceptability of punishing the agent, potential for friendship between the agent and the victim, as well as the emotional state of the victim. The task will also include embedded theory of mind assessments of false belief, location change, and second order false belief.

In addition to evaluating morality and theory of mind in an intergroup context, participants will complete a distinct False Belief Theory of Mind Competency Task with a standard false belief, false contents task as well as a standard false belief location change task. The participants will also be administered a Gender Stereotype Knowledge, Tolerance, and Flexibility Task including gender stereotypic, counter-stereotypic, and neutral toys. These tasks will be assessed as participant variables in order to relate them to the attribution of intentions judgments and justifications.

Hypotheses

The overall goal of this study is to assess children's moral judgments, including attributions of intentionality, with respect to the child's false belief theory of mind competence and their gender stereotype knowledge and flexibility in an intergroup task created for this study.

Attributions of Intentionality and Gender Stereotype Knowledge, Tolerance, and Flexibility. In accord with previous research by Killen and colleagues, Liben and Bigler, as well as Freeman (Freeman, 2007; Killen, et al., 2001; Liben & Bigler, 2002) indicating that children as young as 3 hold gender stereotypes, and that children use those stereotypes when making decisions about toy preference and social decisions, it is expected that children with a high knowledge, low tolerance, and/or

low flexibility of gender stereotypes will indicate more negative intentions and more punishment acceptability in the gender consistent scenario than in the gender inconsistent scenario, as children with a high knowledge, low tolerance, and/or low flexibility of gender stereotypes will have a harder time believing that a child would desire to have a counter-stereotypic toy, indicating that gender stereotypes are impacting the attribution of intentionality.

Attributions of Intentionality and False Belief Theory of Mind. It is expected that, in concordance with Killen and colleagues previous study on Morality and Theory of Mind children's false belief theory of mind competence will impact the attributions of intentions and judgments of punishment acceptability such that children with a false belief theory of mind competence will indicate less negative intentions and less punishment acceptability than children without false belief theory of mind (Killen, Mulvey, et al., 2009). Furthermore, it is expected that children with a false belief theory of mind competence will use more moral justifications indicating issues of lack of negative intentions due to the accidental nature as well as prosocial issues referring to helping the teacher and being cooperative.

False Belief Theory of Mind. It is expected that overall, fewer children will have a false belief theory of mind competence within the Intergroup Attitudes Attribution of Intentions Task, than with the independent measurement of false belief theory of mind competence (Killen, Mulvey, et al., in press). Furthermore it is expected that the embedded theory of mind competence will be affected by the age of the participant (Killen, Mulvey, et al., in press), as well as by the indication of

knowledge, tolerance, or flexibility of stereotypes for the gender inconsistent scenario.

Age of Participant. It is expected that children will attribute less negative intentions and indicate less punishment acceptability with age as they will better be able to see the possible accidental nature of the transgression (Killen, Kelly, & Richardson, in press; Killen, Mulvey, et al., in press). Killen and colleagues have found in previous studies that with age, children and adolescence are more able to see the complexity and ambiguity of situations, and therefore less likely to indicated negative intentions (Killen, Kelly, & Richardson, in press; Killen, Mulvey, et al., in press).

Gender of Participant. In line with what Killen and colleagues have found in previous studies, it is expected that female participants will be less likely to indicate negative intentions and indicate less punishment acceptability than will the male participants, as the females, due to female's greater social experience of being excluded from participating in activities and from using toys due to their gender than males (Killen, Henning, Kelly, Crystal, & Ruck, 2007; Killen, Kelly, Richardson, Crystal, & Ruck, 2009; Killen, et al., 2001).

Expected Contribution to the Field

Children's reasoning about and judgments concerning moral decisions is crucially important to their social development. Children are making social decisions everyday that are influenced by their moral reasoning and by their ability to infer intentionality (Killen, Richardson, & Kelly, 2010). Researching the contextual information impacting a child's ability to reason about intentions, and understanding

what role gender stereotypes and false belief theory of mind competence are playing in a child's ability to reason about intentionality will further our ability to understand moral reasoning and the factors that are impacting that reasoning process. Additionally, gaining a better understanding of the ages that correspond with the effects of having gender stereotypes and a theory of mind competency, as well as the contextual scenarios that encourage possible misattribution of intentions will inevitably broaden our understanding of the intergroup biases of children, as well as how to help ameliorate theses biases.

Understanding the role of false belief theory of mind in moral decision making will deepen our understanding of the cognitive competencies, such as those necessary to have a indicate a false belief theory of mind, which are necessary to make moral decisions such as those concerning intentionality. Additionally, understanding whether having gender stereotypes impact those moral decisions being made will further our understanding of the influence of gender stereotypes, and will enable the creation of interventions that address the comprehensive effect of stereotypes. Furthermore, this research will address if stereotypic expectations are related to children's attributions of intentions and moral judgments. Children who have stereotypic expectations may unknowingly be using these expectations as a reason to assume that a child did not intentionally do something in an ambiguous situation when the action does not conform to a stereotype, and by contrast assign blame to someone when the ambiguous act does conform to stereotypes. For example, when Horn, Killen, and Stangor (1999) examined ambiguous scenarios with adolescents in a high school setting (scenario 1, someone broke the sound equipment

at a party; scenario 2, someone broke into the school's computer system), the participants used stereotypic expectations to assign blame. The participants were more likely to accuse the "jocks" than the "techies" of breaking the sound equipment at the party, and were more likely to accuse the "techies" than the "jocks" of breaking into the computer system. Even though both scenarios were ambiguous, as are the ones in the proposed study, the participants used stereotypic expectations to assign blame.

To date, no research has examined the impact of intergroup attitudes and theory of mind competencies on moral judgments. Despite the lack of research in this area, it is clear that further research will push the field of moral development forward and reveal the new information about the ontogeny of moral reasoning and the emergence of intergroup attitudes. In sum, this proposal seeks to further both the understanding and the interactions of moral reasoning, theory of mind, and gender stereotypes in general, and more specifically how having a theory of mind and holding gender stereotypes impacts children's ability to make decisions about intentionality, and what contextual information is impacting these decisions so as to advance our understanding of moral decision making.

Chapter 2: Literature Review

Introduction

In this literature review I will assess the current state of moral development research with regards to understanding the role of attribution of intentionality on decision making, as well as other contextual factors impacting decision making around intentionality. I will examine this literature through the lens of the Social Domain Model. I will begin with a review the fundamentals of the Social Domain Theory so as to lay a foundation for the theory guiding my research plan. I will then review theory of mind research concerning intentionality as well as the research at the intersection of moral development and theory of mind so as to understand the role theory of mind plays in decision making concerning intentionality. Subsequently, in order to identify the important research questions that need to be addressed research in the area of attribution of intentions, intergroup relations, and the link of moral reasoning and theory of mind the morality and intergroup relations will be thoroughly examined, with a specific focus on the impact of gender stereotypes, preferences, and ingroup/outgroup relationships on cognitive functioning, decision making, and moral reasoning. Finally I will lay out a future line of research which takes all of these literatures into consideration with the goal of furthering the understanding of the field, and filling the holes in current state of the research.

Social Cognitive Domain Model

As indicated above, the social cognitive domain model allows for researchers to explore attribution of intentionality and decision making about familiar events, and

then evaluate what children are reasoning about when making those decisions. More specifically, it uniquely permits researchers to investigate when children are taking stereotypes into account when making a decision about intentionality. This is critically important in gaining an understanding of the ontogeny, formation and the developmental trajectories of prejudice (the tendency to attribute negative characteristics to outgroup members), stereotyping (making judgments about an individual's traits or behaviors based on group membership), and discrimination (differential treatment based on biased beliefs about one's group membership) (Killen, Richardson, & Kelly, 2010).

As discussed earlier, social domain theory provides a basis for investigating different forms of reasoning about complex social issues, and research using this model has empirically demonstrated that individuals use three distinct domains of social reasoning; moral, social conventional, and psychological (personal), and that individuals utilize one or more of these domains when understanding and making decisions about their social worlds, including those decisions made in intergroup situations (Horn, 2003; Killen, Sinno, et al., 2007; Smetana, 2006). Understanding these three domains of social reasoning allows for a thorough examination of the reasoning behind everyday social situations. The moral domain pertains to issue's of others' welfare (harm), justice (comparative treatment and distribution), and rights (Nucci, 1978; Smetana, 2006; Turiel, 1983, 1998).

Individuals judge moral rules to be generalizable and unalterable, and consequently judge moral transgressions as wrong even in the absence of rules and wrong independent of authority dictates (Smetana, 1983). When children and

adolescents are reasoning about intergroup interactions and judge an action, such as the denying of access to school based solely on race, to be wrong, they use moral reasons to justify that judgment (Killen, Lee-Kim, et al., 2002).

Additionally, the social cognitive domain model addresses the social conventional domain, which pertains to issues involving rules, norms, and conventions that coordinate social interactions of individuals within social systems (Horn, 2003). Individuals judge social conventional rules to be both relative to the social context and alterable, and consequently judge social conventional transgressions to be contingent upon the presence of rules and subordinate to authority dictates (Smetana, 1983).Social conventions ensure smooth group functioning and promote group identity.

Finally, the social cognitive domain model includes a psychological or personal domain which pertains to knowledge of interpersonal relationships, the understanding of individuals as psychological systems, and issues in which individuals have personal jurisdiction such as choice of friends, choice of occupation, and privacy (Horn, 2003). The psychological or personal domain appeals to individual preferences or prerogatives, and is therefore not regulated by rules or judgments about transgressions (Killen, Lee-Kim, McGlothlin, & Stangor, 2002). This indicates that built into the model is an understanding of the importance of context. If a rule is wrong regardless of the timing, the people involved, the location, or any of the specifics of the situation, it transcends all context. In the social conventional domain, individuals' judge rules to be both relative to the social context and alterable, and consequently judge social conventional transgressions to be

contingent on the presence of rules and subordinate to authority dictates (Smetana, 1983). The psychological or personal domain appeals to individual preferences or prerogatives, and is therefore not regulated by rules or judgments about transgressions (Killen et al., 2002). The social conventional and psychological domains therefore allow for context to be a deciding factor when decisions and judgments are being made.

Additionally, the social cognitive domain model allows for the application of reasons from one domain (moral or social conventional) or more than one domain (both moral and personal) to a given situation, and acknowledges that the judgments made may include interpretations of specific features of the situation (Killen et al, 2002). When individuals evaluate complex acts and issues, it is assumed that they weigh different considerations and give priority to one perspective or form of reasoning over another, and can vary by physical and situational context. Individuals have to assess the multiple dimensions often present in a context in order to make an evaluation. With this understanding, it becomes obvious how critical it is for researchers to analyze experimental situations in terms of both the components of the context as well as the predictions of how individuals will analyze it.

Previous intergroup research utilizing the social cognitive domain theory has shown that when children and adolescents are reasoning about an intergroup interaction and judge an action, such as the denying of access to a music club or friendship based solely on race or gender, as fair, social conventional reasons of sustaining group identity and group functioning and/or reasons of personal preference for friends are generally used (Killen et al., 2002). It is possible to manipulate

everyday scenarios so that the use of a stereotype in one scenario over another can be seen to influence a decision concerning intentionality. It is in this exploration of the social reasoning used within the manipulated scenarios that the importance of a stereotype in directing an answer can be seen and explored. Additionally a critical examination of everyday situations that children potentially experience can allow us to examine which situations, and which contextual variables in certain situations, are elemental to the use of moral reasoning versus the use of social conventional or personal reasoning when making decisions about intergroup interactions. An extensive line of research has shown that individuals from as early as two years of age differentiate events along these domain distinctions (for reviews, see: Smetana, 1995; Tisak, 1995; Turiel, Killen, & Helwig, 1987).

Furthermore, having a thorough understanding of how and when the cognitive competence of false belief theory of mind is coming into account and impacting a child's ability to make attributions of intentionality in morally relevant scenarios will allow us to further understand the developmental progression of moral reasoning, as well as to understand what cognitive capacities are compulsory to having fully represented social and moral decision making concerning intentionality.

In the current literature, what has not been examined in a thorough or critical manner are the contextual factors that are interacting with a child's ability to attribute intentionality, such as if highly a stereotypical object is introduced, would that affect how a child reason's about the intentions of someone taking that object if it did not belong to them. Are children who hold stereotypes less able to make an accurate attribution of intentionality when a highly stereotyped object is being used?

Additionally, the question of the role of theory of mind competence in the developmental progression of moral reasoning and more specifically the attribution of intentionality in morally relevant scenarios has not been thoroughly investigated. Are children with a theory of mind competence able to attribute intentionality even if they hold gender stereotypes and a stereotypical object is the object of interest? Examining these questions is necessary for a more comprehensive understanding of moral development and moral reasoning.

Moral Development Research

The idea that children construe their own understanding of reality led to the conclusion that people behave and respond in accord with their own interpretations of their experiences, rather than the experiences themselves (Inhelder & Piaget, 1958; Piaget, 1952; Wainryb, 2004). This has spurred research in the role of subjective interpretations in people's social behavior, social interactions, and social adjustment. In the early 1990s, Wainryb and colleagues began to systematically document the different aspects of the interpretive process that go into making moral judgments (Wainryb, 2000; 2004; Wainryb & Brehl, 2006).

Warneken and Tomasello (2009) hold that one such aspect of making moral judgments is altruism, and have found that children as young as 14- to 18-months-of-age will help others, irrespective of any reward. Specifically, they found that a child of this age can observe an actor who is unable to achieve a goal and act altruistically to help the actor achieve that goal, even when there is no immediate benefit for the child (Warneken & Tomasello, 2009). Additionally, Tomasello and colleagues have found that infants from 12- to 18-months-of-age can understand other person's

behaviors in terms of their underlying goals and intentions (Tomasello, Carpenter, Call, Behne, & Moll, 2005) and furthermore, can distinguish purposeful from accidental actions (by imitating twice as many intentional actions by adults, indicated by the adult exclaiming, "There!", than accidental ones, indicated by the adult exclaiming, "Whoops!", indicative of the children differentiating the two) (Carpenter, et al., 1998). These findings are complimented by the work of Nichols, Svetlova, and Brownell who found 18- to 30-month-olds able to help instrumentally in an action-based task (2010). Taken together, these findings provide evidence that children are reasoning about moral decisions and actions from a very early age.

In addition to altruistically helping others from a very early age, researchers have found that by 2-years-of-age, children will respond with empathic concern for others' distress, (Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992) help adults who have dropped or misplaced things and comfort those in distress, (Liszkowski, Carpenter, Striano, & Tomasello, 2006; Warneken & Tomasello, 2006) instrumentally cooperate with adults and peers, (Brownell, Ramani, & Zerwas, 2006; Warneken, Chen, & Tomasello, 2006) as well as voluntarily share valued and desired resources, such as food, when there is no cost to them (Brownell, Svetlova, & Nichols, 2009). All of these behaviors can be interpreted as precursors to moral reasoning and decision making, and seem to occur to the benefit of the physical well-being of others, but are happening before theory of mind competencies and thorough understandings of other people's mental states come in-line. Moral decisions concerning other people's intentions build upon these skills, but there is a lack of understanding of why these infants, toddlers, and children are making the decisions

that they are as it is very difficult for these children to verbalize their justifications and reasoning. Information concerning the reasoning process behind these decisions, such as information discerned through the social domain research's focus on justifications and reasoning would add to the understanding of what aspects of the context are taken into account when making the decision.

Social Cognitive Domain Theory and Moral Development

In the social cognitive domain literature, morality is seen as one of the elements of children's developing social knowledge concerned with justice, welfare, and rights (Smetana, 2006; Turiel, 1983; 1998). These moral issues coexist with concerns for authority, tradition, social conventional issues of norms, privacy, and personal preferences (Smetana, 2006). Children form concepts of fairness and quality as young as 2 ½ - years-of-age (Smetana, 1985). The social cognitive domain model holds that these systems of social knowledge arise from children's experiences in the social environment (Turiel, 1983; 1998). Thus, differing social experiences can differentially impact the social decisions that children are making, and their interpretation of intentionality. If a child is surrounded by friends and family that adhere to strict gender-stereotypical roles, the child can bring that information into mind when making decisions about intentionality, and possibly perceive a boy taking a doll as accidental, as the boys that he encounters do not ever play with dolls, and therefore boys do not like dolls, and would not take one on purpose.

When children are making decisions about the seriousness of a transgression, moral transgressions of all kinds are seen as more serious, more deserving of punishment, more independent of rules, and more generalizably wrong than social

conventional transgressions (Smetana, 2006). Therefore, if a child believes that someone stole something on purpose, issues of fairness and justice would arise, and this act would be judged to be more deserving of punishment than if the action was accidental. Additionally, children have been found to apply concepts of welfare to situations entailing physical harm at an earlier age than situations involving psychological harm, most likely because physical harm is concrete and observable, while psychological harm is abstract and a child must understand someone else's mental state (have a theory of mind) to fully understand psychological harm (Smetana, Schlagman, & Adams, 1993). These findings illustrate the importance of theory of mind competence in a child's ability to understand the full complexity of a social scenario, such as one involving someone else's desires and expectations, as well as the importance of theory of mind competence in a child's ability to attribute intentionality.

Theory of Mind and Social Decision Making

Even though the link between moral development and theory of mind research has only been recently explored in the research (see for exceptions see: Helwig, Zelazo, & Wilson, 2001; Killen, Mulvey, Richardson, Jampol, & Woodward, in press; Leslie, Knobe, & Cohen, 2006; Wainryb & Brehl, 2006; Zelazo, Helwig, & Lau, 1996), understanding how a child's cognitive capacities impact their moral judgments and justifications is crucial to understanding the developmental progression of moral reasoning (Wainryb & Brehl, 2006), and their social decision making in general. Understanding the state of the research of theory of mind and

social decision making in general will facilitate predictions about theory of mind and moral decision making.

Wellman and Liu (2004) define Theory of Mind as children's understanding of other people's mental states (pp. 523). Theory of Mind has been rigorously researched for decades with the majority of the research examining young children's understandings of intentions, emotions, desires, beliefs, false beliefs, and knowledge, but focusing on one single cognitive process at a time (Wellman & Liu, 2004). One of the most recurrently used tasks in the literature, the false belief task, assesses children's ability to use a person's belief states to predict his or subsequent actions when those beliefs differ from reality and from the child's own knowledge (Wimmer & Perner, 1983).

Theory of Mind and Beliefs

The majority of the research into theory of mind competencies holds that prior to the age of 4- to 5-years-of-age, children do not fully understand that beliefs are representations of reality or that different people may have or construe differing representations of, or beliefs about, the same reality (for a review, see Wellman, Cross, & Watson, 2001). It appears that young children are relying on the assumption that perception is the sole basis for belief, with a one-to-one mapping between what one sees or hears and what one knows (Pillow & Mash, 1999). Additionally, 3-year-olds find it difficult to comprehend that someone might not know something that they know to be true (Leslie, 2000), and also find it difficult to report even their own previous mistaken beliefs (Gopnik & Astington, 1988; Harris & Leevers, 2000).

Beginning around 4- to 5-years-of-age, children first begin to develop an understanding that people may have access to different information and may end up with different beliefs. By the age of 7- to 8-years-of-age children begin to comprehend that individuals perceive experiences and the mind then selects, transforms, and organizes that perceptual information (Wellman, et al., 2001). Consequently, 7- to 8-year-olds are able to appreciate, for the first time, that people may have equal access to all the relevant information but still form differing beliefs. Furthermore, it is at the same time that children begin to become aware that prior thoughts or emotions can inform current beliefs and interpretations of experience (Wellman, et al., 2001). Children's burgeoning understanding of others interpretations and construal's of observed experiences also informs their understandings of intentionality.

Theory of Mind and Intentionality

Children begin their journey of understanding intentionality at 12- to 18-months-of-age when it has been shown that children can distinguish purposeful actions from accidental ones (Carpenter, et al., 1998). As previously mentioned, Tomasello and colleagues determined that 12- to 18-month-old children could distinguish intentional actions from accidental ones by imitating twice as many intentional actions by adults, as indicated by an adult exclaiming, "There!", than accidental ones, indicated by an adult exclaiming, "Whoops!", (Carpenter, et al., 1998). Six months later, by the age of two, Wellman and Woolley found that children could conceive of other children and adults as active agents with actions that were directed toward a goal, and could even predict behavior on the basis of the other

person's desire (1990).

In a great deal of research, information concerning intentions is often confounded with information regarding consequences (Wainryb & Brehl, 2006). When this information is not confounded, but rather given explicitly, 5- to 6-year-olds judge intentional acts as more wrong than accidental ones (Shultz, Wright, & Schleifer, 1986; Wellman & Liu, 2004). In fact, when not asked to weigh intentions against consequences, 3-year-olds can distinguish between deliberate and accidental breaches (Harris & Nunez, 1996; Núñez & Harris, 1998; Siegal & Peterson, 1998). Young children's early understandings of intentions and motives are not as complete or concrete as the understandings of older children (Jones & Nelson-Le Gall, 1995; Karniol, 1987; Nelson-le Gall, 1985). Children between 3- and 4-years-of-age still inaccurately report their own intentions to match the actual outcomes of their own actions (Phillips, Baron-Cohen, & Rutter, 1998; Schult, 2002). Furthermore, in studies that do not clearly differentiate the cause of an action, children equate intentions with outcome and confuse fulfilled desire with unfulfilled intentions (Schult, 2002). In fact, young children are likely to assume that if something happens it is because someone intended it to, even when provided with information to the contrary (Kalish, 2006).

Theory of Mind – Developmental Progression

Wellman and Liu (2004), in an attempt to consolidate the wealth of theory of mind knowledge into an age related progression from one theory of mind understanding to another, did a meta-analysis of research comparing different types of mental state understandings such as desires versus beliefs and ignorance versus false

belief. Wellman and Liu (2004) were able to provide evidence for a consistent developmental progression from one theory of mind task to another, whereby if they passed a later task, they would be able to pass the earlier tasks as well. Wellman and Liu (2004) found that children were able to understand other children's desires before other children's beliefs, understanding that two people can have different desires for the same object before they become aware that two people can have different beliefs about the same object. This finding provides evidence that a child could understand that a girl could want to play with a doll, or that a girl could want to play with a truck before they pass the standard false belief theory of mind task. This would mean that a child who passes the false belief theory of mind task should be able to make that distinction. It is possible though, the addition of an object that has emotional, stereotypical, and social significance could alter that progression, making it harder for a child to see that another child has desires different from their own.

Theory of Mind and Social Reasoning

The search to understand the role of theory of mind in social reasoning has lead researchers in many directions. Karpinski and Scullin (2009) researched theory of mind with relation to executive functioning, higher-level action control, and suggestibility with preschoolers. Karpinski and Scullin (2009) interviewed 80 preschoolers (3- to 5-years-of-age), and found that, when controlling for age, children with better executive functioning were overall less suggestible when being interviewed, and older children with a more developed theory of mind were less suggestible. More specifically, they found that executive functioning is more relevant when interviewers try to apply social pressure to get children to change their answers,

and theory of mind is more relevant in a child's ability to resist suggestibility in a pressured interview (Karpinski & Scullin, 2009). Taken together, this implies that the social understanding and decision making can be impacted separately by their executive functioning as well as their theory of mind. It is possible then that having a person trying to sway a decision or a social understanding could be cognitively similar to having a previously held belief, understanding, or stereotype affecting a decision or social understanding, and the impact of theory of mind and possibly the impact of executive functioning on that should also be examined.

In order to directly examine the effect of theory of mind and stereotyped belief about desirability on decision making, Terwogt and Rieffe (2003) interviewed 29 4-year-olds and 29 5-year-olds about the desirability of four different toys, a doll, a toy tea set, a toy plane, and a toy car, which they determined in a pilot study to be stereotypically female and male. The children were then read a story in which the protagonist expressed his or her preference for one toy over another (Terwogt & Rieffe, 2003).

Terwogt and Rieffe (2003) found that both age groups were able to use the desires of the protagonist as a basis for their answer, and also were able to predict that the protagonist would be happy if the outcome was consistent with the expressed desire and unhappy if the outcome was inconsistent with their expressed desires. Additionally, the researchers found that the number of correct answers was significantly lower in inconsistent scenario (a female desired a stereotypically male toy, or a male desired a stereotypically female toy) (Terwogt & Rieffe, 2003). Furthermore, they established that these findings were consistent regardless of the

child's personal toy preference and did not differ by gender, indicating that the children were not biased by their own desires, but by their beliefs or possibly their stereotypes about what is desirable for boys and for girls (Terwogt & Rieffe, 2003). When the researchers looked exclusively at which toy the child desired, the male participants chose stereotype consistent toys 83% of the time, while the female participants chose stereotype consistent toys only 45% of the time (Terwogt & Rieffe, 2003).

This study is informative in many ways. First, it provides evidence that overall, preschool age children are able to make decisions about a person's preferences based on expressed information about their desires. Second, it shows that those predictions about preference are negatively affected by counter-stereotypic information about the desire for toys, but not children's' own toy desires or preferences. This could be because children have a harder time recalling information that is counter-stereotypic (a female doctor) than stereotype consistent (a female teacher) (Liben & Bigler, 2002; Ruble & Martin, 1998). These findings also indicate that a child's decision about another child's intentions to take a toy that is counter-stereotypic can be affected by stereotypic beliefs about boys and girls preferences for toys. It is unclear, though, whether having a theory of mind competence would improve the ability of a child to resist the held stereotypes about male and female toy preferences, as the study by Karpinski and Scullin (2009) might suggest.

Furthermore, Terwogt and Rieffe's finding that the male children were more likely to choose stereotype-consistent (83% of the time) toys than were the female children (45%) of the time, indicates that the male toys may be overall more desirable

to all children than the female-stereotyped toys (2003). This could also impact a child's decision about intentionality, as a child might be more willing to believe that a female desired to have, and therefore took, a male-stereotyped toy than a male desiring to have and therefore taking a female-stereotyped toy. It is unclear though, how this could impact a child's decision about the appropriateness of taking the toy, or how theory of mind may impact these decisions.

Theory of mind researchers have not only explored stereotypes and desirability, but have also separately made headway into issues of fairness, a cornerstone of moral reasoning research. Takagishi and colleagues played a modified version of the Ultimatum Game with preschoolers (with a mean age of 7 years) and also administered a false belief location change theory of mind task (Takagishi, Kameshima, Schug, Koizumi, & Yamagishi, 2010). The Ultimatum Game was set up so that one child was given candies by the experimenter, and then that child was able to divide the candies between him- or her-self and a second child. The second child could either accept or reject the offer of candies made by the first child. If the second child accepted the offer, then both children can kept the candies, but if the second child rejected the offer, no one got any candy. A purely economic model would expect that the second child should accept anything over zero, and the first child should offer as little as possible, but researchers have found that these results differ across cultures (Henrich et al., 2005), with the modal offer made of 50/50 and a mean offer of 60/40 (Camerer, 2003) cross-culturally.

The researchers in this study found that children who had acquired a false belief theory of mind competence proposed a higher mean offer in the game than did

children who had not acquired a false belief theory of mind (Takagishi et al., 2010). The researchers take this to mean that the ability to infer the mental states of others is important to fairness-related behavior (Takagishi et al., 2010). This research provides evidence that having a theory of mind competence effects social decision making in children, and decisions that children come across regularly in their daily lives about sharing and the distribution of resources, all part of moral decision making. What is not clear is how this competence would affect other typical moral decisions made by children in their everyday lives such as the attribution of intentions.

Morality and Theory of Mind Research

There has been only a limited amount of research connecting morality and theory of mind (for exceptions see: Helwig, Zelazo, & Wilson, 2001; Killen, Mulvey, Richardson, Jampol, & Woodward, in press; Leslie, 2006; Leslie, Knobe, & Cohen, 2006; Wainryb & Brehl, 2006; Zelazo, Helwig, & Lau, 1996). Within the research that has connected the two areas of research, there are several limitations. First, many of the studies do not use the same or even similar measures of moral judgment, which makes it difficult to examine the effects across studies. Furthermore, very few studies examine morality and theory of mind within the same task, which limits the ability to determine how theory of mind competence bears on moral judgments, and instead allows for only correlational relationships between the two concepts. Additionally, most theory of mind methodologies provide very limited social or contextual information. For example, in the prototypic theory of mind false belief location change task, (i.e., Crackers have been moved from one container into another, one child being administered the task knows they have been moved, but others do not.

Will a child who did not see the crackers get moved look in the previous location or the correct one?). In this prototypic example, no social information is provided regarding ownership of the objects, the intentions of the person who moved the object, the relationship between the child who was present for the location change and the child who was not (e.g., friends, strangers). It is clear from previous research, such as that conducted by McGlothlin and Killen (2006) and Killen, Kelly, Richardson, and Jampol (2010), that showing contextual information such as the color of a child's skin, the intergroup nature of an interaction, or the social group a child is affiliated with (Horn, et al., 1999) influences decisions that children and adolescents make, illustrating that the context of a scenario, such as the gender or ethnicity of the actors in it, is very important for a child or adolescent when evaluating the facts, and children can weigh information differently based on the context (for a review see Killen, Richardson, & Kelly, 2010).

What we can discern then is that there are several holes in the research that explores the intersection of morality and theory of mind. A few researchers have conducted studies that address some of the previously mentioned concerns. Helwig and colleagues utilized a prototypical moral scenario, but made the intentions of the actors in those scenarios clear, which would not be typical in an everyday scenario.

Moral Judgments in Normal and Noncanonical Scenarios

Helwig and colleagues have explored children's judgments of both psychological and physical harm in both normal and noncanonical situations, and have found that children as young as 3 were able to judge the infliction of physical or psychological harm on unwilling participants as wrong, even when the reactions were

noncanonical (Helwig, et al., 2001; Zelazo, et al., 1996). What was intriguing about these studies on 3-, 4-, and 5-year-olds is that the younger children tended to take outcome only into account when making a punishment decision (if outcome is negative then punish), while the older children were more likely to use an intention rule or a conjunction rule (if outcome is negative and intention is negative, then punish). This indicates that the younger children have a harder time integrating the intention information with the morally relevant information, which was supported by the findings of Killen, Mulvey, and colleagues (in press) when they examined false belief theory of mind tasks embedded in morally relevant scenarios.

While these studies go far in elucidating the process of integrating intention information and outcome information when making moral judgments, they explicitly identify the reaction to the action as well as the intention of the actor. The identification of the intention of the actor is therefore not in question, it is whether this information was used in order to make the moral judgment or not. If the intention of the actor was left ambiguous, the participant would have to use the contextual information to discern intentionality, and then use this information to make a moral judgment, which is much more akin to what children do on a daily basis. Most children do not make their intentions clear before taking action, and therefore an important social skill developed at this age is correctly discerning the positive or negative connotation of intention in an action. Furthermore, it is impossible to note the impact the relevant contextual information such as the time, place, objects, other involved individuals, information about the actor, past behaviors, or information about the action itself, which is used to help discern intentionality in everyday

scenarios when the intentions are not explicitly made known, on the attribution of intentions or the moral judgments. As stated earlier, misattributing intentions in social situations can lead a child to have a harder time making friends, and can lead to exclusion and rejecting, (Crick & Dodge, 1996), so fully understanding how children make this connection between action and intentionality in a morally relevant scenario, and the information that they use to make this connection, is of critical importance to understanding a child's moral reasoning and social interactions.

The Importance of Factual Information and Beliefs when Making Moral Judgments

Wainryb and colleagues have also made important contributions to the literature connecting moral reasoning and theory of mind (Wainryb & Brehl, 2006). In a series of studies, Wainryb and colleagues manipulated information concerning the beliefs (upon which the characters presumably based their behavior) in hypothetical scenarios, and evaluated children's judgments of act acceptability (Wainryb & Brehl, 2006). One such example was a scenario in which a teacher who gave more snacks to the girls did so because she either believed that girls need more food than boys or believed that it is all right to be nicer to girls and not as nice to boys (Wainryb & Brehl, 2006). In this series of studies, Wainryb and colleagues found the 3-year-olds to be unable to understand that the characters in the scenarios had different beliefs than their own, and therefore evaluated the characters' behaviors in terms of what they themselves thought to be the right thing to do (Wainryb & Ford, 1998). The 5- and 7- year olds were able to attribute beliefs that were different than their own, but still nearly half judged the characters' behaviors in accord with their

own factual beliefs, and the majority judged the behaviors based on non-normative moral beliefs to be wrong (Wainryb & Ford, 1998). It wasn't until the children were 8- or 9-years old that they would accept the behavior based on factual beliefs different than their own, but still were much less accepting of the same behavior if it was based on moral beliefs other than their own (Wainryb & Ford, 1998). These studies add strongly to the field, in that they provide valuable details concerning what information children are using when making judgments about act acceptability in a morally relevant scenario, and also provide evidence that children are attending to many details in a scenario when making their judgments about act acceptability. The intentions of the actors in these scenarios are made transparent though, so what is still unclear are the interpretations that would be made and the conclusions that would be drawn if the intentions were not explicit, and what contextual information the characters would use to make those decisions.

Moral Judgments when Intentions are Ambiguous

In 2005, Wainryb and Brehl partially addressed these concerns as they examined the behavioral and moral interpretations that 4-, 7-, and 10-year-olds made when presented with hypothetical scenarios in which one child hurt the feelings of a peer by excluding him/her from a group, making unequal distribution of desired goods, or saying something mean (Brehl & Wainryb, 2005). They found that the younger children were less likely to refer to intentions or beliefs when explaining the behavior than the older children (Brehl & Wainryb, 2005). The authors suspected that this difference was due to the younger children's equating of intention with action, causing them to not discuss the two issues separately, while the older children

were able to understand that someone can intend for something to happen, but have something separate occur, leading them to discuss intentionality more often (Brehl & Wainryb, 2005). In this study, while the intentions of the action are not explicitly made obvious, the outcome is decidedly, and strongly, negative. The action here is not ambiguous, even though the intentions are, and trying to find a plausible reason for why the actor could not have had negative intentions would be a creative challenge for all the participants, and especially for the younger participants simply due to the verbal and imaginative difficulty of it. Additionally, while this study broaches the question of determining intentionality in a scenario that is similar to what one would encounter in their daily lives, it still leaves us with many unanswered questions, such as what the attributions would be if the interpretation of the action was ambiguous, and what contextual information the participant was attending to or using in order to make their attribution of intention.

In a similar vein, Leslie and colleagues conducted a series of studies examining the connection between explicitly foreseen side effects of an intentional act and the positive or negative valence of that foreseen side effect (Leslie, et al., 2006). In the scenarios the participants were told that a child who loved frogs was going to bring a frog over to a friend's house, and in one condition the friend loved frogs, while in the comparison condition, the friend hated frogs. In both conditions, the participant was told that the actor did not care that the friend will be alternately happy or upset. The authors investigated the participant's judgments of intentionality, and found that the participants judged the action to be intentional in the negatively valenced scenario, but not in the positively valenced one (Leslie et al.,

2006). The authors attributed this outcome to what they describe as the side-effect effect (Leslie et al., 2006), and consider the intentionality judgment to be driven by the valence of the scenario, which the authors refer to as the moral judgment of the scenario, instead of the moral judgment being driven by the judgment of intentionality (Leslie et al., 2006).

This line of research does provide significant information concerning what aspects of a scenario children are attending to and using in order to make their judgments of intentionality, but again, the initial intentions of the actor were explicitly stated, and were not ambiguous. Therefore, the question of what aspects of the contextual information would be utilized if the intentions were ambiguous still abound.

Taken together, all of these studies have added greatly to both the moral development literature. Cumulatively, they provide evidence that children are attending to many details in a scenario when making their judgments about act acceptability (Wainryb & Brehl, 2006), they address what aspects of a scenario children are attending to and using in order to make their judgments of intentionality and act acceptability in morally relevant scenarios (Leslie et al., 2006; Wainryb & Brehl, 2006) and provide evidence that younger children have a harder time integrating the intention information with the morally relevant information (Helwig, Zelazo, & Wilson, 2001). There is a great deal left to be understood about this connection of morality and theory of mind though. A study, such as the one conducted by Killen and colleagues (Killen, Mulvey, et al., in press) that used a standard moral judgment task and also utilized a moral dilemma that could occur in

one's daily life, and also examined morality and theory of mind in the same task and additionally added social and contextual information to the embedded theory of mind task addressed many of these concerns, but also opened the door to more questions at the same time, such as what other contextual information is being used to attribute intentionality.

Morality and Theory of Mind Study

As discussed earlier , in a study conducted by Killen and colleagues children from 4 – 8 – years-of-age were interviewed about a series of tasks, one prototypical moral reasoning, one prototypical false belief theory of mind, and one that embeds false belief theory of mind into a morally relevant hypothetical scenario. As stated previously, in the morally relevant theory of mind task, the children were read a vignette involving a child who is helping the teacher clean up the classroom while the rest of the class is outside at recess, and accidentally throws away another child's special cupcake. Participants were asked to make evaluations, judgments, and justifications concerning intentionality, as well as punishment acceptability, and attribution of emotions (Killen, Mulvey, et al., in press). The main findings from this study were that children without theory of mind were more likely to attribute negative intentions and found it more acceptable to punish children in the morally relevant hypothetical scenario than children with theory of mind, and that children attributed negative emotions and intentions to the accidental transgressor up until 8 – years-of-age (Killen, Mulvey, et al., in press). Again, this study was the first to make it clear that integrating moral judgment and theory of mind poses challenges to children.

The study by Killen and colleagues allows us to see some of what is impacting a child's ability to make an attribution of intentionality. It is clear that while a 4- or 5-year-old can correctly distinguish intentionality in a straightforward scenario involving no emotional valence or moral concerns, the addition of a moral premise adds a layer of complexity to the scenario that makes it more difficult to correctly decipher intentionality. What is not known is what other aspects of a scenario may also contribute to increasing difficulty or greater ease in the ability to discern intentionality. Killen and colleagues found that the fundamental determinations of intentionality, a core aspect of moral and social reasoning, is affected by the introduction of a moral premise to a straightforward theory of mind scenario, which falls into line with previous social domain research indicating that children and adolescents use more social conventional reasoning when presented with a multifaceted scenario (Killen, Mulvey, et al., in press). These new data add a wealth of information concerning what children are taking into account when making attributions of intentionality, and additionally informs us that these determinations of intentionality are context dependent.

This study also established that attributions of intentionality in a morally relevant scenario are affected by a child's theory of mind competence, and also that taking the false belief theory of mind task and embedding it into a moral context impact the age at which a child can demonstrate theory of mind competence (Killen, Mulvey, et al., in press). It is still unclear what contextual aspects were impacting the attribution of intentionality, and if adding contextual information that children encounter daily, such as intergroup interactions and stereotyped objects would impact

the attribution of intentions further or impact the age trends of false belief theory of mind competence. Adding contextual features, such as a highly gender stereotyped toy instead if an object which elicits no emotional reaction, to a scenario aimed at examining a child's ability to attribute intentionality will allow for a more thorough investigation of the impact of those contextual features. This would enable us to examine what information children are using when attributing intentionality as well as to examine if theory of mind would impact these decisions or be impacted by the additional contextual information. There is a paucity of research evaluating what other aspects of the context are affecting the decisions surrounding attribution of intentionality. There has been though, a great deal of previous social domain research indicating that group identification, intergroup interactions, and stereotypes all impact exclusion decisions (Killen, Sinno, & Margie, 2006; Smetana, 2006; Horn, 1999; 2003). It is important to explore whether this impact of intergroup interactions, attitudes, and stereotypes will impact decisions surrounding attributions of intentionality.

Developmental Research on Gender Stereotypes

Development of Gender Stereotyping

Children are aware of gender stereotypes as early as preschool, and children's gender stereotyping increases with age (Liben & Bigler, 2002; Ruble & Martin, 1998). Arthur, Bigler, Liben, Gelman, and Ruble (2008) hold that these stereotypes are created based on characteristics of others that adults mark as important, which is often done through labeling. Simply giving an object or a group a common name leads children to make inferences about the properties of those objects (Booth &

Waxman, 2003; Heyman & Gelman, 1999). Additionally, labels are thought to facilitate children's belief that members of a category share important, if non-obvious qualities (Diesendruck, Gelman, & Lebowitz, 1998; Diesendruck & Markson, 2001; Heyman & Gelman, 2000). Gender labeling is very common, with gender labels built into the English language forcing distinctions between male (he) and female (she) (Gelman, Taylor, Nguyen, Leaper, & Bigler, 2004), and even gender labeled occupations (e.g., mailman, actress, cowboy) (Liben, Bigler, & Krogh, 2001). Most children learn gender words between 18 and 24 months (Poulin, Serbin, & Derbyshire, 1998), and there is a strong positive correlation between the acquisition of gender labels and sex typing (Leinbach & Fagot, 1986). In fact, children who acquired gender labels have been shown to have more sex-typed toy preferences than those who had not, while children who were early labelers (27- to 28-months-of-age) had more gender stereotype knowledge at 4 –years-of-age than children who were late labelers (Leinbach & Fagot, 1986).

Martin, Wood, and Little interviewed children 4- to 6-years-of-age and children 8- to 10-years-age and examined how their use of gender stereotyping affected their ability to predict the likelihood of other stereotypically feminine or masculine characteristics when they knew one characteristic of a child whose gender was not mentioned (1990). For the younger children, the target characteristic was toy preference and of the older children, the target characteristic was appearance, personality, occupation, or toy choice (Martin, Wood, & Little, 1990). They found that the younger children had an easier time making associations about toy preference when the characteristic was their own gender, while the older children were able to

make associations for their own gender and the opposite gender. Furthermore, the older children were found to be more extreme in their stereotypical judgments than the younger children. Understanding the developmental progression of gender stereotypes will inform us as to when and how these stereotypes could be impacting the decisions that children are making about intentionality, especially around gender stereotyped toys.

Knowledge of Gender Stereotypes

Knowledge of gender stereotypes for concrete items and activities emerges in the preschool years and reaches ceiling levels by 5- or 6-years-of-age (Ruble & Martin, 1998). Children as young as three begin to understand that items such as hairbrushes, dolls, irons, vacuums, and a needle and thread are associated with females, while bats, balls, shovels, and cars are associated with males (Ruble & Martin, 1998). As children grow older, they continue to develop more detailed knowledge of concrete items, and begin to learn about stereotypes in other domains such as occupations and personality attributes (Ruble & Martin, 1998).

Trautner, Ruble, Kirsten, and Hartmann (2005) reported on a longitudinal study examining children's (interviews starting in kindergarten and continuing for 5 years) gender knowledge of specific objects and activities and children's verbal preference for those same items. In this study, the authors found that stereotypic knowledge is already high at age 5, increases slightly at 6 years, and levels off until age 10. The increase in knowledge from ages 5 to 6 is paired with an increase in rigidity, which begins to decline when kids reach their peak level of knowledge (Trautner et al., 2005). Boys' same-sex preferences remain slightly higher than girls'

throughout the age range of 5- to 6-years-of-age. Additionally boys demonstrate an increase in avoidance of opposite-sex items from age 5- to 9-years-of-age, while girls demonstrate an increase until the age of 7. Furthermore, boys' knowledge of stereotypes at year one was associated with same-sex preference and opposite-sex avoidance at years 1, 2, and 3, while girls' knowledge did not predict preference (Trautner et al., 2005). These male and female differences could be because male toys and items are seen as overall, more desirable, and of a higher status than female toys and items (Antill, Goodnow, Russell, & Cotton, 1996).

These findings prove that children hold gender stereotypes and same-sex preferences as early as 4-year-of-age, use these gender stereotypes to make interpretations about toy preference, activities, and behaviors, indicating that stereotypes could additionally impact decisions that children are making concerning stereotypically male and female toys.

Gender and Status

While it has been shown that both male and female children hold gender stereotypes, Rowley, Kurtz-Costes, Mistry, and Leagans (2007) have found that social status can impact who holds what stereotypes. The authors interviewed 4th, 6th, and 8th grade students who reported on their perceptions of the competence of Black, White, female, and male children in the domains of academia, sports, and music, and found that the children in lower status groups (females and Black children) did not endorse stereotypes that reflected negatively on their own group, but were likely to report stereotypes that favored their social group (Rowley et al., 2007). The children in the high status group (males and White children) endorsed most traditional

stereotypes, whether they were positive or negative, for their social group (Rowley et al., 2007). This effect was found in all the age groups, but the older children were more likely than the younger children to report traditional stereotypes and the status effects were even more pronounced for the older children than the younger children (Rowley et al., 2007).

In addition, female occupations and gender roles are more highly stereotyped than male gender roles and occupations (Shepard & Hess, 1975; Smetana, 1986) most likely reflecting greater societal discomfort with males who exhibit non-traditional gender role behavior than with females who do so (Liben & Bigler, 2002). Furthermore, children as young as elementary school, identify that masculine occupations, as a group, are higher in status than female occupations (Liben, Bigler, & Krogh, 2001). In fact, not only are males higher in social status, masculine items are seen as more desirable than feminine items (e.g., Antill, Goodnow, Russell, & Cotton, 1996). If the social status of the group and the toys as well as the opinion of the gender appropriate behavior of a group is affecting what stereotypes children endorse about their group, it is possible that these differences in status could impact how children evaluate gender stereotyped objects and children acting in counter-stereotypic ways. The female children, in the comparatively lower status group as compared to men, could have a harder time imagining that another female could take a toy on purpose, as it would endorse a negative image of females.

Ingroup/Outgroup Relationships

Ingroup/outgroup research is important to the study of moral reasoning, attribution of intentions, and gender relations as the power of the ingroup can have an impressive impact on the decisions that are made in social situations involving morality, especially in the attribution of intentions, as gender has a dramatically powerful ingroup and outgroup. Patterson and Bigler (2006) were able to create a viable ingroup/outgroup relationship with novel groups of preschoolers (3- to 5-years-of-age) using colors (red or blue). Even though these groups were novel, and either used to organize the classroom (experimental condition), or not at all (control), all the children developed ingroup-biased attitudes. If Patterson and Bigler were able to fabricate novel groups and find ingroup biases that impacted the attitudes of the children, an existing group such as gender with an extreme history of bias and segregation is able to produce immense running ingroup/outgroup attitudes with the power to impact the perception of social interactions and the decisions that are made because of that.

Children, in fact, have been repeatedly shown to exhibit preferences for individuals of their own gender, as well as to self-segregate by gender, and prefer same-sex peers as early as the second and third years of life (Martin, 1989; Martin, Eisenbud, & Rose, 1995; Ruble & Martin, 1998). By their 4th year, children increasingly believe that unfamiliar peers are more likely to have same-sex than other-sex friendships, and this belief is associated with the tendency of children to play exclusively with same-sex peers (Martin, Fabes, Evans, & Wyman, 1999). Martin and colleagues additionally found that children who spent significantly more

time in sex-segregated play became increasingly sex-typed in their play overtime. Powlishta (1995) was additionally able to show that 8- and 9-year-old children report their own gender to have a greater number of positive, and a smaller number of negative traits than the other gender.

There is a long history of general ingroup/outgroup research in the social and developmental literatures (Brewer & Silver, 1978; Foddy, Platow, & Yamagishi, 2009; Patterson & Bigler, 2006; Platow, McClintock, & Liebrand, 1990; Rowley, et al., 2007). The literature consistently points to an individual's favoritism toward their ingroup, with simply referring to a collection of individuals as a group sufficient to produce ingroup favoritism, even when the group membership was random (Brewer, 1979; Tajfel, Billig, Bundy, & Flament, 1971; Tajfel & Turner, 2001). People tend to judge ingroup members as generally better, nicer, and more helpful than outgroup members, as well as more generous, trustworthy, and fair (Boldizar & Messick, 1988; Brewer & Silver, 1978; McAllister, 1995; Platow, et al., 1990). Ingroup members even have expectations of altruistic and fair behavior from ingroup members (Kiyonari, 2002; Yamagishi, Jin, & Kiyonari, 1999; Yamagishi & Kiyonari, 2000). In fact, when Foddy and colleagues examined ingroup relationships, they found that people are more likely to place their trust in members of an ingroup than a more relevant outgroup member, and additionally preferred the ingroup even when there was a stereotype for the ingroup that was more negative than that of the outgroup (Foddy et al., 2009).

Recent research indicates that ingroup favoritism depends on group status, as disadvantages groups frequently favor higher status outgroups (Jost, Banaji, & Nosek,

2004). It is possible that this is because members of low-status groups often internalize negative ingroup stereotypes as a means of justifying the existing social order (Jost & Banaji, 1994). With gender stereotypes, women often associate traits like dependence with females over males (Banaji, Hardin, & Rothman, 1993; Blair & Banaji, 1996) and implicitly prefer male over female leaders (Rudman & Kilianski, 2000). This ingroup favoritism could affect how children are attributing intentions in situations involving both males and females and gender stereotyped toys. Children could have a hard time admitting that same sex peers would purposely take a toy that did not belong to them, indicating that the same sex peer would have taken the toy by accident, thus showing an ingroup bias.

Gender Stereotypes and Punishment Decisions

Karinol and Aida (1997), examined if punishment contingent to the accidental breaking of a toy was affected by the gender stereotype of the toy. They interviewed 80 second grade children from 7- to 81/2-years-of-age concerning two stories in which a neutral (cup), male-stereotyped (truck), or female-stereotyped (doll) toy was held by a target child (male or female) and then accidentally dropped and broken (Karniol & Aida, 1997). Each participant judged one neutral and one sex-stereotyped story, with half of the participants receiving a same-sex story (girl dropping a doll, boy dropping a truck) and half of the participants received an opposite-sex story (girl dropping a truck, boy dropping a doll). The participants were asked to judge how much to punish the target child (scale from no punishment to severe punishment) and to justify this judgment (Karinol & Aida, 1997).

Karinol and Aida (1997) found that both male and female children judged the breaking of the neutral toy in the same manner, not recommending punishment, but when compared to the neutral story, the target children who violated gender stereotypes were judged more severely than those who did not. Additionally, females judged both male and female toy breakers who violated gender stereotypes more severely than those who did not, while males did not (Karinol & Aida, 1997). When the male participants justified their judgments, they referred to toy ownership for everyone and intentions when the toy breaker was male, while the female participants did not refer to toy ownership or intentionality (Karinol & Aida, 1997). When the authors co-varied out toy ownership, they found the differences between males and females disappear, as the boys judged the toy breakers who violated the gender stereotypes as not owning the toy, and the toy breaking to be presumably accidental, as the females held that children should not want a toy that violates gender stereotypes, and judged those who broke an inconsistent toy more severely (Karinol & Aida, 1997).

These results confirm that gender stereotypes are impacting the decisions that children are making in everyday situations involving gender stereotyped toys. Additionally, there were differences in the male and female participants' responses, as the females judged the gender stereotype violators more severely, possibly do to confusion about toy ownership and the intentionality of breaking the toy. Unfortunately, we are unable to disentangle this result as the participants were never asked to judge intentionality, simply punishment. Additionally, since the participants were asked to judge punishment without judgments of intentionality, it is possible

that the children assumed that the breaking of the toy was intentional because they were asked to judge punishment. Having specific questions concerning the intentions of the target child, followed by justifications for those judgments as well as questions and justifications about punishment would allow for a more thorough examination of these affects.

It is also unclear whether any of these children hold gender stereotypes, and how these stereotypes impacted the punishment decisions that were make. Also, unfortunately, no age related trends were found, as the age range was very limited. A more broad range of ages would allow for greater exploration of these possible differences.

Social Domain Research on Gender

Research delving into intergroup relationships with the social-cognitive domain theory began with the exploration of gender related stereotypes (Carter & Patterson, 1982; Killen & Stangor, 2001; Smetana, 1986; Stoddart & Turiel, 1985; Theimer, Killen, & Stangor, 2001). The majority of the research has found that children generally consider conforming or not conforming to gender stereotypes as part of the social conventional or personal reasoning domain, as preschoolers have been shown to consider sex-role deviations as informed by social rules, and not an issue of fairness or justice (Smetana, 1986). Carter and Patterson (1982) asked elementary school-aged children to reason about the flexibility and cultural implications of gender-stereotypic toys and occupations, in addition to table manners and a natural law. The study showed that children evaluated cross-gender behavior to be a social conventional issue, not a moral one, because toys and occupations seen as

gender appropriate were seen that way because that was what most people believe and think, not because it was a moral issue (Carter & Patterson, 1982).

Furthering this line of research, Stoddart and Turiel (1985) interviewed children from 5-years-of-age to 13-years-of-age about their acceptance of cross-gender activities. Stoddart and Turiel discovered a U-shaped curve, as the youngest and oldest children found participation in gender-inconsistent activities as more wrong than did children in middle childhood (1985). The authors assert that this is because the maintenance of gender identity is defined in physical terms for kindergarteners, and thus if a girl was to play with trucks or a boy was to play with dolls, other children might question their gender. In adolescence, the authors found gender identity closely linked with psychological characteristics, and rather than others thinking that you may be the other gender for participating in gender-inconsistent activities, others may exclude you socially for this (Stoddart & Turiel, 1985). These studies indicate that gender-inconsistent behavior is noted by children from preschool up until adolescence, and it is reasoned about using social conventional reasoning. Furthermore, gender-inconsistent behavior is deemed unacceptable by the majority of children because of the reactions of their peers to this behavior (Carter & Patterson, 1982; Smetana, 1986; Stoddart & Turiel, 1985). These beliefs about cross-gender behavior could then impact the decisions that children are making about the intentionality of a child taking a gender-consistent or gender-inconsistent toy.

Killen and colleagues made their first foray into intergroup research through the examination of gender stereotypes (Killen & Stangor, 2001; Theimer, Killen, &

Stangor, 2001). Through this research, it became obvious that children as young as 4-years-of-age can hold stereotypes, specifically gender related stereotypes, and that moral considerations were more salient to these preschool children than stereotypic considerations (Killen et al., 2001). Killen and colleagues interviewed children 3 ½ -, 4 ½ -, and 5 ½ - years-of-age, and asked if it was all right for girls to exclude a boy from playing with dolls or for boys to exclude a girl from playing with trucks (Killen et al., 2001). They found that with age, children judged it wrong to exclude because it would be unfair. Additionally, when the children were asked to make a forced-choice judgment, (if there were to be only room for one more child to play, should the group pick a boy or a girl), stereotypical judgments increased with age, and the majority of the justifications for these judgments fell into the social conventional reasoning category with the children and adolescents noting issues of group functioning (Killen et al., 2001). Furthermore, Killen and colleagues found that the younger preschoolers, 4.5-years-of-age, were more likely than the older preschoolers, 5.5-years-of-age, to base initial judgments on stereotypic knowledge (Killen et al., 2001).

In 2001, Killen and Stangor furthered this line of research and interviewed 9-, 13-, and 15- year-olds regarding exclusion decisions about stereotypic peer activities based on gender and race that were both straightforward and exclusively moral in nature, as well as multifaceted decisions, possibly moral, social conventional, or personal (e.g., a group that excludes: a girl from a baseball club, a boy from a ballet club, an African-American child from a math club, and a European-American child from a basketball club). In the multifaceted situations, two children desired to enter a

group at first with equal qualifications for entrance (e.g. they were both equally good at math when desiring entrance to the math club) and then they were presented with the scenario where the children had unequal qualifications and the child who was consistent with the stereotype associated with the group was better at the activity than the non-stereotypic child (e.g. the European American was better at math than the African American child) (Killen & Stangor, 2001). It was found that the vast majority of straightforward exclusion was regarded as wrong, with the children and adolescents citing moral reasons of unfairness (Killen & Stangor, 2001). Furthermore, it was found that in the multifaceted situations, both moral and social conventional reasons were given for the exclusion decision and when the exclusion decision was moderated by a set of qualifications, (Killen & Stangor, 2001) the participants were more likely to make the non-stereotypic choice in the equal qualifications context than in the context where the qualifications were unequal, with the reasoning often referencing group functioning as why (Killen & Stangor, 2001). The participants therefore chose the stereotypic choice when the functioning of the group was being threatened.

These findings strengthen the claim that not only do young children hold gender stereotypes, they view gender-inconsistent behavior negatively, and these stereotypes and beliefs about gender-inconsistent behaviors can influence decisions that children are making concerning stereotypically male and female toys and the intentionality of a child taking a gender-consistent or gender-inconsistent toy.

Negative Bias in Attribution of Intentions

As noted earlier, children take intentions into account when making interpretations about moral relevance. Understanding the cues that are being used to attribute intentions will help in gaining an understanding of when moral norms are applied to social exchanges and encounters as well as the contextual factors that are impacting the attribution of intentions. Children and adolescents have been shown to have stereotypes that become more entrenched with age (Stangor & Schaller, 1996), and younger adolescents and European Americans have been shown to make accusations based on stereotypic assumptions with less available information than older adolescents, who reserve judgment until more evidence is presented (Horn, Killen, & Stangor, 1999; Killen, Kelly, Richardson, & Jampol, in press). These children and adolescents may use this stereotypic information to make inaccurate attributions of intentions based on gender or ethnic stereotypes. Gaining a better understanding of the ages that this occurs, and the contextual scenarios that encourage this misattribution of intentions will inevitably broaden our understanding of the biases that children and adolescents hold, as well as help to indicate what can be done to ameliorate theses biases.

The majority of the research concerning attribution of intentions focuses on the individuals with the social deficits, and not on the victims of the misattribution of bias (Juvonen, Graham, & Schuster, 2003), missing out on the social and relational contexts within which the victimization is occurring (Graham, 2006). Graham and colleagues have been able to show that ethnic diversity, where no one group is of the numerical ethnic majority, may have distinct psychological benefits that create

"attributional ambiguity", where multiple social cues are present and various appraisals of causality of social predicaments are possible (Graham, 2006). Their research suggests that with these myriad of social cues to draw from in ethnically diverse situations, individuals have a more comprehensive understanding of the various behaviors of individuals and therefore do not focus simply on the superficial and inconsequential physical or situational attributes of an individual or a social encounter, but rather the specifics of that particular encounter, reducing a misattribution of intentions or an over-attribution of bias based on race (Bellmore, Witkow, Graham, & Juvonen, 2004; Graham, 2006; Juvonen, Nishina, & Graham, 2006). The same may apply to gender biases as well; Bigler, Brown, and Markell (2001) have shown that teachers routinely differentiate boys and girls through explicit and implicit messages about gender-specific abilities, and Patterson and Bigler (2006) have found that simply differentiating between two groups can create an ingroup-bias and ingroup-favoritism. Creating classrooms that are diverse in gender and ethnicity may allow for the focus of social encounters to be on the encounter, and not on the ethnicity or gender of the participants in the encounter.

Additional research with children focusing on the relational context associated with hostile attribution bias in children, has shown that attribution of bias varies by the relationship with the child (enemy, friend, neutral) even when social and behavioral reputation is accounted for (Peets, Hodges, Kikas, & Salmivalli, 2007). This indicates that attribution bias is not just a matter of an aggressive child misunderstanding social cues, but can also be a product of stereotypes concerning

contextual variables, and the relationships associated with the interpretation of bias are critical to understanding attributional biases.

Studies conducted by McGlothlin and colleagues explored attributions of intentions with preschool children through an ambiguous picture card tasks featuring both male and female children depicted as either “Black” or “White” (McGlothlin & Killen, 2006). McGlothlin found that European American children from homogeneous school, with low levels of intergroup contact, attributed more negative intentions to the Black character in the card than to the White character (McGlothlin & Killen, 2006). Additionally, with age, and irrespective of school composition, the majority of European-American children were less likely to expect that an interracial peer dyad could be friends.

Together, these findings indicated that school experience and social experience of the participant played a role in the attributions of intentions in peer exchanges. These findings were not about the European American children from the homogeneous schools being more aggressive or having social deficits in their abilities to read social cues, but more likely were about these children being at risk for holding racial biases, possibly due to a lack of social experience with people of different ethnicities. Additionally, another study conducted by Killen and colleagues found that attribution of intention varied by the context of the ambiguous scenario, the ethnicity of the characters in the scenario, the ethnicity of the participant, as well as the age of the participant (Killen, Kelly, Richardson, & Jampol, in press). This provides strength to the argument that social experience plays a role in a misattribution of bias and additionally provides evidence that context needs to be

explored more fully in order to have a deeper understanding of how and why children are attributing intentions in social scenarios and what factors in the context are contributing to the decisions being made.

Current Study

It is clear from the abundant research on moral development, theory of mind, and gender stereotypes, that these are areas ripe for further research. To date, no research has examined the impact of intergroup attitudes and theory of mind competencies on moral judgments and justifications of intentionality. It is proposed, that this line of research will push the field of moral development forward and reveal the new information about the ontogeny of moral reasoning, the impact of theory of mind competency on moral reasoning, and the impact of intergroup stereotypes on moral reasoning. As seen through the previously discussed research the social-cognitive domain model provides a useful heuristic for investigating these issues, and allows for an in depth critical analysis of the evaluations, judgments, and justifications that children and adolescents make concerning everyday situations.

This will allow for a more thorough understanding of children's attribution of intentionality and thus their moral reasoning, and will address if stereotypic expectations are related to children's attributions of intentions and moral judgments. Children who have stereotypic expectations may unknowingly be using these expectations as a reason to assume that a child did not intentionally do something in an ambiguous situation when the action does not conform to a stereotype, and by contrast assign blame to someone when the ambiguous act does conform to stereotypes. This proposed study will examine how theory of mind and intergroup

attitudes, such as gender stereotypes, impact with a child's social evaluations, judgments, and reasoning about intentionality.

Chapter 3: Methodology

Participants

Participants were 44 3 - 4.5 year olds, 40 4.5 - 6 year olds, and 43 6 - 8 year olds (N = 127) from private nursery schools, kindergartens, and elementary schools in the Mid-Atlantic. Participants were evenly divided by gender (63 female, 64 male), and are representative of the diverse metropolitan from which they live (18.9% African American, 60.6% European American, 2.4% Latino, 18% Asian, and 5% other). Only students receiving parental consent were interviewed (see Appendix A for consent forms). The return rate was approximately 75%. Those children who chose not to participate where nearly identical in demographic breakdown to that of the overall school they attended as well as to the demographic breakdown of the sample from the present study. All children whose parents gave permission agreed to be interviewed. Two participants were excluded from the data as they were unable to complete the interview. Power analysis revealed that the sample size of this study was sufficient for a medium effect size at the .05 significance level (Cohen, 1992).

Design

The study involved between-subjects and within-subjects factors for an overall design that included a 2 (gender: Female, Male) X 3 (age: 3-4, 5-6, 7-8 years) X 2 (false belief theory of mind competence: yes, no) X 2 (gender stereotype consistency: consistent, inconsistent) X 2 (identification with agent; ingroup, outgroup) X 2 (status of object; high (truck), low (doll)) model with repeated measures on the last three factors. These variables were not analyzed at the same

time. This statement is exclusively for the purpose of explaining all the variables to be measured.

Three tasks were administered. The main task of this study is a 2 (gender stereotype consistency: consistent, inconsistent) X 2 (identification with the agent: ingroup, outgroup) X 2 (status of object of interest: high (truck), low (doll)) model, with repeated measures on the last two factors. This task includes four scenarios that are varied in order to systematically examine the impact of gender stereotype consistency of the agent in the scenario and the object of interest, ingroup versus outgroup identification of the participant with the agent, and the status of the object of interest on the assessments of intentionality, friendship, punishment, and emotionality of the victim (see Table 1 for task design).

All of the participants were administered both a gender consistent scenario as well as a gender inconsistent scenario. The gender consistent scenario was matched in gender to the gender of the participant. Half of the participants (n = 64) were administered a gender inconsistent scenario in which the agent of the scenario is male and the object of interest is of low status (doll). The other half of the participants (n = 63) were administered a gender inconsistent scenario in which the agent of the scenario is female and the object of interest is of high status (truck).

This design allows for an examination of the distinct impact of ingroup/outgroups identification, participants' identification with the agent, as well as identification with the object (toy).

A gender stereotype task as well a false belief theory of mind task were administered as participant variables for gender stereotype knowledge, tolerance, and flexibility as well as false belief theory of mind competence.

Procedure

The three tasks were administered by a trained researcher in a quiet room at each school. The research assistants went through a rigorous two week training in order to be able to interview the children. The training involved readings of previously done transcripts of interviews completed by a trained research assistant, observations of interviews done by a previously trained research assistant, multiple practice runs of the interview with a trained research assistant, reviews of their own taped practice interviews with a trained interviewer with review of all areas in need of improvement, observations of the children they will be interviewing, as well as a review of their first interview completed on their own. The research assistant will additionally be trained to notice when a child is not paying attention to the interview, how to re-engage the child, how to tell if a child will need to be excluded from the protocol (due to a pre-existing developmental condition such as ADHD, a significant language delay, or any other problem which would interfere with the completion of the interview).

With an age appropriate vocabulary, participants were told that there are no right or wrong answers and that all of their responses are anonymous and confidential. Additionally, participants were told that their participation is voluntary and that they may chose to stop the assessment at any time. Prior to the beginning of the assessment, the children were also be given a chance to practice giving a Likert

score for judgments, being shown a scale with happy faces ranging from a long frown to a large smile (See Appendix D for Likert Scale). The assessment did not continue until it was clear that the child understood how to use a Likert scale. The assessment took an average of 25 minutes to complete.

Measures

The assessment consisted of three tasks (for the complete interview, see Appendix B). Each assessment followed the same order of tasks*: Intergroup Attitudes Attribution of Intentions Task, Theory of Mind Task*, and then the *Gender Stereotype Task. The Intergroup Attitudes Attribution of Intentions Task* as well as the *Theory of Mind Task* were modified from Killen, Mulvey, Richardson, Jampol, & Woodward (in press). The coding categories were correspondingly adapted from the Killen et al., study (in press) as well. The *Gender Stereotype Task* was modified from Signorella, Bigler, and Liben (1993).

Intergroup Attitudes Attribution of Intentions Task

The Intergroup Attitudes Attribution of Intentions Task consists of four possible hypothetical scenarios in which a child, ambiguous as to intent, places a doll (or truck) in his or her own backpack instead of the backpack of the classmate that owns the doll (or truck) while helping the teacher clean up the classroom (see Tables 1 and 2 for more scenario and design details). Each child heard only two scenarios, a gender stereotype consistent scenario in which the two children in the scenario are of the same gender as the participant and the toy is gender stereotype consistent (e.g., a truck for boys and a doll for girls), or a gender stereotype inconsistent scenario which

includes two children of mixed gender, and a toy that is gender stereotype inconsistent for the potential transgressor (e.g., a girl taking a truck or a boy taking a doll) (see Table 1 for task design).

The first scenario was always the gender stereotype inconsistent scenario. Half of the participants were read a gender stereotype inconsistent scenario where the potential transgressor is in the participant's ingroup (e.g., a male participant was read the scenario in which a boy puts a doll into his backpack or a female participant was read a scenario in which a female is putting a truck into her backpack). The other half of the participants were read a gender stereotype inconsistent scenario where the potential transgressor is in the participants' outgroup (e.g. a male participant has the scenario where a girl puts a truck into her backpack or a female participant has the scenario where a boy puts a doll into his backpack). These splits allow for an examination of the importance of ingroup/outgroup identity when making decisions about intentionality in a morally relevant scenario.

The second scenario was matched with the gender of the participant and either involved a girl putting a doll into her own backpack when it belongs to another girl in the class, or a boy putting a truck into his own backpack when it belongs to another boy in the class. This allows for a comparison group to the gender stereotype inconsistent scenario.

Each child had the story read aloud to them, and additionally had picture cards of the children and objects depicted in the stories shown to them in correspondence to the actions of the stories.

Dependent Measures for the Intergroup Attitudes Attribution of Intentions Task

The two scenarios in this measure contain the same 9 assessments, so as to allow for a direct comparison between scenarios. The assessment, *Intentions of Transgressor*, includes one question which reads, 1) "Whose backpack did (transgressor's name) think he/she was putting the toy into?" Two choices are provided; the transgressor's backpack and the victim's backpack. This assessment showed what the participant believes the intentions of the transgressor are.

Judgment and Justification of Transgressor's Intentions, includes 3 questions. The questions read, 2) "When (transgressor's name) put the toy truck (or doll) into the backpack did she (or he) think she (or he) was doing something that was all right or not all right?" 3) "How all right (or not all right depending on the answer they provided for the previous question) did (transgressor's name) think he/she was for doing that?" 4) "Why?" This assessment allows the participant to judge the transgressor's intentions as all right or not all right, place this judgment on a Likert scale (1, not all right to 4, all right), and then justify the judgment. This assessment provided information about the participants' moral judgments of the transgressor's intentions as well as their reasoning to support this justification, and can provide insight into what information they are using to make their judgments. The coding categories for the justifications were based on previous research used to analyze social reasoning (Killen et al., 2001; Smetana, 1995) and were modified from the Killen et al., 2009 (see Table 3 for coding examples).

Judgment and Justification of Transgressor's Action by Participant, asks the participant first to judge if they think the transgressor was doing something that was all right or not all right, then to place this judgment on a Likert scale from 1 (not all right) to 4 (all right), and then to justify this judgment. The questions read, 5) "When (transgressor's name) put the truck (or doll) into the backpack, do you think he (or she) was doing something that was all right or not all right?" 6) "How all right /not all right do you think she was for doing that?" 7) "Why?" The justifications were coded using the same coding categories as the fourth assessment. This assessment provided further information as to the participant's own evaluation of the intentions of the transgressor, and the reasoning that they used to support this judgment.

Friendship Judgment and Justification asks the participant to determine how much they think the transgressor and the victim like each other based on a Likert scale from 1 (not at all) to 4 (a lot), and then to justify this judgment. The questions read, 8) "How much do you think (transgressor's name) and (victim's name) like each other?" 9) "Why?" The justification for this assessment was coded using a modified version of the coding categories used in the previous assessments. This assessment provided information about how the participant views the relationship between the transgressor and the victim, which may have implications as to the intentionality judgments and justifications. In the cross gender scenario, scenario 1, this question also provided information as to the participant's beliefs about males and females being friends, which could be related to their stereotype assessment.

Theory of Mind Location Change, asks the participant to determine if the victim will know where to look for his/her toy when he/she re-enters the classroom,

and to justify their determination. The questions read, 10) "Where will (victim's name) look for his (or her) truck (or doll)?" 11) "Why will (victim's name) look there?" Three possible choices are provided for the location decision, on the floor (which is where he (or she) left it and should look for it), in transgressor's backpack (where it actually is), in victim's own backpack). This false belief theory of mind assessment is considered first order false belief theory of mind, as the participant is asked to interpret the scenario from just one person's point of view. The coding categories for the justification were modified from the Killen et al., 2009 study.

Emotion Judgment of Victim, includes one questions which reads, 12) "When (victim's name) finds out his (or her) truck (or doll) is not where he (or she) left it, how will (victim's name) feel?" This assessment provides information as to the participants' belief about the victim's emotional reaction to the loss of the toy, which may have implications as to the judgments of intentionality and punishment of the transgression.

Punishment of Transgressor, includes three questions which read, 13) "Should (transgressor's name) get in trouble for putting the truck (or doll) into the backpack (researcher points to transgressor's backpack)?" 14) (If the participant answers yes) "How much trouble should (transgressor's name) get in?" 15) "Why?" (or why not depending on the answer to the previous question)? A choice of "a little" or "a lot" is provided for the second question in this assessment. The justifications for the final question in the assessment were coded using the same coding categories as the justifications of intentionality. This assessment provided additional information as to

the participants' judgments of intentionality of the transgressor's actions, as well as the participants' feelings about the severity of this transgression.

Second Order Theory of Mind, includes two questions which read, 16) "What did (victim's name) think that (transgressor's name) thought he (or she) was trying to do?" 17) "Why?" Two choices were provided for this assessment, "Put (victim's name) truck (or doll) in (transgressor's name) backpack," or "Put (victim's name) into (victim's name) backpack". The justifications were coded using the same coding categories as the False Belief Theory of Mind, Location Change assessment.

Emotion Judgment and Justification after Knowledge of Transgression, includes three questions which read, 18) "How will (victim's name) feel about (transgressor's name) now that he (or she) knows (transgressor's name) put the truck (or doll) into this backpack (researcher points to transgressor's backpack)?" 19) "How good (or bad, depending on the participant's answer to the previous question) will (victim's name) feel about (transgressor's name)?" 20) "Why?" A Likert scale was provided for the second question in this assessment from 1 (bad) to 4 (good). The justifications were coded using the same coding categories as the Liking justification. This assessment provided further information as to their feelings concerning the transgression, the participants' determinations of intentionality, and the participant's interpretation of the severity of the transgression.

Coding Categories for Justifications

The coding categories for the justification were based on previous research used to analyze social reasoning (Killen et al., 2001; Smetana, 1995) and were modified from the Killen et al., 2009 study. There were 3 superordinate categories,

each with subcategories. Coding is conducted to categorize children's spontaneous judgments and justifications. The 3 superordinate categories are: (A) Moral, (B) Social Conventional, (C) Personal, and (D) Other (see Table 3 for further explanation and delineations).

A justification of (A) Moral was a response that involved a victim and included a focus on the subcategories of 1) psychological harm (e.g., "She'll be sad that she doesn't have her doll"); 2) negligence (e.g., "He should have looked at the initials"), 3)ownership and wrongfulness of stealing (e.g., "He wanted to take it home", "She took it to her house", "It's not okay to steal"); 3) prosocial behavior (e.g. "She was trying to help the teacher."); and 5) accident/lack of negative intentions (e.g., "She put it in the wrong backpack by accident."). A justification of (B) Social Conventional was a response that involved social rules and conventions and included a focus on the subcategories of 6) deference to the rules, a response that was based on rules and authority and involved a focus on deference to the rules (e.g., "It is against the rules to take something that is not yours", "She was just doing what she was told"); and 7) gender stereotypes, a response that referred to the gender stereotypes (e.g., "Girls don't want to take trucks home; it must have been an accident" or "Girls don't play with trucks, she shouldn't have done that.", "Boys don't steal"). A justification of (C) Personal was a response that was based on a personal decision, under no authority dictates and included a focus on the subcategory of 8) selfish desires (e.g., "She will have what she wants"); and justifications that did not fit into the previous categories was designated as (D) Other with the subcategories of 9)

undifferentiated (e.g., “I don’t know; It’s bad; Because it’s good”) or 10) incomplete or inaudible.

Reliability of the coding system was calculated using two coders who independently code all 10 coding categories for 20% of the surveys. Inter-rater reliability was 98%, and was determined by the percent agreement between the coders as well as the more conservative Cohen’s kappa statistic, 0.98, which adjusts for chance agreement.

False Belief Theory of Mind Task

The False Belief Theory of Mind task (Killen et al., 2009) consisted of two short scenarios that are prototypical Theory of Mind Tasks for false belief/false contents and change of location. The first scenario, false contents, read, “See this box (pointing to a crayon box)? This is a crayon box. Now, here is Sarah. She is cleaning up the classroom and puts some crackers in the empty crayon box.” The change of location scenario read, “Lenny is using markers before recess over at the art table. Lenny goes outside to play and the teacher, Mr. Jones puts the markers in the cabinet.” The scenarios were read aloud to the participants, and photos of the objects within the scenarios were shown in concordance with the reading of the scenario.

Dependent Measures for the False Belief Theory of Mind Task

The first assessment, *False Belief Theory of Mind*, read 1) “When the other children come back in from playing outside, what will they think is in the crayon box?” This assessment discerned if the participant could see someone else’s point of

view when they had a false belief. The second assessment, *False Contents – Information Accessible,* read, 2) "Did the children who were playing outside see Sarah put the crackers in the box?" This assessment determined if the participant could access the information that they were just read concerning the crackers being put into the crayon box. The third assessment, *False Contents – Real Contents,* read, 3) "What is really in the crayon box?" This assessment determined if the participant could indicate the true contents of the crayon box. These three assessments, taken together, provided an indication as to the participant's false belief theory of mind with respect to a false contents scenario.

The fourth assessment, *False Belief Theory of Mind – Location Change,* read, 1) "When Lenny comes back inside from recess, where will he look for the markers?" This assessment discerned if the participant could see someone else's point of view when they had a false belief. The fifth assessment, *Location Change – Information Accessible,* read, 2) "Did Lenny see where Mr. Jones put the markers?" This assessment determined if the participant could accurately remember the information that was read to them in the scenario. The sixth assessment, *Location Change – Real Location,* read, 3) "Where are the markers really located?" and determined if the participant could indicate the true location of the markers. These three assessments taken together, provided an indication as to the participant's false belief theory of mind in a location change scenario, and when combined with the previous three assessments, provided a strong indication as to the participants' theory of mind for false belief tasks in general.

Gender Stereotype Task

The *Gender Stereotype Task* consisted of the presentation of 9 photos of toys; three that were stereotypically female (beads, kitchen set, jump rope), three that were stereotypically male (airplane, football, toolset), and three that were gender neutral (marbles, puzzles, and paint set). The toys were chosen based on previous research by Liben and Bigler (2002) and the task was based on a survey by Signorella and colleagues (1993).

Dependent Measures for the Gender Stereotype Task

The first assessment, *Rule Knowledge*, read, 1) "Who usually plays with this toy?" with the choices of boy, girl, or both provided. This assessment provided information as to the participant's knowledge of gender stereotypes with regards to common toys. The second assessment, *Rule Flexibility*, read, 2) "Who can play with this toy?" with the choices of boy, girl, and both provided again. This assessment provided information as to the participant's flexibility with gender stereotyped toy play. The third assessment, *Tolerance*, read, 3) "How much would you like it if a (read opposite gender as one provided for answer to the second assessment) wanted to play with this toy?" A Likert scale was provided for this assessment from 1 (not at all) to 4 (a lot). This assessment provided information as to the participant's tolerance of the opposite gender playing with the gender stereotyped toy (see appendices C and D for interview and drawings).

Plan for Analysis

Hypotheses were tested using ANOVA (see Table 4 for further details concerning the hypotheses). Follow up tests were conducted using Bonferroni to control for Type I errors, and the Bonferroni adjustment was made automatically when needed. Dichotomous responses were coded 0 or 1. Justifications were indicated as proportions of responses for each respective coding category. The False Belief Theory of Mind Task was transformed into an independent variable by summing the score of the three false content assessments and the three location change assessments and converted into a dichotomous variable of either possessing a Theory of Mind or not possessing a Theory of Mind. Participants had to answer all six assessments correctly in order to indicate that they possess a false belief theory of mind. A 3 (age: 3-4, 5-6, 7-8 years) X 2 (Theory of Mind, no Theory of Mind) is hypothesized to indicate that with age, children acquire Theory of Mind. The *Gender Stereotype Task* was converted into a scale indicating degree of gender stereotype tolerance and flexibility (a high score will indicate greater flexibility and tolerance), and additionally transformed in a dichotomous high/low variable using a median split, in order to incorporate it as an independent variable in the analysis of the *Intergroup Attitudes Attribution of Intentions Task.* A 3 (age: 3-4, 5-6, 7-8 years) X 2 (gender: male, female) is hypothesized to indicate that stereotype flexibility increases with age, and varies by gender, with males consistently indicating lower gender stereotype flexibility and tolerance than females.

Intergroup Attitudes Attribution of Intentions Task

A 3 (age of participant: 3-4, 5-6, 7-8) X 2 (gender of participant: female, male) X 2 (Theory of Mind, No Theory of Mind) X 2 (gender stereotype flexibility and tolerance: high, low) X 2 (scenario: stereotype inconsistent, stereotype consistent) ANOVA with repeated measures on the last factor was conducted on the assessments in the *Intergroup Attitudes Attribution of Intentions Task*.

Hypotheses regarding the *Intentions of Transgressor,* the *Judgment of the Transgressor's Intentions,* and the *Judgment of Transgressor's Action by Participant* assessments were tested by conducting 3 (age: 3-4, 5-6, 7-8 years) X 2 (gender: male, female) X (theory of mind, no theory of mind) X 2 (gender stereotype flexibility and tolerance, high, low) X 2 (scenario; gender inconsistent, gender consistent) ANOVAs with repeated measures on the final factor for each dependent variable. It was hypothesized that with *Intentions of Transgressor,* the *Judgment of the Transgressor's Intentions,* and the *Judgment of Transgressor's Action by Participant* participants for the gender inconsistent story as well as the gender consistent story, would indicate a main effect for age, gender, false belief theory of mind, and gender stereotype flexibility and tolerance. It was hypothesized that participants would indicate the action to be more all right (less negative intentions) as age increased. Also, it was hypothesized that participants who have a false belief theory of mind would indicate the action to be more all right, as they would be more likely to be able to see the accidental nature of the transgression.

Additionally, it was hypothesized that there would be a main effect for stereotype flexibility, as it was expected that participants low in gender stereotype

flexibility and tolerance would be more likely to indicate that the action is more not all right, or less likely to be able to see the accidental nature of this transgression. Furthermore, a 2 (scenario: gender inconsistent, gender consistent) X 2 (gender stereotype flexibility and tolerance, high, low) ANOVA was hypothesized to show that participants that are low in gender stereotype flexibility and tolerance would see more negative intent in the gender consistent scenario than in the gender consistent scenario. Also, it was hypothesized that participants will find the action to be more all right for ingroup transgressors than for outgroup transgressors for the stereotype inconsistent scenario and the stereotype consistent scenario.

Additionally, follow-up testing was hypothesized to indicate that participants with low stereotype flexibility and tolerance would indicate the action to be more all right for the stereotype inconsistent scenario than for the stereotype consistent scenario.

Hypotheses regarding the *Justification of Transgressor's Intentions, and Justification of Transgressor's Action by Participant* were tested by conducting individual 3 (age: 3-4, 5-6, 7-8 years) X 2 (gender: male, female) X 2 (Theory of Mind, no Theory of Mind) X 2 (Gender stereotype flexibility and tolerance; high, low) X 2 (scenario; gender inconsistent, gender consistent) ANOVAs with repeated measures on the final factor for each reasoning category (moral, social convention – deference to the rules, social conventional – gender stereotype, or personal). A main effect for age was hypothesized to reveal more social conventional reasoning indicating deference to the rules, as age increases, and with the indication of a false belief theory of mind. Additionally, it was hypothesized that there would be more

social conventional reasoning indicating gender stereotypes for male participants, as well as those participants with lower gender stereotype flexibility and tolerance. It was also hypothesized that there would be more "Social Conventional – gender stereotype" reasoning in the gender inconsistent scenario than in the gender consistent scenario.

Hypotheses regarding the *Liking Judgment* assessment were tested by conducting a 3 (age: 3-4, 5-6, 7-8 years) X 2 (gender: male, female) X 2 (Theory of Mind, no Theory of Mind) X 2 (Gender stereotype flexibility and tolerance; high, low) X 2 (scenario; gender inconsistent, gender consistent) ANOVA with repeated measures on the final factor. Main effects for age and stereotype flexibility and tolerance were expected, as it was expected that with age, participants would think that male and female children cannot be friends and that participants who have less stereotype flexibility and tolerance would think that male and female children cannot be friends than participants with more flexibility and tolerance.

Hypotheses regarding the *Theory of Mind Location Change* assessment were tested by conducting a 3 (age: 3-4, 5-6, 7-8 years) X 2 (gender: male, female) X 2 (Theory of Mind, no Theory of Mind) X 2 (Gender stereotype flexibility and tolerance; high, low) X 2 (scenario; gender inconsistent, gender consistent) ANOVA with repeated measures on the final factor. A main effect of age and false belief theory of mind was hypothesized, as with age, children acquire false belief theory of mind for location change in a multifaceted scenario, and with a general indication of false belief theory of mind, participants would be more likely to correctly indicate

where the victim should look for the toy, thus indicating that they have a false belief theory of mind in a morally relevant, multifaceted scenario.

Hypotheses regarding the *Emotion Judgment of Victim* assessment were tested by conducting 3 (age: 3-4, 5-6, 7-8 years) X 2 (gender: male, female) X 2 (Theory of Mind, no Theory of Mind) X 2 (Gender stereotype flexibility and tolerance; high, low) X 2 (scenario; gender inconsistent, gender consistent) ANOVA with repeated measures on the final factor. It was expected that all participants would say that the victim felt badly when the toy was moved. It was also expected that participants would indicate that the victim would feel more badly in the gender consistent scenario than in the gender inconsistent scenario, and that the children with low gender stereotype flexibility and tolerance as well as the male participants would indicate that the victim would feel more badly than the participants with high gender stereotype flexibility and tolerance, and the female participants.

Hypotheses regarding the *Punishment of the Transgressor* assessment were tested by conducting a 3 (age: 3-4, 5-6, 7-8 years) X 2 (gender: male, female) X (theory of mind, no theory of mind) X 2 (gender stereotype flexibility and tolerance, high, low) X 2 (scenario; gender inconsistent, gender consistent) ANOVA with repeated measures on the final factor. It was expected that participants would indicate a main effect for age, gender, theory of mind, and gender stereotype flexibility and knowledge. The participant was expected to indicate not to punish, or to punish to a lesser degree as age increases and with an indication of having false belief theory of mind.

Additionally, it was expected that participants with low stereotype flexibility and tolerance, and the male participants, would indicate the action to be more punishable for the stereotype inconsistent scenario than for the stereotype consistent scenario. Furthermore, it was expected that participants will find the action to be more all right for an ingroup transgressor than for outgroup transgressors for the stereotype inconsistent scenario and the stereotype consistent scenario.

Hypotheses regarding the *Punishment of Transgressor, Justification* were tested by conducting individual 3 (age: 3-4, 5-6, 7-8 years) X 2 (gender: male, female) X 2 (Theory of Mind, no Theory of Mind) X 2 (Gender stereotype flexibility and tolerance; high, low) X 2 (scenario; gender inconsistent, gender consistent) ANOVAs with repeated measures on the final factor for each reasoning category (moral, social convention – deference to the rules, social conventional – gender stereotype, or personal). A main effect for age was expected to reveal more social conventional reasoning – deference to the rules as age increases, and with the indication of a false belief theory of mind. Additionally, it was expected that there would be more social conventional reasoning indicating gender stereotypes for male participants, as well as those participants with lower gender stereotype flexibility and tolerance. It was also expected that there would be more social conventional reasoning using gender stereotypes in the gender inconsistent scenario than in the gender consistent scenario.

Hypotheses regarding the *Second Order Theory of Mind* assessment were tested by conducting a 3 (age: 3-4, 5-6, 7-8 years) X 2 (gender: male, female) X 2 (Theory of Mind, no Theory of Mind) X 2 (Gender stereotype flexibility and

tolerance; high, low) X 2 (scenario; gender inconsistent, gender consistent) ANOVA with repeated measures on the final factor. There was expected to be a main effect of age and false belief theory of mind, as children acquire a second order theory of mind after they acquire a first order theory of mind, and thus theory of mind is requisite for being able to accurately answer this assessment. An additional 2 (scenario: inconsistent gender, consistent gender) X 2 (scenario: inconsistent gender, consistent gender) ANOVA was conducted for this assessment in comparison to *Judgment of Transgressor's Intentions* as it was expected that participants with second-order theory of mind would think that transgressor has less negative intentions than those participants who do not have a second order theory of mind.

Hypotheses regarding the *Emotion Judgment and Justification after Knowledge of Transgression* assessment were tested by conducting 3 (age: 3-4, 5-6, 7-8 years) X 2 (gender: male, female) X 2 (Theory of Mind, no Theory of Mind) X 2 (Gender stereotype flexibility and tolerance; high, low) X 2 (scenario; gender inconsistent, gender consistent) ANOVA with repeated measures on the final factor for the judgment decision and for the judgment Likert.

Chapter 4: Results

Introduction

Hypotheses were tested by conducting Univariate ANOVAs for the dependent measures within each scenario of the Intergroup Attitudes Attribution of Intentions Task, and using repeated measures ANOVAs in order to make comparisons between the two scenarios within the Intergroup Task (see Table 5 for a summary of the dependent variables, Table 6 for a summary of the independent variables, and Table 4 for a summary of the hypotheses). Follow-up tests to examine interaction effects were done using one-way ANOVAs.

The results of the False Belief Theory of Mind Task and the Gender Stereotype Task will be discussed first, as the responses to these tasks were used as independent variables that were then used to examine the various dependent variables in the Intergroup Attitudes Attribution of Intentions Task. The results of the Intergroup Attitudes Attribution of Intentions Task is organized so as to first review the hypotheses and results pertinent to the intentionality related assessments first followed by the hypotheses and results pertinent to the punishment acceptability assessment. Second, the hypotheses and results associated with the justifications for each of those assessments will be addressed. Third, the hypotheses and results of the false belief theory of mind embedded assessment will be attended to.

False Belief Theory of Mind Task

The participant responses to the false belief theory of mind task was converted into an independent variable by summing the score of the three false content assessments and the three location change assessments and converting the task into a dichotomous variable of either possessing a false belief theory of mind or not possessing a false belief theory of mind. Participants had to answer all six assessments correctly to indicate that they possessed a false belief theory of mind. A Univariate ANOVA for age (age: 3-4, 5-6, 7-8 years) was conducted for the false belief task. A significant effect for age, F (2, 118) = 42.79, p = .000, η_p^2 = .43, (M = .10, SD = .30; M = .44, SD = .50; M = .88, SD = .33 for ages 3-4, 5-6, 7-8 years respectively) was found indicating that children were more likely to pass the false belief task with age (see Table 7).

The Gender Stereotype Task

The *Gender Stereotype Task* was converted into three individual scales separately indicating degree of gender stereotype tolerance, flexibility, and knowledge. The scales, gender stereotype flexibility (How much would you like it if a child of the opposite gender were to play with this toy? (Likert scale: 1 = not at all, 4 = a lot); gender stereotype tolerance (Who can play with this toy? girls, boys, either); and gender stereotype knowledge (Who usually plays with this toy? girls, boys, either), were then each converted into separate dichotomous high/low variables using median splits. For the variable gender stereotype flexibility, high gender stereotype flexibility indicated the participant was flexible in liking any child playing with any toy regardless of the gender stereotype, and low flexibility indicated the participant

was not flexible in liking a child playing with counter-stereotypic toys. High gender stereotype tolerance indicated the participant was tolerant of any child playing with any toy, regardless of the gender stereotype, while a low gender stereotype tolerance indicated the participant was intolerant of children playing with counter-stereotypic toys. Finally, high gender stereotype knowledge indicated the participant associated gender with the toys along stereotypic categories. Low gender stereotype knowledge indicated the participant did not associate gender with the stereotypic toys.

These dichotomous variables were then incorporated as independent variables in the analysis of the *Intergroup Attitudes Attribution of Intentions Task.* Three separate ANOVA analyses were conducted; one for gender stereotype tolerance, one for gender stereotype flexibility, and one for gender stereotype knowledge. It was hypothesized that participants would indicate an increase in gender stereotype knowledge, tolerance, and flexibility with age, as well as with an indication of having a false belief theory of mind. It was also hypothesized that male participants would indicate lower gender stereotype tolerance and flexibility than female participants.

A 3 (age: 3-4, 5-6, 7-8 years) X 2 (gender: male, female) X 2 (False belief: fail, pass) ANOVA was conducted on gender stereotype tolerance, and in concert with the hypotheses, a significant effect was found for age, F (2, 121) = 12.57, p = .000, η_p^2 = .17, (M = .33, SD = .48; M = .28, SD = .45; M = .738, SD = .44 for ages 3-4, 5-6, 7-8 years respectively), revealing that with age, children were more tolerant of gender counter-stereotypic play. A main effect for false belief, F (1, 115) = 14.29, p = .000, η_p^2 = .11, (M = .30, SD = .46; M = .63, SD = .49 for no false belief and false belief, respectively), revealed that children who passed the false belief task were more

tolerant of counter-stereotypic play. These main effects indicate that, with age, there was an increase in tolerance of counter-stereotypic play, and additionally as one passes the false belief task, there is also an increase in tolerance of counter-stereotypic play. No significant effects for participant sex and no interaction effects were found.

For gender stereotype flexibility, contrary to hypotheses, no significant effects or interaction effects for age, false belief theory of mind, or participant sex were found.

A 3 (age: 3-4, 5-6, 7-8 years) X 2 (gender: male, female) X 2 (False belief: fail, pass) ANOVA was conducted on gender stereotype knowledge, and the hypotheses for age and gender were confirmed. A significant effect for age, F (2, 121) = 42.72, p = .000, η_p^2 = .24, (M = .07, SD = .26; M = .28, SD = .45; M = .62, SD = .49 for ages 3-4, 5-6, 7-8 years respectively), indicated that as age increased so did knowledge about gender stereotypes. Furthermore a significant effect for participant sex was found, F (1, 122) = 6.17, p = .014, η_p^2= .05, (M = .42, SD = .50; M = .22, SD = .42 for female and male participants respectively), indicating that female participants had more knowledge of gender stereotypes than did the male participants.

A significant effect for false belief was also found, F (1, 115) = 21.78, p = .000, η_p^2 = .16, (M = .15, SD = .36; M = .53, SD = .50), revealing, contrary to the hypothesis, those who pass the false belief task are more likely to have a higher knowledge of gender stereotypes. No significant interaction effects were found.

Intergroup Attitudes Attribution of Intentions Task

Judgment of Transgressor's Intentions

Overall results. In order to discern if the participants, across the gender inconsistent and gender consistent scenarios, attributed negative intentions to the act of placing another child's toy into his/her own backpack (not all right, indicated by a 1 or 2 on the Likert scale) or positive (all right, indicated by a 3 or 4 on the Likert scale) frequency data were run for the combined answers to the judgment of the transgressor's intentions. The frequency data revealed 72% of the participants indicated the action to be not all right. No significant interaction effects were found, and therefore the individual Univariate ANOVAs are reported.

Two separate assessments of intentionality were given to the participants in order to examine the participants' interpretation of the intentions of the transgressor in the scenarios. The judgment of transgressor's intentions assesses if the participant believes that the transgressor in the story thinks that he/she did something all right or not all right. The second assessment of intentionality, judgment of the transgressor's action by the participant, is examined and reported separately. That assessment measures if the participant him/herself believes the transgressor did something all right or not all right.

Age related significant effects. In order to test the hypothesis that children would attribute less negative intentions with age two separate Univariate ANOVAs were conducted. The Univariate ANOVA for age (age: 3-4, 5-6, 7-8 years) on the intention judgment in the gender inconsistent scenario confirmed the hypothesis, F (2, 124) = 8.44, p = .000, η_p^2 = .12, (M = 1.66, SD = 1.01; M = 1.75, SD = 1.19; M =

2.54, *SD* = 1.05 ages 3-4, 5-6, 7-8 years respectively), indicating that with age, the children were less likely to attribute negative intention. The Univariate ANOVA for age (age: 3-4, 5-6, 7-8 years) on the intention judgment in the gender consistent scenario, F (2, 122) = 7.05, p = .001, η_p^2 = .10, (*M* = 2.00, *SD* = 1.06; *M* = 2.30, *SD* = 1.17; *M* = 2.86, *SD* = .98 for ages 3-4, 5-6, 7-8 years respectively), additionally indicated that with age, children were less likely to attribute negative intentions (see Figure 1).

False belief significant effects. It was hypothesized that children who passed the false belief task would indicate less negative intent on the judgment of intention assessment. With respect to this hypothesis, two separate Univariate ANOVAs were conducted. The Univariate ANOVA for false belief (false belief: pass, fail) on the intention judgment (intent judgment: Likert scale ranging from 1, not all right to 4, all right) confirmed the hypothesis, F (1, 116) = 14.71, p =.000, η_p^2 = .11, (*M* = 1.64, *SD* = .97; *M* = 2.42, *SD* = 1.24 for no false, and false belief respectively), revealed that children who passed the false belief task were more likely to indicate less negative intentions than children who did not pass. The Univariate ANOVA for false belief (false belief: pass, fail) conducted on the intention judgment in the gender consistent scenario, F (1, 115) = 10.28, p = .002, η_p^2 = .08, (*M* = 2.08, *SD* = 1.09; *M* = 2.74, *SD* = 1.14 for no false belief and false belief respectively) additionally confirmed the hypothesis by revealing that children who passed the false belief task were more likely to indicate less negative intent (see Figure 2).

Gender stereotype significant effects. In regards to the hypothesis that children would indicate less negative intent when they indicate high levels of

stereotype tolerance and flexibility, and low stereotype knowledge, separate Univariate ANOVAs were conducted for each of the stereotype variables, tolerance, flexibility, and knowledge.

The Univariate ANOVA (gender stereotype tolerance: low, high) on the intention judgment for the gender inconsistent scenario confirmed the hypothesis, F (1, 122) = 10.60, p = .001, η_p^2 = .08, (M = 1.68, SD = 1.04; M = 2.32, SD = 1.16 for participants with low stereotype tolerance and high stereotype tolerance respectively), and revealed that participants with high tolerance of counter-stereotypic play indicated less negative intent than those with low tolerance. Additionally, a Univariate ANOVA (gender stereotype tolerance: low, high) on the intention judgment for the gender consistent scenario confirmed hypothesis, F (1, 121) = 9.10, p = .003, η_p^2 = .07, (M = 2.10, SD = 1.07; M = 2.70, SD = 1.09 for participants with low gender stereotype tolerance and high stereotype tolerance respectively), and signified again that participants with high gender stereotype tolerance indicated less negative intent than those participants with low stereotype tolerance.

This hypothesis was further confirmed for stereotype flexibility for the gender inconsistent scenario and consistent scenarios. The Univariate ANOVA (gender stereotype flexibility: low, high) on the intention judgment for the gender inconsistent scenario, F (1, 122) = 4.93, p = .028, η_p^2 = .039, (M = 1.78, SD = 1.10; M = 2.24, SD = 1.15 for participants with low gender stereotype flexibility and high stereotype flexibility respectively), indicated that participants with high flexibility with gender stereotypes revealed less negative intent than those with low flexibility. Furthermore, the Univariate ANOVA (gender stereotype flexibility: low, high) on the intention

judgment for the gender consistent scenario, F (1, 121) = 4.12, p = .043, η_p^2 = .03, (M = 2.21, SD = 1.19; M = 2.62, SD = 0.97 for participants with low stereotype flexibility and high stereotype flexibility respectively), additionally indicated that participants with high stereotype flexibility revealed less negative intent than those participants with low stereotype flexibility.

The Univariate ANOVA (gender stereotype knowledge: low, high) on the intention judgment for the gender inconsistent scenario, F (1, 122) = 6.96, p = .009, η_p^2 = .05, (M = 1.79, SD = 1.07; M = 2.35, SD = 1.21 for participants with low gender stereotype knowledge and high stereotype knowledge respectively), contrary to what was hypothesized, signified that participants with high knowledge of gender stereotypes indicated less negative intent than those with low knowledge of gender stereotypes. The Univariate ANOVA (gender stereotype knowledge: low, high) on the intention judgment for the gender consistent scenario, F (1, 121) = 10.37, p = .002, η_p^2 = .08, (M = 2.16, SD = 1.08; M = 2.83, SD = 1.08 for participants with low stereotype knowledge and high stereotype knowledge respectively), in concert with what was found with gender stereotype tolerance and flexibility, but contrary to hypothesis, also showed that participants with high stereotype knowledge indicated less negative intent than those participants with low stereotype knowledge.

Differences between gender consistent and gender inconsistent scenarios. In order to test the hypothesis that overall, participants would view the action of a child placing another child's toy into their own backpack as more negative for the gender consistent scenario than for the gender inconsistent scenario, a repeated measures ANOVA for scenario (scenario: gender inconsistent, gender consistent) was

conducted on the intentionality judgment. A significant effect was found for the intentionality judgments, F (1, 124) = 13.86, p = .000, η_p^2 = .10, (M = 1.98, SD = 1.15; M = 2.38, SD = 1.12 for participants on the gender inconsistent scenario and the gender consistent scenario respectively), and indicated that, contrary to the hypothesis, participants found the action in the gender inconsistent scenario more negative than the action in the gender consistent scenario (see Figure 3).

Ingroup and outgroup differences. In order to test the hypothesis that participants would find the action more all right for ingroup transgressors than for outgroup transgressors, a repeated measures ANOVA for scenario (scenario: male transgressor (ingroup), female transgressor (outgroup)) was conducted for the male participants on the intention judgment, however no significant effects were found. An additional repeated measures ANOVA for scenario (scenario: male transgressor (outgroup), female transgressor (ingroup)) was conducted for the female participants on the intention judgment, and again, no significant effects were found.

Judgment of Transgressor's Action by Participant

Overall results. This second assessment of intentionality, judgment of the transgressor's action by the participant, measures if the participant him/herself believes the transgressor did something all right or not all right (Likert scale: 1, not all right; 4, all right). In order to discern if the participants, across the gender inconsistent and gender consistent scenarios, attributed negative intentions to the act of placing another child's toy into his/her own backpack (not all right, indicated by a 1 or 2 on the Likert scale) or positive (all right, indicated by a 3 or 4 on the Likert scale) frequency data were run for the combined answers to the judgment of the

transgressor's action by the participant. The frequency data revealed 87.8% of the participants indicated the action to be not all right. No significant interaction effects were found, and therefore the individual Univariate ANOVAs are reported.

Age related main effects. Regarding the hypothesis that children would indicate the ambiguous act would be more all right (indicate less negative intent) with age, a Univariate ANOVA for (age: 3 - 4, 5 - 6, 7 - 8 years) was conducted for age on the intention judgment in the gender inconsistent scenario. No significant effects were found. An additional Univariate ANOVA for (age: 3 - 4, 5 - 6, 7 - 8 years) was conducted for age on the intention judgment in the gender consistent scenario. Again, no significant effects were found.

False belief theory of mind main effects. In order to test our expectations that children who pass the false belief task would indicate less negative intent, a Univariate ANOVA was conducted for false belief (false belief: pass, fail) on the intent judgment in the gender inconsistent scenario, F (1, 116) = 4.43, p = .038, η_p^2 = .038, (M = 1.42, SD = .81; M = 1.88, SD = .93 for no false belief and false belief respectively), and indicated that children who passed the false belief task in fact did indicate less negative intent (see Figure 2). The Univariate ANOVA conducted for false belief (false belief; pass, fail) on the intent judgment in the gender consistent scenario did not reveal any significant effects.

Gender stereotype main effects. As with the previous intent judgment, the hypotheses regarding gender stereotype tolerance, flexibility, and knowledge for the intentions of the transgressor as predicted by the participant were tested by

conducting separate Univariate ANOVAs for each stereotype variable. Significant effects were found for gender stereotype flexibility only.

It was expected that those participants with high flexibility with gender stereotypes would indicate less negative intent. This hypothesis was tested with a Univariate ANOVA for tolerance (gender stereotype tolerance: low, high) on the intent judgment for the gender inconsistent scenario, F (1, 122) = 4.03, p = .047, η_p^2 = .03, (M = 1.62, SD = .82; M = 1.96, SD = 1.05 for participants with low stereotype flexibility and high stereotype flexibility respectively), and indicated that participants with high stereotype flexibility did indicate less negative intent than those participants with low stereotype flexibility. Another Univariate ANOVA for tolerance (gender stereotype tolerance: low, high) on the intent judgment for the gender consistent scenario was conducted as well, however no significant effects were found.

Differences between gender consistent and gender inconsistent scenarios. A repeated measures ANOVA for scenario (scenario: gender inconsistent, gender consistent) was conducted for the intentionality judgment in order to examine if, as expected, participants would demonstrate less negative intent in the gender inconsistent scenario than in the gender consistent scenario. Contrary to the original hypothesis, F (1, 123) = 4.19, p = .043, η_p^2 = .03, (M = 1.76, SD =.93; M = 1.96, SD = 1.03 for participants on the gender inconsistent scenario and the gender consistent scenario respectively), it was demonstrated that participants found the action in the gender inconsistent scenario as more negative than the action in the gender consistent scenario (see Figure 3).

Ingroup and outgroup differences. It was expected that the male participants would find the action to be more negative for a transgressor in his outgroup (female transgressor) than for a transgressor in his ingroup (male transgressor). In order to test this likelihood, a Univariate ANOVA for scenario (scenario: male transgressor (ingroup), female transgressor (outgroup)) was conducted for the intent judgment for male participants, but no significant effects were found. A Univariate ANOVA for scenario (scenario: male transgressor (outgroup), female transgressor (ingroup) was conducted on the intent judgment for the female participants as well, F (1, 62) = 4.96, p = .030, η_p^2 = .07, (M = 1.94, SD = .83; M = 1.45, SD = .89 for participants examining their outgroup or their ingroup, respectively), and revealed, contrary to predictions, that females indicated less negative intentions for their outgroup (male transgressor) than for their ingroup (female transgressor).

Punishment of the Transgressor Decision

Overall results. In order to discern if the participants, across the gender inconsistent and gender consistent scenarios, indicated the action should be punished or not (0, no punishment; 1, a little punishment; 3, a lot of punishment), frequency data were run for the combined answers to the punishment acceptability assessment. The frequency data revealed 76.7% of the participants indicated the action should be punished either a little, or a lot. No significant interaction effects were found, and therefore the individual Univariate ANOVAs are reported.

Age related main effects. Age differences for punishment judgments were analyzed by conducting separate Univariate ANOVAs for age (age: 3–4, 4–5, 6–7 years) on the punishment judgment for the gender inconsistent and gender consistent

scenarios. The Univariate ANOVA for the gender inconsistent scenario, F (2, 124) = 18.37, p = .000, η_p^2 = .23, (M = 1.52, SD = .67; M = 1.68, SD = .66; M = .77, SD = .84 for ages 3-4, 5-6, 7-8 years respectively), confirmed the hypothesis that with age, participants would endorse less punishment. Furthermore, the Univariate ANOVA for the gender consistent story, F (2, 121) = 12.56, p = .000, η_p^2 = .17, (M = 1.3, SD = .77; M = 1.46, SD = .72; M = .69, SD = .72 for ages 3-4, 5-6, 7-8 years respectively), additionally demonstrated that with age, participants endorsed less punishment (see Figure 4).

False belief theory of mind main effects. In concordance with the hypotheses for the two previous intentionality assessments, it was hypothesized that children who pass the false belief task would endorse less punishment than those children who did not past the false belief task. A Univariate ANOVA for false belief (false belief; fail, pass) on the punishment judgment was conducted for the gender inconsistent scenario, F (1, 116) = 10.12, p = .002, η_p^2 = .08, (M = 1.54, SD = .70; M = 1.07, SD = .90 for no false belief and false belief respectively), and showed that, as expected, those participants who passed the false belief task indicated less negative intentions than those who did not pass. The Univariate ANOVA for false belief (false belief: fail, pass) on the punishment judgment in the gender consistent scenario, F (1, 115) = 7.70, p = .006, η_p^2 = .06, (M = 1.33, SD = .75; M = .93, SD = .82 for no false belief and false belief respectively), revealed concordantly that participants who passed the false belief task were less likely to endorse punishment than the participants who had not passed the false belief task (see Figure 5).

Gender stereotype main effects. The hypotheses for the effects of the gender stereotype variables on the punishment judgment were in accord with the hypotheses for the two intentionality judgments. Participants indicating high tolerance and flexibility with gender stereotypes were expected to endorse less punishment than those with low tolerance and flexibility, and alternately that those participants with high knowledge of gender stereotypes were expected to endorse more punishment than those with low knowledge.

The Univariate ANOVA for gender stereotype tolerance (tolerance: low, high) on the punishment judgment in the gender inconsistent scenario, F (1, 122) = 5.37, p = .022, η_p^2 = .04, (M = 1.49, SD = .74; M = 1.14, SD = .90 for participants with low stereotype tolerance and high stereotype tolerance respectively), revealed, as predicted, that those participants with high tolerance of counter-stereotypic play endorsed less punishment than those participants with low tolerance. This result was further demonstrated by the Univariate ANOVA for tolerance (tolerance: low, high) on the punishment judgment in the gender consistent scenario, F (1, 121) = 6.05, p = .015, η_p^2 = .05, (M = 1.31, SD = .78; M = .96, SD = .79 for participants with low stereotype tolerance and high stereotype tolerance respectively), indicating again that those participants with high tolerance of gender stereotypes were less likely to endorse punishment than those participants with low tolerance of gender stereotypes.

Results revealed for gender stereotype flexibility mirrored those found for tolerance. A Univariate ANOVA for flexibility (flexibility; low, high) on the punishment assessment in the gender inconsistent scenario, F (1, 122) = 4.51, p = .036, η_p^2 = .04, (M = 1.46, SD = .80; M = 1.14, SD = .86 for participants with low

stereotype flexibility and high stereotype flexibility respectively), as well as Univariate ANOVA for flexibility (flexibility; low, high) on the punishment assessment for the gender consistent scenario, F (1, 121) = 6.32, p = .013, η_p^2 = .05, (M = 1.30, SD =.79; M = .94, SD = .77, indicated that those participants with high flexibility with gender stereotypes were less likely to endorse punishment than those participants with low flexibility with gender stereotypes.

The Univariate ANOVA for gender stereotype knowledge (knowledge: low, high) on the punishment judgment for the gender inconsistent scenario, F (1, 122) = 9.97, p = .002, η_p^2 = .08, (M = 1.49, SD = .77; M = 1.00, SD = .88 for participants with low stereotype knowledge and high stereotype knowledge respectively), as well as the Univariate ANOVA for knowledge (knowledge: low, high) on the punishment judgment for the gender consistent scenario, F (1, 121) = 10.87, p = .001, η_p^2 = .08, (M = 1.31, SD = .76; M = .82, SD = .78, both, contrary to hypothesis, but in conformity to what was found with the previous two intention judgments, indicated that participants with high stereotype knowledge were less likely to endorse punishment than those participants with low stereotype knowledge.

Differences between gender consistent and gender inconsistent scenarios. The hypothesis that participants would endorse more punishment in the gender consistent scenario than in the gender inconsistent scenario was examined with a repeated measures ANOVA for scenario (scenario: gender inconsistent, gender consistent) on the punishment judgment, F (1, 124) = 6.85, p = .010, η_p^2 = .10, (M = 1.32, SD =.83; M = 1.15, SD = .80 for participants on the gender inconsistent scenario and the gender consistent scenario respectively), and indicated, contrary to

hypotheses, but in accord with what was found with the initial intentionality judgment, that participants endorsed more punishment for the gender inconsistent scenario than for the gender consistent scenario (see Figure 6).

Ingroup and outgroup differences. A repeated measures ANOVA for scenario (scenario: male transgressor, female transgressor) on the punishment judgment was conducted separately for male participants and female participants, in order to test the hypothesis that participants would endorse less punishment for ingroup transgressors than for outgroup transgressors. No significant effects were found.

Justifications for the Judgment of Transgressor's Intentions

Justifications were provided by the participants for the two intentionality assessments as well as for the punishment acceptability assessment for both the gender inconsistent as well as the gender consistent scenarios, for a total of 6 assessments. Additionally, there are a total of 8 justifications (Moral: psychological harm, negligence, ownership, prosocial, accident; Social Conventional: deference to the rules, gender stereotypes; Personal: selfish desires) (for a full description of the justification categories see Table 3), each a dichotomous variable (0, not utilized; 1, utilized). No analyses were conducted with justifications that were utilized by fewer than 10% of the participants. This eliminated the moral justifications of psychological harm, and prosocial, as well as the social conventional justification of deference to the rules, and the personal justification of selfish desires. Each of the remaining justifications was examined separately for each assessment in each scenario in terms of its relationship to the independent variables (age, false belief,

stereotype tolerance, stereotype flexibility, stereotype knowledge) the differences between ingroup and outgroup populations and between the inconsistent and consistent scenarios (see Tables 8 and 9 for overall proportions of coding categories for each the intentionality and punishment acceptability variables). These analyses, in concert with the hypotheses, revealed findings overwhelmingly similar for the two intention judgments as well as the punishment judgment. In order to limit the redundancy of the findings, only the results for the first intentionality judgment will be reported here. Furthermore, no significant interaction effects were found, and therefore the individual Univariate ANOVAs are reported.

Age related main effects. It was hypothesized that participants would use more moral reasoning indicating no negative intentions with age. The moral justifications indicating no negative intentions were negligence (the action was due to the transgressor neglecting to look carefully at the initials on the backpacks), and accident (the action was an accident on the part of the transgressor). A Univariate ANOVA was conducted for age (age: 3- 4, 5 -6, 7 – 8 years) on the justification of negligence for the gender inconsistent scenario, F (1, 120) = 3.23, p = .043, η_p^2 = .05, (M = .09, SD = .25; M = .22, SD = .39; M = .28, SD = .41 for 3 – 4 year olds, 5 – 6 year olds, and 7 – 8 year olds respectively), indicating that with age, participants were more likely to reason the action was due to negligence. A Univariate ANOVA was additionally conducted for (age: 3- 4, 5 -6, 7 – 8 years) on the justification of negligence for the gender consistent scenario, however no significant effects were found.

In furtherance of the same hypothesis, a Univariate ANOVA was conducted for (age: 3- 4, 5 -6, 7 – 8 years) on the justification of accident for the gender inconsistent scenario, F (1, 120) = 4.94, p = .009, η_p^2 = .08, (M = .07, SD = .23; M = .06, SD = .20; M = .23, SD = .38 for 3 – 4 year olds, 5 – 6 year olds, and 7 – 8 year olds respectively), indicating that with age, participants were more likely to reason the action was due to an accident. In concordance, a Univariate ANOVA as conducted for age (age: 3- 4, 5 -6, 7 – 8 years) on the justification of accident for the gender consistent scenario, F (1, 120) = 4.66, p = .011, η_p^2 = .07, (M = .06, SD = .22; M = .14, SD = .31; M = .27, SD = .41 for 3 – 4 year olds, 5 – 6 year olds, and 7 – 8 year olds respectively), again indicating that with age, participants were more likely to reason the action was due to an accident. Taken together, with age, participants justified the action as more likely to be due to moral reasoning indicating no negative intentions by reasoning it was due negligence or an accident.

False belief theory of mind main effects. To illustrate the expectation that participants who pass the false belief task would use more moral reasoning indicating no negative intentions a Univariate ANOVA was conducted for false belief (false belief: fail, pass) on the justification of negligence for the gender inconsistent scenario, F (1, 112) = 15.46, p = .000, η_p^2 = .12, (M = .11, SD = .29; M = .37, SD = .41 for those without false belief and those with false belief respectively), demonstrating that participants who passed the false belief task were more likely to reason the action was due to negligence. Additionally, a Univariate ANOVA was conducted for false belief (false belief: fail, pass) on the justification of negligence for the gender consistent scenario, F (1, 112) = 5.75, p = .018, η_p^2 = .05, (M = .13, SD =

.32; M = .30, SD = .41 for those without false belief and those with false belief respectively) confirming the hypothesis that participants who passed the false belief task were more likely to reason the action was due to negligence.

Furthermore, a Univariate ANOVA was conducted for false belief (false belief: fail, pass) on the justification of accident for the gender inconsistent scenario, F (1, 112) = 11.28, p = .001, η_p^2 = .09, (M = .03, SD = .13; M = .20, SD = .39 for those without false belief and those with false belief respectively), demonstrating again, that participants who passed the false belief task were more likely to indicate the action was due to an accident. Taken together, those participants who passed the false belief task were more likely to reason the action was a result of negligence or an accident, not negative intentions.

Gender stereotype main effects. In concert with the previous findings for the gender stereotype variables, it was expected that participants with high gender stereotype tolerance and flexibility, and low gender stereotype knowledge would indicate the action would be more likely due to moral reasoning indicating no negative intentions. These expectations were tested with a Univariate ANOVA for gender stereotype tolerance (tolerance: low, high) for the justification of negligence in the gender inconsistent scenario, as well as a Univariate ANOVA for gender stereotype tolerance (tolerance: low, high) for the justification of negligence in the gender consistent scenario, however no significant effects were found for either scenario.

In order to further examine this hypothesis, a Univariate ANOVA for gender stereotype tolerance (tolerance: low, high) for the justification of accident in the

gender inconsistent scenario, as well as a Univariate ANOVA for gender stereotype tolerance (tolerance: low, high) for the justification of accident in the gender consistent scenario, however again, no significant effects were found for either scenario.

The hypotheses regarding gender stereotype flexibility were also examined. A Univariate ANOVA for gender stereotype flexibility (flexibility: low, high) for the justification of negligence in the gender consistent scenario, as well as a Univariate ANOVA for gender stereotype flexibility (flexibility: low, high) in the gender consistent scenario, however again, no significant effects were found for either scenario. Additionally, a Univariate ANOVA for gender stereotype flexibility (flexibility: low, high) for the justification of accident in the gender consistent scenario, as well as a Univariate ANOVA for gender stereotype flexibility (flexibility: low, high) in the gender consistent scenario were conducted, however again, no significant effects were found for either scenario.

Finally, the effects of gender stereotype knowledge on the justifications were examined. A Univariate ANOVA for gender stereotype knowledge (knowledge: low, high) for the justification of negligence in the gender inconsistent scenario was conducted, F (1, 118) = 6.42, p = .013, η_p^2 = .05, (M = .17, SD = .06; M = .35, SD = .11 for participants with low stereotype knowledge and high stereotype knowledge respectively), demonstrated that participants with a high knowledge of gender stereotypes were more likely to reason the action was due to negligence than those participants with low knowledge of gender stereotypes. In addition, a Univariate ANOVA for gender stereotype knowledge (knowledge: low, high) for the

justification of negligence in the gender consistent scenario was conducted, F (1, 118) = 6.66, p = .011, η_p^2 = .05, (M = .14, SD = .08; M = .32, SD = .13 for participants with low stereotype knowledge and high stereotype knowledge respectively), and also indicated that those with a high knowledge of gender stereotypes were more likely to reason the action was due to negligence than those participants with a low knowledge of gender stereotypes.

Furthermore, this hypothesis was analyzed with respect to the justification of accident for gender stereotype knowledge. A Univariate ANOVA for gender stereotype knowledge (knowledge: low, high) for the justification of accident in the gender inconsistent scenario, F (1, 118) = 5.82, p = .017, η_p^2 = .05, (M = .08, SD = .09; M = .22, SD = .12 for participants with low stereotype knowledge and high stereotype knowledge respectively) signified that participants with high gender stereotype knowledge were more likely to reason the action was due to an accident than those with low stereotype knowledge. Finally, a Univariate ANOVA for gender stereotype knowledge (knowledge: low, high) for the justification of accident in the gender consistent scenario, F (1, 118) = 5.07, p = .026, η_p^2 = .04, (M = .11, SD = .04; M = .26, SD = .14 for participants with low stereotype knowledge and high stereotype knowledge respectively) revealed that participants with high gender stereotype knowledge were more likely to reason the action was due to an accident than those with low gender stereotype knowledge. These results taken together, although different than initial hypotheses, are consistent with the findings of intentionality and gender stereotype knowledge, and demonstrate that those participants with high

knowledge of gender stereotypes indicated the transgression to be less negative than those participants with a low knowledge of gender stereotypes.

Differences between gender consistent and gender inconsistent scenarios. In order to test the hypothesis that participants would be more likely to use moral reasoning indicating no negative intentions for the gender inconsistent scenario than for the gender consistent scenario, a repeated measures ANOVA for scenario (scenario: gender inconsistent, gender consistent) for the justification of negligence was conducted. No significant effects were found. In order to further examine this hypothesis, a repeated measures ANOVA for scenario (scenario: gender inconsistent, gender consistent) on the justification of accident was conducted. Again, no significant effects were found.

Ingroup and outgroup differences. Regarding the hypothesis that participants would be more likely to justify their judgments of intentionality with moral reasoning indicating no negative intentions in scenarios that depicted their ingroup rather than their outgroup as the potential transgressor, a repeated measures ANOVA for scenario (scenario: male transgressor; female transgressor) on the justification of negligence was conducted for male participants, and then again separately conducted for female participants. No significant effects were found. The hypothesis was further examined by conducting a repeated measures ANOVA for scenario (scenario: male transgressor; female transgressor) on the justification of accident for male participants, and then again separately for female participants. Again, no significant differences were found for the ingroup or the outgroup in their justifications of the intentionality judgment.

Embedded False Belief Task

The false belief task administered separate from the Intergroup Attitudes Attribution of Intentions Task and converted into an independent variable, is the standard false belief task used in the false belief theory of mind research. The embedded false belief task was created for the sole purpose of the current study. The embedded false belief task is a part of the Intergroup Attitudes Attributions of Intentions Task, and as such, is embedded in a morally relevant scenario, rich with contextual information such as the highly stereotyped toys. No significant interaction effects were found, and therefore the individual Univariate ANOVAs are reported.

Age related main effects. It was hypothesized that, similar to the independent false belief task, participants would be more likely to pass the embedded false belief task with age. This hypothesis was examined by conducting a Univariate ANOVA for age 3 (age: 3-4, 5-6, 7-8 years) on the embedded false belief task in the gender inconsistent scenario, F (2, 124) = 17.30, p = .000, η_p^2 = .22, (M = .023, SD = .15; M = .17, SD = .38; M = .49, SD = .51 for ages 3-4, 5-6, 7-8 years respectively), confirming hypotheses that participants were more likely to pass the false belief task with age. An additional Univariate ANOVA for age 3 (age: 3-4, 5-6, 7-8 years) on the embedded false belief task in the gender consistent scenario, F (2, 124) = 27.03, p = .000, η_p^2 = .30, (M = .023, SD = .15; M = .13, SD = .33; M = .56, SD = .50 for ages 3-4, 5-6, 7-8 years respectively), also confirmed hypotheses that participants were more likely to pass the embedded false belief task with age (see Table 7).

False belief theory of mind main effects. In order to test the hypothesis that participants would indicate a false belief theory of mind at an earlier age on the

independent measure of false belief than on the embedded task, a 3 (age: 3-4, 5-6, 7-8 years) X 2 (false belief: fail, pass) ANOVA was conducted on the embedded false belief theory of mind task for the gender inconsistent scenario and then again separately for the gender consistent scenario. While separate main effects for age were found for both the false belief task as well as the embedded false belief task, no interaction between the two was found for either the gender inconsistent or the gender consistent scenarios.

Despite this, the frequency data do signify that the participants were indicating a false belief theory of mind on the independent task at an earlier age than on the embedded false belief theory of mind assessment. For the 3 – 4 year olds, 10% indicate having an independently measured false belief theory of mind, while only 2.3% indicate having a false belief theory of mind in the embedded measure. Similarly, for the 5 – 6 year olds, 44% indicate having an independently measured false belief theory of mind, while 15% indicate having a false belief theory of mind in the embedded measure. Furthermore, for the 7 – 8 year olds, 88% indicate having an independently measured false belief theory of mind, while 52.3% indicate having a false belief theory of mind in the embedded measure. Additionally, overall, 48.3% of the participants indicate having an independently measured false belief theory of mind, while only 23.6% of the participants indicate having a false belief theory of mind in the embedded measure. This demonstrates that, as expected, participants were able to indicate a false belief theory of mind more easily and at an earlier age for the independently measured task than for the embedded task.

Chapter 5: Discussion

Introduction

The goal of this study was to investigate how intergroup attitudes, such as gender stereotypes, in combination with a false belief theory of mind competency, impact moral judgments. More specifically, the goal was to examine whether gender stereotypes impact children's attributions of intentionality in peer interactions when the intention is ambiguous. To measure the role of cognitive judgment on these judgments, children were also measured for their false belief theory of mind ability, and whether this ability contributed to judgments about attributions of intentions in contexts in which gender stereotypes were made salient. These issues are important to examine because when children's gender stereotypes are affecting how they attribute intentions then children may be assigning blame inappropriately in peer contexts. This over attribution of negative intentions can lead to peer rejection and exclusion (Crick & Dodge, 1996; Dodge & Newman, 1981). To date, no research has examined the impact of intergroup attitudes and false belief theory of mind competencies on moral judgments.

There were many novel findings from this study. Gender stereotypes and false belief theory of mind were shown to be related to children's attributions of intentionality, and children's gender stereotypes impacted their decisions about intentionality. Additionally, it was shown that false belief theory of mind facilitated the amelioration of the impact of those gender stereotypes. Furthermore, gender

stereotypes impacted children's ability to indicate a false belief theory of mind. Children found it difficult to make a false belief judgment when it was embedded in a task involving intentions that were ambiguous. Children were not able to pass the false belief theory of mind task until a significantly later age, with less than half passing at 7 -8-years-of-age, instead of the majority passing at 4 – 5-years-of-age (see Figure 7) (Wellman & Liu, 2004).

The impact of the gender inconsistent scenario in contrast to the gender consistent scenario did not emerge as predicted. The original hypothesis predicted the participants would indicate less negative attributions of intentions (indicating that the action was less wrong or more all right) in the gender inconsistent scenario in which a boy put a girl's doll into his backpack or a girl put a boy's truck into her backpack. This prediction was based on the assumption that participants would view the taking of the counter-stereotypic toy as accidental, assuming that the child would not want to actually steal or take home a counter-stereotypic toy. What was found instead was participants indicated more negative intentions (indicating the action was more wrong) in the gender inconsistent scenario (see Figure 3). This result has to be viewed in light of the data indicating that overall, the overwhelming majority of the participants found the action to be not all right, and indicated negative intentions. If the children were viewing this action negatively, they were not paying attention to the possible accidental placement of the toy into the wrong backpack. Instead, the children were focusing on the taking of the toy. When focusing on the taking of the toy, children were judging the taking of a counter-stereotypic toy, the one in the

gender inconsistent scenario, as more wrong than the taking of the gender consistent toy.

This result is not surprising as Karinol and Aida's (1997) examination of judgments associated with children breaking counter-stereotypic toys yielded similar results. Karinol and Aida (1997) found that when judging the breaking of a neutral toy, punishment was not recommended. However, the participants expressed that children should not want a toy that violates gender stereotypes, and subsequently judged those who broke gender inconsistent toys more severely (Karinol & Aida, 1997). This is consistent with what was found in the current study, as the participants judged the child who took the inconsistent toy more severely than the child who took the gender consistent toy.

Overall, the current study also contributes to social cognitive domain theory of moral development by illustrating the importance as well as the role of false belief theory of mind competency in making attributions of intentionality in general as well as in morally relevant intergroup scenarios. In concordance with the hypotheses, the participants with a false belief theory of mind competency were able to indicate less negative intentionality (indicated the action to be more all right) across both scenarios, regardless of gender stereotypes, and were thereby better able to see the ambiguity of the scenarios and withhold judgment. This finding indicates that having a false belief theory of mind is critical to being able to make sound moral decisions regarding intentionality, and can even ameliorate the impact of having gender stereotypes when making decisions about intentionality. In addition, when the false belief theory of mind task was embedded into the scenario that was not only morally

relevant, but included contextual variables that were gender stereotypic, children were not able to pass until a much later age. This indicates that the addition of the gender stereotype contextual information made it significantly more difficult for the children to be able to indicate a false belief. These novel findings, along with others will be discussed below in greater detail, and with respect to the other variables in this study.

False Belief Theory of Mind Task

In concordance with previous findings (Wellman & Liu, 2004) as well as hypotheses made, the findings revealed that with age, children acquire a false belief theory of mind as indicated by the separate and independent false belief theory of mind task. The overwhelming majority of participants from 3 – 4 years-of-age were unable to pass the false belief theory of mind task, while, as expected (Wellman, Cross, & Watson, 2001), by 5 – 6 years of age, many of the participants were able to pass (see Table 7). Finally, the vast majority of the participants from 7 – 8-years-of-age, also as expected, were able to pass the independent false belief theory of mind task. These findings serve to confirm previous findings as well as validate the age range chosen for the current study.

Gender Stereotype Task

In the current study, three gender stereotype scales were assessed. For the gender stereotype flexibility scale, high gender stereotype flexibility indicated the participant was flexible in liking any child playing with any toy regardless of the gender stereotype, and low flexibility indicated the participant was not flexible in liking a child playing with a counter-stereotypic toy. High gender stereotype tolerance

indicated the participant was tolerant of any child playing with any toy, regardless of the gender stereotype, while a low gender stereotype tolerance indicated the participant was intolerant of children playing with counter-stereotypic toys. Finally, high gender stereotype knowledge indicated the participant could accurately predict which toys were gender stereotyped for boys and which were stereotyped for girls, and low gender stereotype knowledge indicated the participant was not able to accurately specify the gender stereotyped toys.

It was predicted, and found, that children with high gender stereotype flexibility and tolerance would indicate less negative attribution of intentions, as these children would be able to see past the stereotypic nature of the scenario to better see the ambiguity and withhold judgment. It was also predicted that children with a high knowledge of gender stereotypes would indicate more negative attributions of intentions, as it was expected that these children with a high knowledge would also endorse stereotypes. This was not the case. Children with high knowledge of gender stereotypes were in line with the children with high tolerance and flexibility of gender stereotypes. They could accurately specify the gender stereotype of toys, but did not necessarily endorse the stereotype of only that gender of child playing with the gender stereotyped toys.

The inclusion of these gender stereotype scales uniquely allowed for the examination of the direct impact of gender stereotypes on attributions of intentionality, punishment decisions, the social reasoning supporting those decisions, as well as the impact of having a false belief theory of mind. Previously in the literature, it has been found that stereotype knowledge, paired with stereotype

rigidity, increases with age up until a ceiling is reached by 5- or 6-years-of-age (Ruble & Martin, 1998; Trautner, Ruble, Kirsten, & Hartmann, 2005). In the current study, in accord with the previous literature, it was found that gender stereotype knowledge increased with age. For the standard measure of false belief competency, most studies demonstrate that this ability is achieved by 5- or 6-years-of-age. Yet, in this study, as was shown in a previous study on morally-relevant theory of mind, children as old as 8 years of age did not apply their false belief knowledge to the gender stereotype context. In addition, children's tolerance for gender stereotypes increased with age, again with no ceiling emerging.

False belief theory of mind competency is a cognitive competency that indicates children can accurately view or predict what another person's perspective would be, and predict that person's actions even when the other person's perspective or beliefs differ from reality and the child's own beliefs (Wimmer & Perner, 1983). This is a complex cognitive capacity which facilitates social development as it allows children to see social situations from another person's perspective and acknowledge that the other person may be making decisions based on inaccurate information. Due to the cognitive complexity of the task, the current study predicted that children who were able to accomplish this and pass a false belief theory of mind task would also be more likely to have greater flexibility with gender stereotypes as well as greater tolerance with gender stereotypes. This hypothesis was partially confirmed as a greater tolerance of gender stereotypes was found with the indication of a false belief theory of mind competency. The findings for gender stereotype flexibility were however not significant and therefore inconclusive. It is possible that because it has

been found that rigidity of gender stereotypes increases with age (Turiel & Stoddard, 1985), as does the acquisition of a false belief theory of mind (Wimmer & Perner, 1983), the impact of the cognitive phenomenon are cancelling each other out for the age range that has been explored.

The acquisition of a false belief theory of mind was also seen to coincide with the acquisition of greater knowledge of gender stereotypes, contrary to the initial hypothesis. The original hypothesis was that gender stereotype knowledge would coincide with gender stereotype use and endorsement (Trautner at al., 2005), and therefore be correlated with a low gender stereotype tolerance as well as low gender stereotype flexibility, but as previously mentioned, this was not the case, high knowledge correlated with high tolerance and flexibility. It was therefore logical that as children passed the false belief theory of mind task, they would additionally be more likely to have a greater understanding of gender stereotypes, and therefore also have a high knowledge of gender stereotypes along with a high tolerance of gender stereotypes.

The gender stereotype scale additionally revealed a significant finding for sex of participant such that the female participants were revealed to have more knowledge of gender stereotypes than were the male participants. This is in concordance with previous findings that male toys are seen overall as more desirable and of a higher status than female toys (Antill et al., 1996). Therefore males and females would be more likely to desire to play with male stereotyped toys, and the females would be more encounter more discouragement from playing with toys that are considered counter-stereotypic, and additionally would be able to predict what would be seen as

socially acceptable for them to play with and what is not as children seen playing with opposite-sex toys is rated by other children as relatively bad (Smetana, 1986).

Attribution of Intention and Punishment Acceptability Judgments

Differences between Gender Consistent and Gender Inconsistent Scenarios

As briefly outlined above, a significant difference between the judgments made in the gender inconsistent scenario (a boy puts a doll into his backpack) and the gender consistent scenario (a girl puts a truck into her backpack) was found not only for the attribution of intention assessment, but additionally for the punishment acceptability assessment as well (see Figures 3 and 6). The two scenarios were identical except for the use of a gender stereotypic inconsistent versus gender consistent toy in the scenarios. These findings show for the first time that children are using stereotypic information to make attributions of intentionality, over attribute negative intentions to children engaging in counter-stereotypic behavior, and make more severe punishment acceptability decisions for the children engaging in counter-stereotypic behavior based on this over attribution of negative intentionality.

While these findings are novel, they are supported by the literature. Gender-inconsistent behavior has been shown to be deemed unacceptable by the majority of children because of the reactions of their peers to this behavior (Carter & Patterson, 1982; Smetana, 1986; Stoddart & Turiel, 1985). Stoddard and Turiel (1985) discovered a U-shaped curve as the youngest (5-years-of-age) and oldest children (13-years-of-age) in their study found participation in gender-inconsistent activities more wrong than did children in middle childhood. Additionally they discovered that children in the youngest age group as well as children in the oldest age group thought

that participation in a gender-atypical activity was wrong. The authors concluded that in kindergarten the maintenance of gender identity is defined in physical terms, so if a girl was to play a male-stereotypical game, other children might question her gender, which would be highly undesirable for a young child (Stoddart & Turiel, 1985). The findings from the current study reinforced this as the participants of all three age groups, from age 3 – 8-years-of-age found the action in the gender counter-stereotypic scenario as more wrong (more negative intention) than the gender consistent scenario, but also added to this research as the participants did not just evaluate the counter-stereotypic behavior as negative, they evaluated the intentions of the actor as negative simply because he/she was engaging in counter-stereotypic behavior.

Additionally, Killen and Stangor (2001) found that children viewed straightforward exclusion, based on gender alone, from a stereotype inconsistent activity (boys from a ballet club, girls from a baseball club) as wrong from a moral viewpoint. However when the children were asked to make a choice between a stereotypic child (presented as more qualified) and a counter-stereotypic child, the stereotypic child was chosen, and group functioning was referenced for why. This indicates that the participants viewed the gender-inconsistent activity negatively (Killen & Stangor, 2001). This was again confirmed in the current study, and the results furthered as the participants did not just exclude the counter-stereotypic child from an activity, they judged that child to be doing something more wrong than a child who engaged in the exact same activity but was stereotype consistent. The

participants' attributions of intentions were highly influenced by the stereotypic nature of the toy, which is an over attribution of negative intentions.

The current study additionally serves to further Horn, Killen, and Stangor's (1999) research on adolescents. Horn, Killen, and Stangor (1999) examined ambiguous scenarios with adolescents in a high school setting (scenario 1, someone broke the sound equipment at a party; scenario 2, someone broke into the school's computer system). In this study, the participants used stereotypic expectations to assign blame. The authors revealed participants to be more likely to accuse the "jocks" than the "techies" of breaking the sound equipment at the party, and were more likely to accuse the "techies" than the "jocks" of breaking into the computer system (Horn, Killen, & Stangor, 1999). Even though both scenarios were ambiguous, as are the ones in the current study, the participants used stereotypic expectations to assign blame. The current study though, was able to establish a similar finding, but with a much younger population, showing that children are using gender stereotypes to assign blame from as early as 3 years of age.

These studies, taken together, revealed that children and adolescents used stereotypes to make exclusion decisions as well as to attribute intentions. Not only did this over attribution of negative intentions indicate that these participants were allowing gender stereotypes and bias to influence their decisions, these decisions about intentionality were leading to prejudicial behavior about the acceptability of punishment of this action. Importantly, over attribution of negative intentions has been linked to peer rejection and exclusion (Crick & Dodge, 1996; Dodge & Newman, 1981). These findings are particularly significant, as children and

adolescents have been shown to have stereotypes that become more entrenched with age (Stangor & Schaller, 1996), and the current study demonstrated that children are making decisions about not just intentionality, but also punishment using gender stereotypes as early as 3-years-of-age. It is therefore critical to continue this line of research so as to understand as much about children's over attribution of negative intentions based on stereotypes as possible, as well as to understand when and how it begins, and to recognize and discern what other decisions children could be making that are influenced by these stereotypes and others.

Impact of Gender Stereotypes on Intent and Punishment Judgments

In line with the original prediction that children would view the counter-stereotypic scenario with less negative intent than the stereotype consistent scenario, it was also originally hypothesized that participants with a low tolerance and flexibility for gender counter-stereotypic play would indicate more negative intent for protagonists in the stereotype consistent scenario. As the original prediction was not realized, and rather the opposite occurred, likely due to the overall negative evaluation of the actions in both scenarios, the assumption concerning the impact of gender stereotypes was not realized either.

The findings for gender stereotype flexibility and tolerance are in accord with the newly informed interpretation of the data. The participants with high tolerance and flexibility with gender stereotypes indicated less negative intent and endorsed less punishment for both the gender inconsistent as well as the gender consistent scenario, with no differences between the two scenarios. The lack of differentiation between the two scenarios is possibly due to the stereotypic nature of both scenarios,

with those children with a high tolerance and flexibility better able to see the ambiguity in both scenarios, not hindered by the stereotypic nature of the toy, thereby able to indicate less negative intent and endorse less punishment for both.

No hypotheses were made with regard to high knowledge of stereotypes, as originally it was hypothesized that high knowledge of stereotypes would indicate a high use and acceptance of gender stereotypes. As this scale of gender stereotype knowledge did not indicate use or endorsement of gender stereotypes, but rather an understanding of how typical gender stereotypes are applied, it is not surprising that in harmony with the previous finding for high tolerance and high flexibility, those participants with a high knowledge of gender stereotypes also indicated less negative intent and endorsed less punishment for both the gender inconsistent as well as the gender consistent scenario, with no differences between the two scenarios. Again, this is possibly because those participants with a high knowledge of gender stereotypes were better able to see the ambiguity in both scenarios, thereby able to indicate less negative intent and endorse less punishment for both.

Ingroup and Outgroup Differences

It was hypothesized that participants would find the action to be more all right (indicate less negative intentions), and endorse less punishment for ingroup transgressors than for outgroup transgressors in concert with the rich literature indicating a strong ingroup favoritism (Boldizar & Messick, 1988; Brewer & Silver, 1978; McAllister, 1995; Platow, et al., 1990). However, this was not found. The first judgment of intentionality as well as the punishment decision yeilded no significant effects, however the attribution of intention by the participant assessment did indicate

the female participants found the action to be more all right for the male (outgroup) potential transgressor than for the female (ingroup), indicating more negative intentions, not less, for the ingroup. As this finding was not revealed for the first attribution of intention assessment or the punishment judgment, it could be that this finding is an anomaly. It is also possible that this finding is in accord with recent research that indicates ingroup favoritism depends on group status (Jost, Banaji, & Nosek, 2004), as well as the research that demonstrates that some females implicitly endorse certain male stereotypes (Rudman & Kilianski, 2000) while preferring higher status male items and toys (Antill, Goodnow, Russell, & Cotton, 1996). In that case, the females would understand why a girl would want a truck, and would therefore believe that a girl took the truck on purpose because it was desirable. A measure with the more explicit goal of a comparison of ingroup and outgroup attributions of intentions is needed in order to further extricate and disentangle this finding.

Impact of False Belief Theory of Mind

The current study verified the hypotheses made for false belief theory of mind, as it was hypothesized that those participants who indicated having a false belief theory of mind would be more likely to see the accidental nature of the transgression in either the gender inconsistent or gender consistent scenarios and thereby indicate less negative intention for both intention judgments, and also endorse less punishment for the action. In accord with these hypotheses, the participants in the current study passed the false belief task did indicate for both attribution of intention judgments, less negative intentions, as well as less endorsement of punishment. The

participants with a false belief were better able to see the accidental nature of the scenarios than were those participants without a false belief.

This hypothesis and the concurrent findings are in accord with the vast literature on false belief theory of mind that indicates that a child able to pass a false belief task is able to understand that another individual can have beliefs that are not only different from their own, but different from reality, and subsequently able to predict decisions made by an individual based on those false beliefs (Wellman & Liu, 2004). This is a very complex cognitive achievement, and Wellman and Liu (2004) have also found that before a child can understand false beliefs, that child understand that individuals are able to have differing desires from their own, and can also predict what choices and decisions those individuals would make based on those differing desires. This would indicate that children who have a false belief theory of mind competency additionally have the ability to see that someone could want a toy that they themselves would not desire, and therefore would not be hindered in their judgments by the stereotypic nature of either of the scenarios.

Additionally, recent research by Killen and colleagues (in press) directly speaks to the current findings as they revealed that children without a false belief theory of mind viewed it as more acceptable to punish an accidental transgressor than did participants with a false belief theory of mind (Killen, Mulvey, et al., in press). The current findings not only support what was already known about false belief theory of mind, but are also able to move the fields of false belief theory of mind and moral development forward. The current study was able to determine that having a false belief theory of mind not only facilitates less negative attributions of

intentionality in an ambiguous scenario, but also in a scenario steeped in gender stereotypes. These findings not only speak to the critical nature of this cognitive capacity in general, but also speak to its importance in moral decision making, and additionally its impact on the maintenance and influence of gender stereotypes on the decision making process. This is accomplished as the current study showed that children with a false belief theory of mind were able to see past the stereotypic nature of the scenarios in order to indicate less negative intentionality, thus ameliorating the negative impact of the gender stereotypes on the attribution of intentionality.

Age Related Findings

The current study demonstrated age related findings regarding the attribution of intentionality (see Figure 2). It was hypothesized that with age, participants would indicate less negative intentions and endorse less punishment, as with age, the participants would be better able to see the accidental nature of the transgression, and in fact, that hypothesis was confirmed. For the first attribution of intention assessment as well as the punishment judgment, with age, children indicated less negative intention as well as less endorsement of punishment for both the gender inconsistent as well as the gender consistent scenarios.

These findings are highly supported by what has been previously found in the social domain literature as Killen and colleagues were able to show that children and younger adolescents were more likely to make accusations based on stereotypic assumptions with less available information than older adolescents, who reserve judgment until more evidence is presented (Horn, Killen, & Stangor, 1999; Killen, Kelly, Richardson, & Jampol, in press). Additionally, Killen and colleagues found

that the younger preschoolers, 4.5-years-of-age, were more likely than the older preschoolers, 5.5-years-of-age, to base initial judgments on stereotypic knowledge (Killen et al., 2001).

Social Reasoning about Attribution of Intentions and Punishment Acceptability Judgments

A unique aspect of the research of the social domain model is the focus on social reasoning data which elucidates individuals' evaluations of social interactions. The current study was able to support the findings already made in this study and others as well as move the field of moral development and false belief theory of mind forward with these social reasoning data.

Age Related Findings

The current study's age related findings that children were less likely to attribute negative intentions and less likely to endorse punishment with age were supported by the social reasoning data. Overall, the participants used more moral reasoning indicating the action was due to an accident or the negligence of the actor to pay attention to the initials on the backpacks which indicated ownership, both moral reasons indicating a lack of negative intentions. It was hypothesized though that participants would use more reasoning, indicating deference to the rules (indicating no negative intentions) with age for the attribution of intention judgments as well as for the punishment judgment, but very few participants reasoned about the action using deference to the rules. The inclusive hypothesis however, indicating individuals would use more reasoning that indicates lack of negative intention, was

revealed for both the attribution of intention judgments as well as the punishment judgment across both the inconsistent and consistent scenarios.

These findings are supported by what has been found in the social domain literature to date, as Killen and colleagues have found younger children to use more reasoning intended on blaming a potential transgressor in ambiguous scenarios than older children, while the older children were more likely to refer to a lack of negative intentions to explain their attribution of intentions (Horn, Killen, & Stangor, 1999; Killen, Kelly, Richardson, & Jampol, in press; Killen, Pisacane, Lee-Kim, and Ardila-Rey 2001).

False Belief Theory of Mind

The study of false belief theory of mind very rarely is inclusive of social reasoning data (for exceptions see: Killen, Mulvey, et al., in press). The social reasoning data in the current study served to support the false belief theory of mind findings that those participants with a false belief were able to indicate less negative intentionality as well as indicate less punishment than those participants with no false belief. The participants with a false belief used more moral reasoning indicating the action was an accident, and not intended at all, as well as reasoning that the action was due to the negligence of the actor to look closely at the bags, also indicating no negative intentions, than those participants that did not indicate a false belief theory of mind.

The hypothesis that participants with a false belief would use more reasoning indicating a lack of negative intentions than those participants without a false belief was additionally supported by the recent study of morality and theory of mind by

Killen and colleagues which found that children without a false belief were more likely to attribute negative intentions to an accidental transgressor the children with a false belief (Killen, Kelly, Richardson, & Jampol, in press). This study however, did not include a gender stereotypic context as the current study did.

Differences between Gender Inconsistent and Gender Consistent Scenarios

Very few participants explicitly noted a gender stereotype in their justification of their judgments of intention or punishment. Therefore, no differences were found between the gender inconsistent and gender consistent scenario as far as reasoning about gender stereotypes. Additionally, no differences were found at all between the reasoning used for the gender inconsistent and gender consistent scenarios. It is possible that no differences were found due to the general nature of the justification asked for. Therefore, had the justification not simply asked why it was all right or not all right for the actor to have put the toy into the wrong backpack, but instead included more probing questions, it is possible more distinctions between the two scenarios could have been made.

Gender Stereotype Findings

The findings from the social reasoning data again served to support the earlier findings that those individuals with high tolerance, high flexibility, and high knowledge of gender stereotypes made less negative attributions of intentionality and endorsed less punishment than those with low flexibility, tolerance, and knowledge. The participants with high tolerance, flexibility, and knowledge were more likely to use moral reasoning indicating no negative intentions than were those with low

tolerance, flexibility, and knowledge. More specifically, participants who indicated a high tolerance and knowledge of and flexibility with gender stereotypes reasoned more often than those low in gender stereotype knowledge, tolerance, and flexibility that the action was without negative intentions, indicating that the action was likely an accident, or that the action was due to the negligence of the actor to look closely at the initials on the bags, both of which indicate a lack of negative intentions.

It was originally hypothesized that participants with were low in gender stereotype flexibility and tolerance would use more reasoning using gender stereotypes than participants that indicated a high tolerance of and flexibility with gender stereotypes, but again, as very few participants used reasoning indicating gender stereotypes, this hypothesis was not corroborated by the data.

These findings do much to support the attribution of intentionality findings. The children who were highly knowledgeable, tolerant of, and flexible with gender stereotypes were not only indicating less negative intentions for the judgments, they were more likely to reason that the action was due to either an accident, or the negligence of the transgressor to pay attention to which bag the toy was going into. This supports the contention that these participants who are knowledgeable, tolerant and flexible with gender stereotypes are clearly better able to see the accidental nature of the transgression, and are not hindered by the stereotypical or counter-stereotypical nature of the toy that is being taken, and are not over attributing negative intentions to the transgressor. The over attribution of negative intentions due to gender stereotype inflexibility and intolerance can have serious consequences in social relationships in the future, as these children can be more likely to be excluded and rejected (Crick &

Dodge, 1996; Dodge & Newman, 1981) and more likely to hold stereotypes and prejudices as they age as stereotypes become more entrenched with age (Stangor & Schaller, 1996).

Embedded False Belief Theory of Mind Task

The embedded false belief theory of mind task was of great importance to the current study, as it was able to further the very limited research connecting theory of mind to moral development. This was accomplished by replicating the findings from Killen and colleagues study on morality and theory of mind (Killen, Mulvey, et al., in press), which indicated that children without false belief competency were more likely to attribute negative intentions and found it more acceptable to punish children in the morally relevant hypothetical scenario than children with false belief. Additionally, it was found that children attributed negative emotions and intentions to the accidental transgressor up until 8 – years-of-age and made it clear for the first time that integrating moral judgment and false belief theory of mind poses challenges to children (see Table 7 and Figure 7) (Killen, Mulvey, et al., in press).

The study by Killen and colleagues revealed a portion of what is impacting a child's ability to make an attribution of intentionality. It is clear that while a 4- or 5-year-old can correctly distinguish intentionality in a straightforward scenario involving no emotional valence or moral concerns, the addition of a moral premise adds a layer of complexity to the scenario that makes it more difficult to correctly decipher intentionality. What had not yet been explored was if adding more contextual features, such as a highly gender stereotyped toy, instead of an object which elicits no emotional reaction, to a scenario aimed at examining a child's ability

to attribute intentionality would allow for a more thorough investigation of the impact of those contextual features.

The current study revealed that adding the layer of complexity and contextual information of the gender stereotyped toys influenced the participants' ability to discern intentionality. In the embedded false belief task, the participants were more likely to pass the task with age. The ages that the participants were able to indicate a false belief theory of mind were much later as a whole than the age at which they were able to indicate a false belief in the independent task, which included no contextual or social information unlike situations that are likely to arise on an everyday basis where the objects and people involved are identifiable and known. Furthermore, it was revealed that even by the oldest age range, 7 -8-years-of-age, only 49% of the participants were able to pass the embedded false belief theory of mind task. This is not only much later than the 4 – 5-year-old age range that is seen in the literature for false belief (Wellman & Liu, 2004), but it is additionally later than what was found in independent false belief theory of mind task in this measure, and what was found for the embedded task for Killen and colleagues morality and theory of mind study (Killen, Mulvey, et al., in press). The morality and theory of mind study was able to indicate that most of the participants had indicated a false belief theory of mind on the embedded task by the age of 8-years-of-age (Killen, Mulvey, et al., in press). In our task, which included not only contextual and social information in a morally relevant scenario, but additionally included gender stereotyped information, only 49% of the 7 – 8-year-old participants were able to indicate a false belief in the embedded task. This further supports the claim that the addition of the

gender stereotyped information is making it more difficult for children to discern intentionality.

The fundamental determinations of intentionality, a core aspect of moral and social reasoning, is affected not only by the introduction of a moral premise to a straightforward false belief theory of mind scenario, but also strongly affected by the introduction of gender stereotyped contextual information. These new data show that children are taking the contextual information into account when making attributions of intentionality and punishment decisions, and additionally these data inform us that determinations of intentionality are context dependent.

Conclusion

The current study had many unique aspects with the inclusion of an embedded false belief theory of mind task into a scenario of moral relevance with a gender stereotyped context, as well as the inclusion of measures of gender stereotype tolerance, flexibility, and knowledge. There are therefore many novel finding that are able to push the fields of moral development and false belief theory of mind forward. First, the finding of a difference between the gender inconsistent and gender consistent scenarios reveals for the first time that children as young as three years of age are having their decisions concerning attribution of intentionality, a core aspect of moral development research, impacted by gender stereotypes and the inclusion of gender stereotypical objects. Furthermore, it was able to be seen that children's gender stereotypes, as directly measured with the tolerance, flexibility, and knowledge variables, impact children's decisions and reasoning about intentionality

and punishment in scenarios involving gender stereotyped objects. It was clear that those children with gender stereotypes were more likely to over attribute negative intentions than those without gender stereotypes.

These findings are critically important as gender stereotypes tend to become more entrenched with age (Stangor & Schaller, 1996), and they impact critical social decisions such as the attribution of intentionality, which when negatively over attributed can lead to peer exclusion and rejection (Crick & Dodge, 1996; Dodge & Newman, 1981). Children that perceive negative intentions tend to react to those perceived negative intentions with a concordant negative reaction and possible aggression, and so therefore children who are over attributing negative intentions perceive and react to negativity even when it is not there. This can lead to peer rejection and exclusion because of the child's avoidable negative reactions and possible aggressive behavior (Crick & Dodge, 1996; Dodge & Newman, 1981). These findings emphasize the importance of exploring this phenomenon even further with respect to age as well as additional contextual variables which could be impacting these decisions.

The current study was able to discern that having a false belief theory of mind did help to ameliorate the negative impact of the inclusion of gender stereotyped objects into the scenario, by allowing the children to see past the gender stereotyped objects and see the ambiguity of the scenario and indicate less negative intentions. This cognitive capacity could therefore hold promise in the field of moral development as it is possible that by encouraging the development of a false belief

theory of mind could facilitate children's ability to see past their stereotypes when making social decisions.

The other major finding to come out of the current study directly speaks to the field of false belief theory of mind, as it was found that embedding a false belief task into a scenario that is both morally relevant as well as contextually rich with gender stereotyped objects significantly impacts the age at which a child can pass the task. The task becomes more complex because of the abundance of contextual information, but everyday scenarios are inclusive of the same level of contextual information, and adding that into the false belief theory of mind task therefore represents a more realistic view of the amount and kind of information that a child is sorting through in order to make decisions about another's false belief, and beliefs in general.

Limitations

The current study has done a great deal to add to the relevant fields of study, but also has several limitations that should be noted, and possibly addressed in a follow-up study. The gender inconsistent scenarios and gender consistent scenarios did differ in the inclusion of a cross-gender interaction for the gender inconsistent scenario. While it is unlikely that the findings were due to the introduction of an opposite gendered peer, that possibility should be ruled out with the inclusion of a scenario involving a gender-inconsistent toy but a same sex peer (a girl taking another girl's truck) or the inclusion of a scenario that includes a boy and a girl, but with a neutral toy (a girl taking another boy's puzzle). The inclusion of these scenarios would allow for the exclusion of the possibility that the increased attribution of

negative intentions in the gender inconsistent scenario was due to the cross-gender nature of the scenario and not the use of the counter-stereotypic toy.

No differences were found between the scenarios for social reasoning. Including more probing for the justifications would allow for a more rich level of reasoning and therefore more information would be available to discern if there are or are not differences between how the participants were reasoning about the intention and punishment decisions in the scenarios.

Furthermore, since the ingroup and outgroup differences did not come out as expected, or with abundance, a measure should be created with the more explicit goal of comparing ingroup and outgroup attributions of intentions is needed in order to further extricate and disentangle the finding that female participants found the action to be more all right for the male (outgroup) potential transgressor than for the female (ingroup), indicating more negative intentions, not less, for the ingroup.

Another point to address is that the toys used in the current study, a doll and truck, varied not only on their gender stereotypes, but also on their level of personalization. The doll is highly personalized, as a child could identify themselves or a close other with the doll, while the truck is an inanimate object that the child cannot identify with. Using a less personalized female stereotyped toy such as a kitchen set, or a more personalized toy for the male such as an action figure, would eliminate that discrepancy as well as any impact it may have on the findings.

Overall, it is also possible that the younger children, and other children with less capable language skills, were not able to fully understand the intricacies and complexities of the scenarios presented to them. Smetana and Braeges (1990)

conducted research very similar in methodology to the current research. Similar assessment questions and scenarios were presented to the participants, and comparable content and complexity were employed in their interview. Smetana and Braeges (1990) additionally utilized a language comprehension task in their research, and were able to indicate that language comprehension was not a concern for these kinds of questions and scenarios past the age of 3 years of age. They did find that children younger than 3 years-of-age were impacted by less capable language skills as those children had more difficulty discriminating the fine details of a scenario (the difference between moral and social conventional judgments) (Smetana & Braeges, 1990). The inclusion of a language comprehension measure would allow for the impact of language deficiencies to be explored or ruled out in the attribution of intentionality decisions.

Additionally, the demographics of this study are limited. The current study is inclusive of low to middle class subjects only, and can therefore be generalized to this population only. It is possible that children within a lower socio-economic status could have differing age trends and could be impacted differently by the inclusion of gender stereotyped contextual information, which could impact the findings. The inclusion of this sample in a follow-up study would allow for us to examine if a lower SES sample would differ in their attributions of intentionality or the age at which they pass both the independent as well as the embedded false belief task.

Finally, the ethnicity of the participant was not addressed in this study, and the ethnicity of the characters used was randomized, not matched, both of which could be addressed in a follow-up study. It is possible that the ethnicity of the participant, due

to differing social experiences, could be impacting their attributions of intentionality. Killen and colleagues have been able to find in previous studies that the differing social experiences of individuals in differing social and ethnic groups' impacts their social decision making (Killen, Kelly, et al., in press). Therefore, the inclusion of ethnicity as a variable of measure could illustrate an impact of participant ethnicity on attribution of intentionality that is distinct from the findings presented here for the current study. Additionally, it is possible that the participants indicated differing attributions of intentionality or punishment decisions due to the ethnicity of the characters that were presented to them in the scenarios. The ethnicity of the characters was randomized so as to reduce their impact on decision making, however, if the ethnicity was controlled and entered as a variable of measure the possibility of the ethnicity of the characters impacting the judgments and justifications made, could be eliminated.

Future Directions

The most critical future direction of the current study is the exploration of the over attribution of negative intentions that was found. It is critical to understand at what age, and in what manner this over attribution of negative intention is presenting itself, so that a means of ameliorating the problem can be explored. It is possible that pushing the age range of this study back so as to be inclusive of even younger children is not truly an option, as three years of age is possibly the youngest age at which a child could comprehend and acknowledge the complexity and intricacies of the scenarios and assessments presented. If the age range were to be pushed back, a different methodology would have to be created. Utilizing methodologies employed

by researchers in the moral development field who are drawing upon samples of children as young as 14 – 18-months-of-age may allow for the examination of some of the rudimentary aspects of moral decision making that could be impacted by gender stereotypes (Warneken & Tomasello, 2009). In addition, the field of theory of mind has broached a much younger sample as well, exploring the cognitive capacities of infants (Woodward & Needham, 2009). These two lines of research could be paired utilizing a much younger sample than the current study so as to gain further insight into the ontogeny of moral decision making, and the impact of a theory of mind on that decision making.

Furthermore, finding that gender stereotypes were impacting attributions of intentionality and punishment decisions and reasoning, indicates that there could be many other contextual variables that additionally impact these judgments. A full exploration of other contextual variables such as but not limited to ethnicity, social class, social group, and religious affiliation could elucidate the myriad of contextual factors that could be impacting a child's ability to make an attribution of intentionality. Exploring these various possibilities will facilitate our growing understanding of the fields of moral development and false belief theory of mind beyond the scope of what the current study could explore.

Tables

Table 1: Design of Intergroup Attitudes Attribution of Intentions Task Justifications

Description of Design for Intergroup Attitudes Attribution of Intentions Task

Story Code	Potential Transgressor	Victim	Object	Action	Stereotype information in relation to potential transgressor and object
A	Boy	Girl	Doll (low status)	Boy puts doll in boy backpack. Doll belongs to girl.	Stereotype inconsistent
B	Girl	Boy	Truck (high status)	Girl puts truck in girl backpack. Truck belongs to boy.	Stereotype inconsistent
C	Girl	Girl	Doll (low status)	Girl 1 puts doll in Girl 1 backpack. Doll belongs to Girl 2.	Stereotype consistent
D	Boy	Boy	Truck (high status)	Boy 1 puts truck in Boy 1 backpack. Truck belongs to Boy 2.	Stereotype consistent

Table 2: Descriptions of Intergroup Attitudes Attribution of Intentions Task
Descriptions of Intergroup Attitudes Attribution of Intentions Task

Scenario	Description
Gender Inconsistent *Female Transgressor*	Todd and Melissa are children in the same classroom who have backpacks that look the same except Todd's has a T on the front and Melissa's has an M on the front. They are playing with their toys when the teacher asks them to get their backpacks ready to go home. Todd puts his truck next to a pile of backpacks near the door, where both his and Melissa's backpacks are. The children go outside to play, but Melissa stays to help the teacher clean up. Melissa see's Todd's truck and puts it into this backpack (researcher points to Melissa's backpack). Researcher then asks the child to show them where the toy is to make sure that the participant is following the story).

(Table 2 continues)

(Table 2 continued)

Scenario	Description
Gender Inconsistent *Male Transgressor*	Tina and Mark are children in the same classroom who have backpacks that look the same except Tina's has a T on the front and Mark's has an M on the front. They are playing with their toys when the teacher asks them to get their backpacks ready to go home. Tina puts her doll next to a pile of backpacks near the door, where both her and Mark's backpacks are. The children go outside to play, but Mark stays to help the teacher clean up. Mark see's Tina's doll and puts it into this backpack (researcher points to Mark's backpack). Researcher then asks the child to show them where the toy is to make sure that the participant is following the story).
Gender Consistent (male)	Tim and Martin are children in the same classroom who have backpacks that look the same except Tim's has a T on the front and Martin's has an M on the front. They are playing with their toys when the teacher asks them to get their backpacks ready to go home. Tim puts his truck next to a pile of backpacks near the door, where

(Table 2 continues)

(Table 2 continued)

Scenario	Description
	both his and Martin's backpacks are. The children go outside to play, but Martin stays to help the teacher clean up. Martin see's Tim's truck and puts it into this backpack (researcher points to Martin's backpack). Researcher then asks the child to show them where the toy is to make sure that the participant is following the story).
Gender Consistent (female)	Tara and Michelle are children in the same classroom who have backpacks that look the same except Tara's has a T on the front and Michelle's has an M on the front. They are playing with their toys when the teacher asks them to get their backpacks ready to go home. Tara puts her doll next to a pile of backpacks near the door, where both her and Michelle's backpacks are. The children go outside to play, but Michelle stays to help the teacher clean up. Michelle see's Tara's doll and puts it into this backpack (researcher points to Michelle's backpack). Researcher then asks the child to show them where the toy is to make sure that the participant is following the story)

Table 3: Descriptions of Social Reasoning Responses to Intergroup Attitudes

Attribution of Intentions Task Justifications

Description of Social Reasoning Responses to Intergroup Attitudes Attribution of Intentions Task Assessments on Intentions and Punishment

Social Reasoning Category	Social Reasoning Subcategory	Example of Response
A. Moral	1. Psychological harm	She'll be sad; He'll be angry that someone did that to him; She'll get angry
	2. Negligence	He should have looked at the initials; She didn't ask him before she put it in; He should have been more careful
	3. Ownership and Wrongfulness of Stealing	He wanted to take it home; She took it to her house; He shouldn't have put it in his bag; She won't have it because he has it
	4. Prosocial	Well, she was trying to help the teacher clean up the classroom; He was just trying to help
	5. Accident/Lack of Negative Intentions	She should be able to take the doll home
B. Social Conventional	6. Deference to the rules	It's against the rules to put it in your bag; She was just doing what she was told to do
	7. Gender Stereotypes	Girls don't want to take trucks home; it must have been an accident; Girls don't play with trucks; Boys shouldn't play with dolls; Girls don't steal
C. Personal	8. Selfish Desires	She will have what she wants
D. Uncodable	9. Undifferentiated	I don't know; It's bad; Because it's good
	10. Incomplete. inaudible	

Table 4: Summary of Independent Variables

Independent Variable	Criteria
1. Age	3 age groups: 1 (3 – 4.5 years); 2 (4.5 – 6 years); (6 – 8 years)
2. Gender	2 genders: Male; Female
3. Scenario	2 scenarios: gender inconsistent (transgressor takes a toy that is inconsistent with their own gender, girl takes a truck, boy takes a doll); gender consistent (transgressor puts a toy into their own backpack that is consistent with their own gender, boy takes a truck, girl takes a doll)
4. False Belief Theory of Mind	Dichotomous pass (correctly answered all 6 questions on False belief theory of mind task)/fail
5. Gender Stereotype Tolerance	Dichotomous high/low (as decided by a median split for Gender stereotype tolerance scale)
6. Gender Stereotype Flexibility	Dichotomous high/low (as decided by a median split for Gender stereotype flexibility scale)
7. Gender Stereotype Knowledge	Dichotomous high/low (as decided by a median split for Gender stereotype knowledge scale)

Table 5: Summary of Hypotheses

Summary of Hypotheses

Intergroup Attitudes Attribution of Intentions Task

Judgments of Intentionality

1. Participants will indicate the action to be more all right (less negative intentions) as age increases.

2. Participants who have a false belief theory of mind will indicate the action to be more all right, as they will be more likely to be able to see the accidental nature of the transgression.

3. Participants low in gender stereotype flexibility and tolerance will be more likely to indicate that the action is more not all right, or less likely to be able to see the accidental nature of this transgression.

4. Participants that are low in gender stereotype flexibility and tolerance will see more negative intent in the gender consistent scenario than in the gender inconsistent scenario.

5. Participants will find the action to be more all right for ingroup transgressors than for outgroup transgressors for the stereotype inconsistent scenario and the stereotype consistent scenario.

Justifications for Intention Assessments

6. Participants will use more social conventional reasoning – deference to the rules and moral reasoning - no negative intentions as age increases.

7. Participants will use more social conventional reasoning – deference to the rules and moral reasoning - no negative intentions with the indication of a false belief.

8. Male participants will use more social conventional reasoning using gender stereotypes.

9. Participants with lower gender stereotype flexibility and tolerance will use more social conventional reasoning using gender stereotypes.

10. Participants will use more Social Conventional – gender stereotype reasoning in the gender inconsistent scenario than in the gender consistent scenario.

11. Participants will be more likely to think that male and female children can be friends with age.

(Table 5 continues)

(Table 5 continued)

Liking Judgment and Justification

12. Participants who have less stereotype flexibility and tolerance will think that male and female children cannot be friends than participants with more flexibility and tolerance.

False Belief Theory of Mind, Location Change

13. Participants will correctly indicate where the victim should look for the toy with age as children acquire a Theory of Mind for multifaceted scenarios with age.

14. Participants with a false belief will more likely be able to correctly indicate where the victim should look for the toy thus indicating that they have a theory of mind in a morally relevant, multifaceted scenario.

Emotion Judgment of Victim

15. Participants will all indicate that the victim feels bad when the toy is moved.

16. Participants will indicate that the victim will feel more badly in the gender consistent scenario than in the gender inconsistent scenario.

17. Participants with low gender stereotype flexibility and tolerance as will indicate that the victim will feel more badly than the participants with high gender stereotype flexibility and tolerance.

18. Male participants will indicate that the victim will feel more badly than female participants.

Punishment of Transgressor

19. Participants will indicate not to punish, or to punish to a lesser degree as age increases.

20. Participants will indicate not to punish, or to punish to a lesser degree with an indication of having a false belief.

21. Participants with low stereotype flexibility and tolerance will indicate the action to be more punishable for the stereotype inconsistent scenario than for the stereotype consistent scenario.

22. Male participants will indicate the action to be more punishable for the stereotype inconsistent scenario than for the stereotype consistent scenario.

(Table 5 continues)

(Table 5 continued)

23. Participants will find the action to be more all right for an ingroup transgressor than for outgroup transgressors for the stereotype inconsistent scenario and the stereotype consistent scenario.

Justifications for Punishment Judgment

24. Participants will use more social conventional reasoning – deference to the rules and moral reasoning - no negative intentions as age increases.

25. Participants will use more social conventional reasoning – deference to the rules and moral reasoning - no negative intentions with the indication of a false belief.

26. Male participants use more social conventional reasoning using gender stereotypes.

27. Participants with lower gender stereotype flexibility and tolerance will use more reasoning social conventional reasoning using gender stereotypes.

28. Participants will use more social conventional reasoning using gender stereotype in the gender inconsistent scenario than in the gender consistent scenario.

Second Order Theory of Mind

29. Participants will indicate that the transgressor was trying to put the toy into the victim's backpack with age and with an indication of having a Theory of Mind as children acquire a second order Theory of Mind after they acquire a first order theory of mind, and thus theory of mind is requisite for being able to accurately answer this assessment.

30. Participants with second-order Theory of Mind will think that transgressor has less negative intentions than those participants who do not have a second order Theory of Mind.

<u>False Belief Theory of Mind Task</u>

31. Participants will correctly indicate what others will think is in the crayon box with age.

32. Participants will correctly indicate if the children outside saw what was in the crayon box with age.

(Table 5continues)

(Table 5 continued)

33. Participants will correctly indicate the actual contents of the crayon box with age.

34. Participants will correctly indicate where Lenny should look for the markers with age.

35. Participants will correctly indicate if Lenny saw where the teacher put the markers.

36. Participants will correctly indicate where the markers are located.

Gender Stereotype Task

37. Participants will increase in stereotype knowledge, tolerance, and flexibility with age.

38. Male participants will indicate lower gender stereotype flexibility and tolerance than female participants.

Table 6: Summary of Dependent Variables in Intergroup Attitudes Attribution of Intentions Task

Description of Dependent Variables in Intergroup Attitudes Attribution of Intentions Task

Dependent Variable	**Question**	**Potential Answers**
1. Intentions of Transgressor	"Whose backpack did (transgressor's name) think he/she was putting the toy into?"	Two choices: Transgressor's backpack; Victim's backpack
2. Judgment of Transgressor's Intentions, dichotomous	"When (transgressor's name) put the toy truck (or doll) into the backpack did she (or he) think she (or he) was doing something that was all right or not alright?	Two Choices: All right, Not all right
3. Judgment of Transgressor's Intentions, Likert	"How alright (or not alright depending on the answer they provided for the previous question) did (transgressor's name) think he/she was for doing that?"	Likert Scale: 1, Not all right; 4, All right
4. Justification of Transgressor's Intention	"Why?"	Open ended verbal response coded dichotomously as 0 (not present) or 1(present) for each justification category (Psychological harm, Negligence, Ownership and Wrongfulness of Stealing, Prosocial, Accident/Lack of Negative Intentions, Deference to the rules, Gender Stereotypes, Selfish Desires, Undifferentiated, Uncodable)

(Table 6 continues)

(Table 6 continued)

Dependent Variable	Question	Potential Answer
5. Judgment of Transgressor's Action by the Participant, dichotomous	"When (transgressor's name) put the truck (or doll) into the backpack, do you think he (or she) was doing something that was alright or not alright? 6) How alright/not alright do you think she was for doing that?"	Two Choices: All right, Not all right
6. Judgment of Transgressor's Action by the Participant, Likert	"How alright/not alright do you think she was for doing that?"	Likert Scale: 1, Not all right; 4, All right
7. Justification of Transgressor's Action by the Participant	"Why"	Open ended verbal response coded dichotomously as 0 (not present) or 1(present) for each justification category
8. Friendship Judgment	"How much do you think (transgressor's name) and (victim's name) like each other?"	Likert scale: 1, Not at all; 4, A lot
9. Friendship Justification	"Why"	Open ended verbal response coded dichotomously as 0 (not present) or 1(present) for each justification category
10. False Belief Theory of Mind, Location Change	"Where will (victim's name) look for his (or her) truck (or doll)?"	Three choices: on the floor (which is where he (or she) left it and should look for it); in transgressor's backpack (where it actually is); in victim's own backpack
11. False Belief Theory of Mind, Location Change Justification	"Why will (victim's name) look there?"	Open ended verbal response coded dichotomously as 0 (not present) or 1(present) for each justification category

(Table 6continues)

(Table 6 continued)

Dependent Variable	Questions	Potential Answers
12. Emotion Judgment of the Victim	"When (victim's name) finds out his (or her) truck (or doll) is not where he (or she) left it, how will (victim's name) feel?"	Two choices: Good, Bad
13. Punishment of Transgressor, Yes/No	"Should (transgressor's name) get in trouble for putting the truck (or doll) into the backpack (researcher points to transgressor's backpack)?"	Two Choices: Yes, No
14. Punishment of Transgressor, Amount	(If the participant answers yes to previous question) "How much trouble should (transgressor's name) get in?"	Two Choices: A little; A lot
15. Justification of Punishment of Transgressor	"Why"	Open ended verbal response coded dichotomously as 0 (not present) or 1(present) for each justification category
16. Second Order Theory of Mind	"What did (victim's name) think that (transgressor's name) thought he (or she) was trying to do?	Two Choices: Put (victim's name) truck (or doll) in (transgressor's name) backpack; Put (victim's name) into (victim's name) backpack
17. Justification for Second Order Theory of Mind	"Why"	Open ended verbal response coded dichotomously as 0 (not present) or 1(present) for each justification category
18. Emotion Judgment after Knowledge of Transgression, dichotomous	"How will (victim's name) feel about (transgressor's name) now that he (or she) knows (transgressor's name) put the truck (or doll) into this backpack (researcher points to transgressor's backpack)?"	Two Choices: Good; Bad

(Table 6 continues)

(Table 6 continued)

Dependent Variable	Question	Potential Answers
19. Emotion Judgment after Knowledge of Transgression, Likert	"How good (or bad, depending on the participant's answer to the previous question) will (victim's name) feel about (transgressor's name)?"	Likert scale: 1, bad; 4, good
20. Justification for Emotion Judgment after Knowledge of Transgression	"Why"	Open ended verbal response coded dichotomously as 0 (not present) or 1(present) for each justification category

Table 7: Proportions of Reasoning about Evaluations of Punishment of Transgressor

Summary of Proportions of Justifications for Judgment of Punishment of Transgressor

	Moral					Social Conventional	
	Negative Intentions Indicated		Positive Intentions Indicated			Positive Intentions Indicated	
	Psych. Harm	Wrong-fulness of Stealing	Negli-gence	Pro-social	Accident/ Lack of Negative Intentions	Defer-ence to the Rules	Gender Stereo-types
Gender incon-sistent	.081 (.251)	.264 (.426)	.175 (.351)	.029 (.147)	.159 (.34067)	.102 (.300)	.020 (.134)
Gender consistent	.074 (.254)	.217 (.406)	.139 (.317)	.049 (.207)	.180 (.358)	.119 (.322)	.004 (.045)

Table 8: Percentage and Amount of Participants Who Passed the False Belief Tasks

Summary of Percentage and Amount of Participants Who Passed the False Belief Tasks

Age Range	Independent False Belief Task		Stereotype Inconsistent Embedded False Belief Task		Stereotype Consistent Embedded False Belief Task	
	n	%	*n*	%	*n*	%
3 – 4 years	4	10%	1	2.3%	1	2.3%
5 – 6 years	16	44.4%	7	17.5%	4	12.5%
7 – 8 years	37	88.1%	21	48.8%	24	55.8%

Table 9: Proportions of Reasoning about Evaluation of Transgressor's Intentions

Summary of Proportions of Justifications for Judgment of Transgressor's Intentions

	Moral					Social Conventional	
	Negative Intentions Indicated		Positive Intentions Indicated			Positive Intentions Indicated	
	Psych. Harm	Wrong-fulness of Stealing	Negli-gence	Pro-social	Accident/ Lack of Negative Intentions	Defe-rence to the Rules	Gender Stereo-types
Gender in-consistent	.142 (.336)	.252 (.426)	.220 (.369)	.065 (.239)	.122 (.296)	.065 (.23919)	.016 (.127)
Gender consistent	.081 (.267)	.240 (.407)	.195 (.360)	.077 (.256)	.155 (.334)	.057 (.224)	.000 (.000)

Figures

Figure 1: Judgment of Intentionality by Age

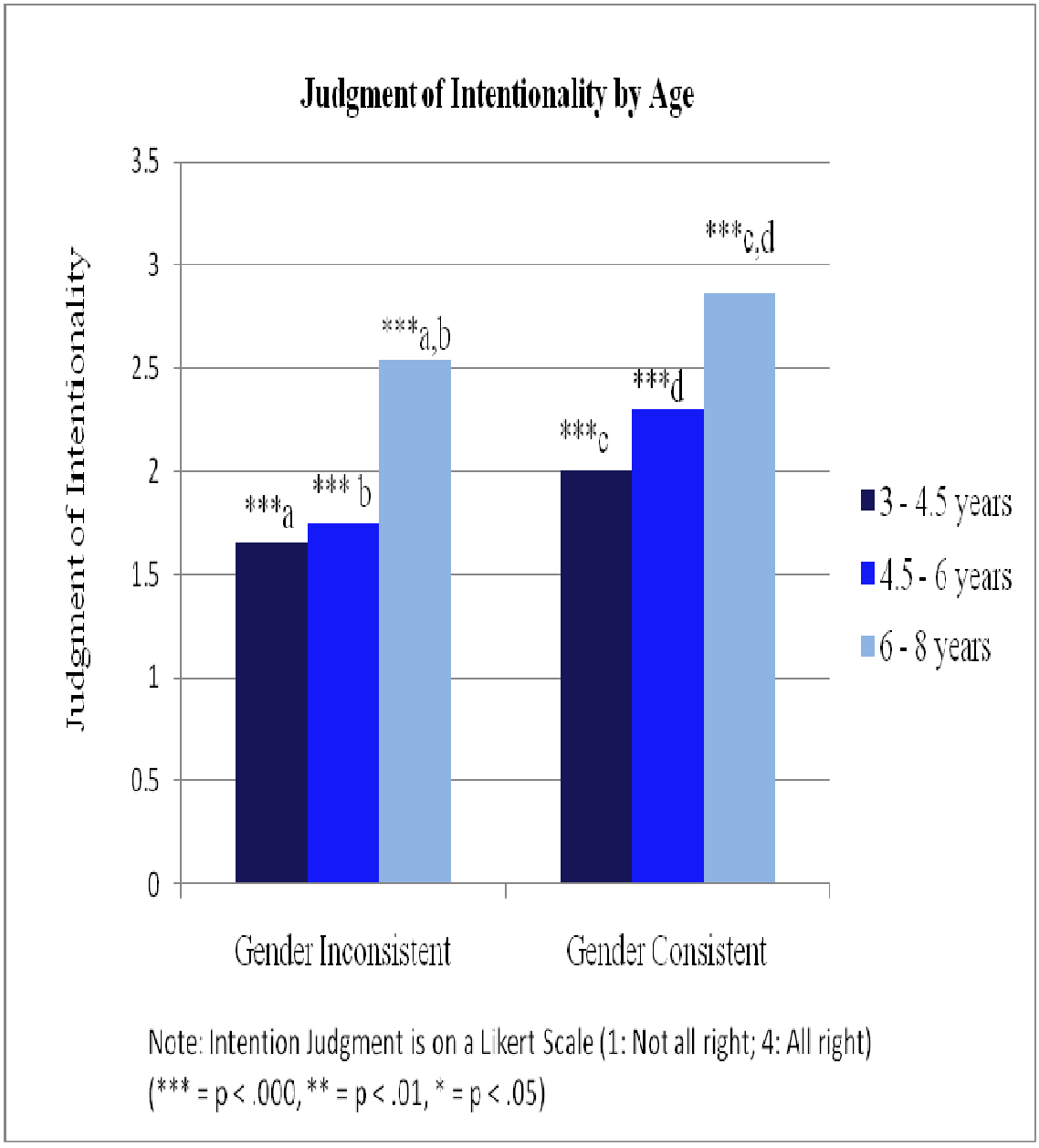


Figure 2: Judgment of Intentionality by False Belief Theory of Mind

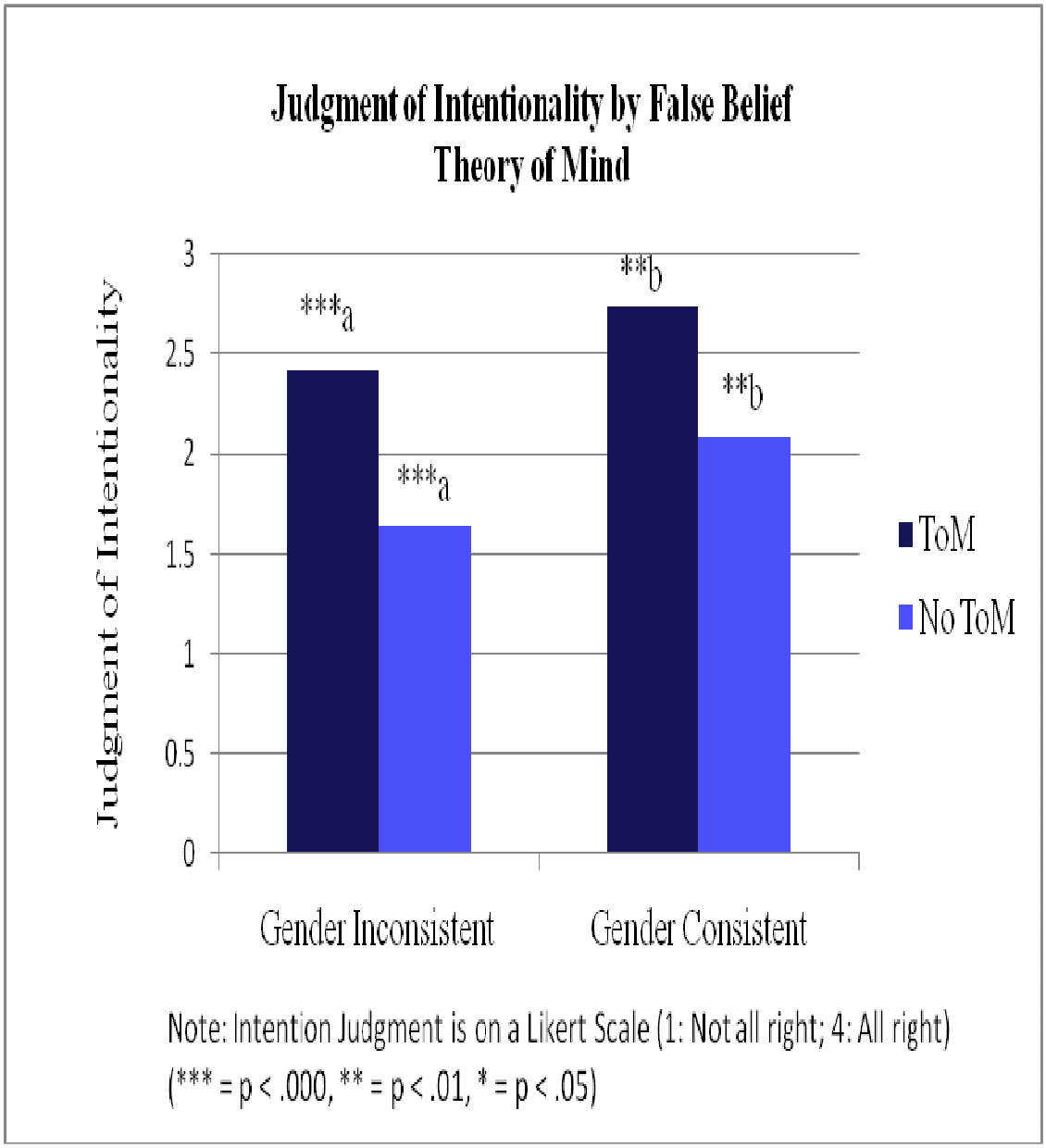


Figure 3: Judgment of Intentionality by Scenario

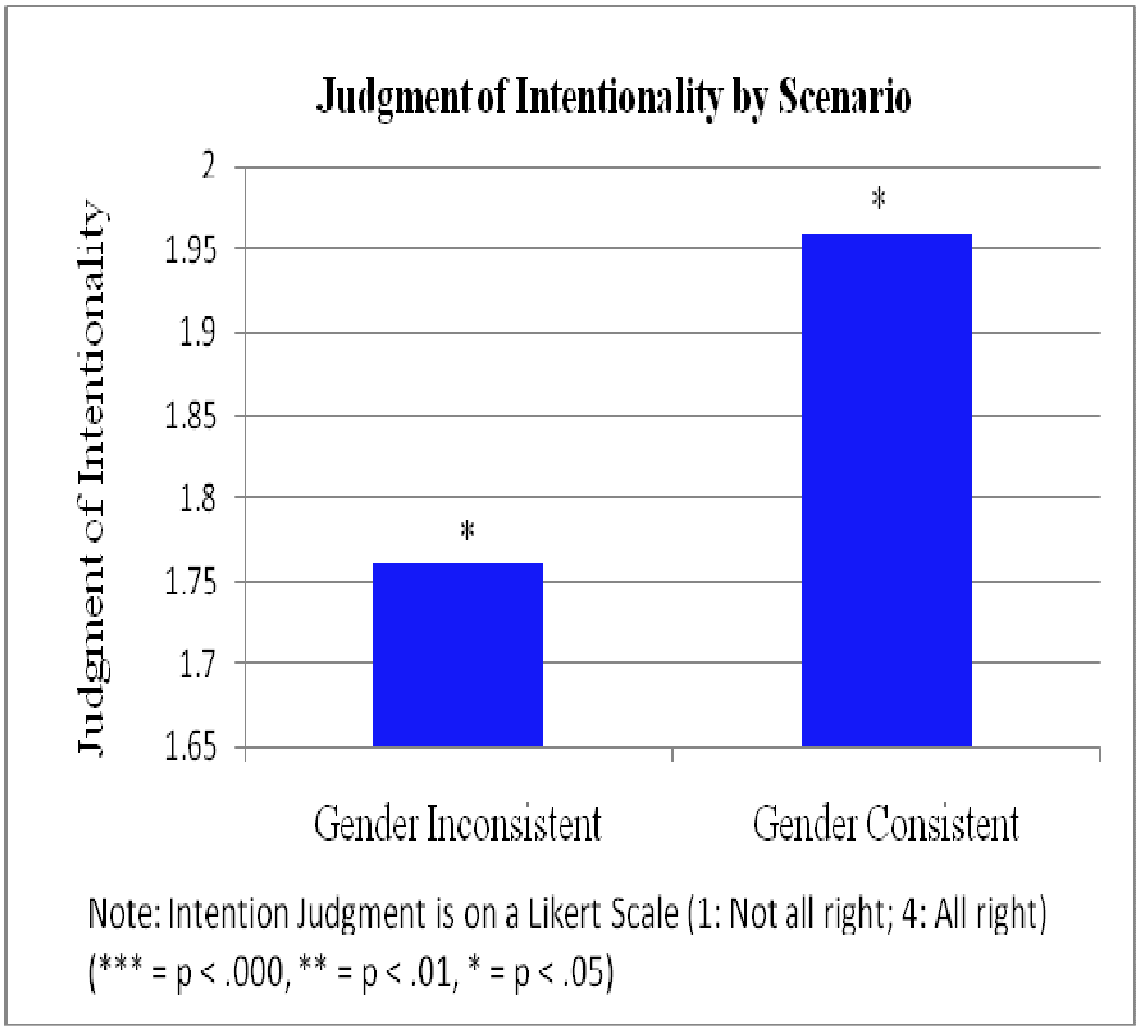


Figure 4: Judgment of Punishment Acceptability by Age

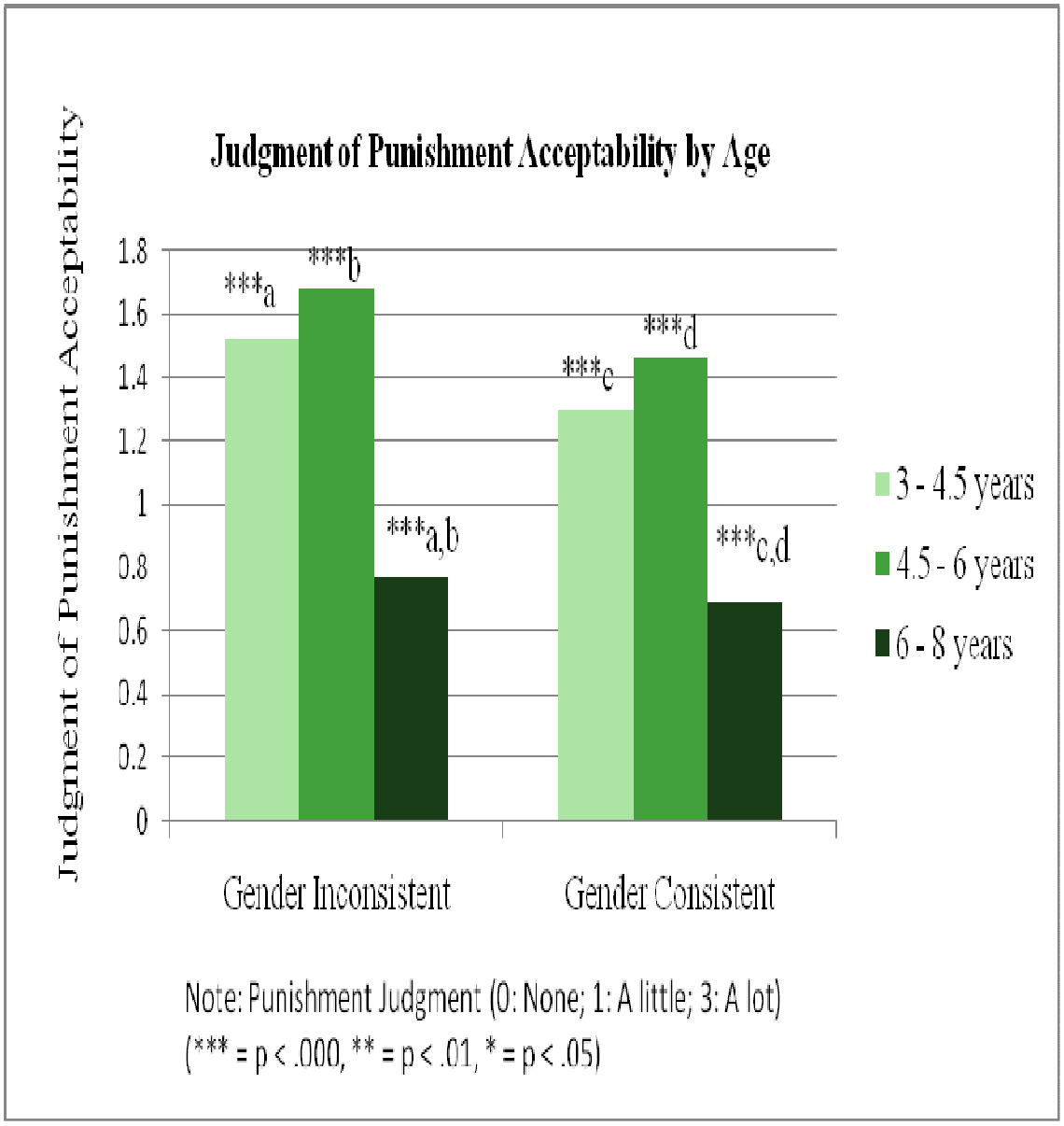


Figure 5: Judgment of Punishment Acceptability by False Belief Theory of Mind

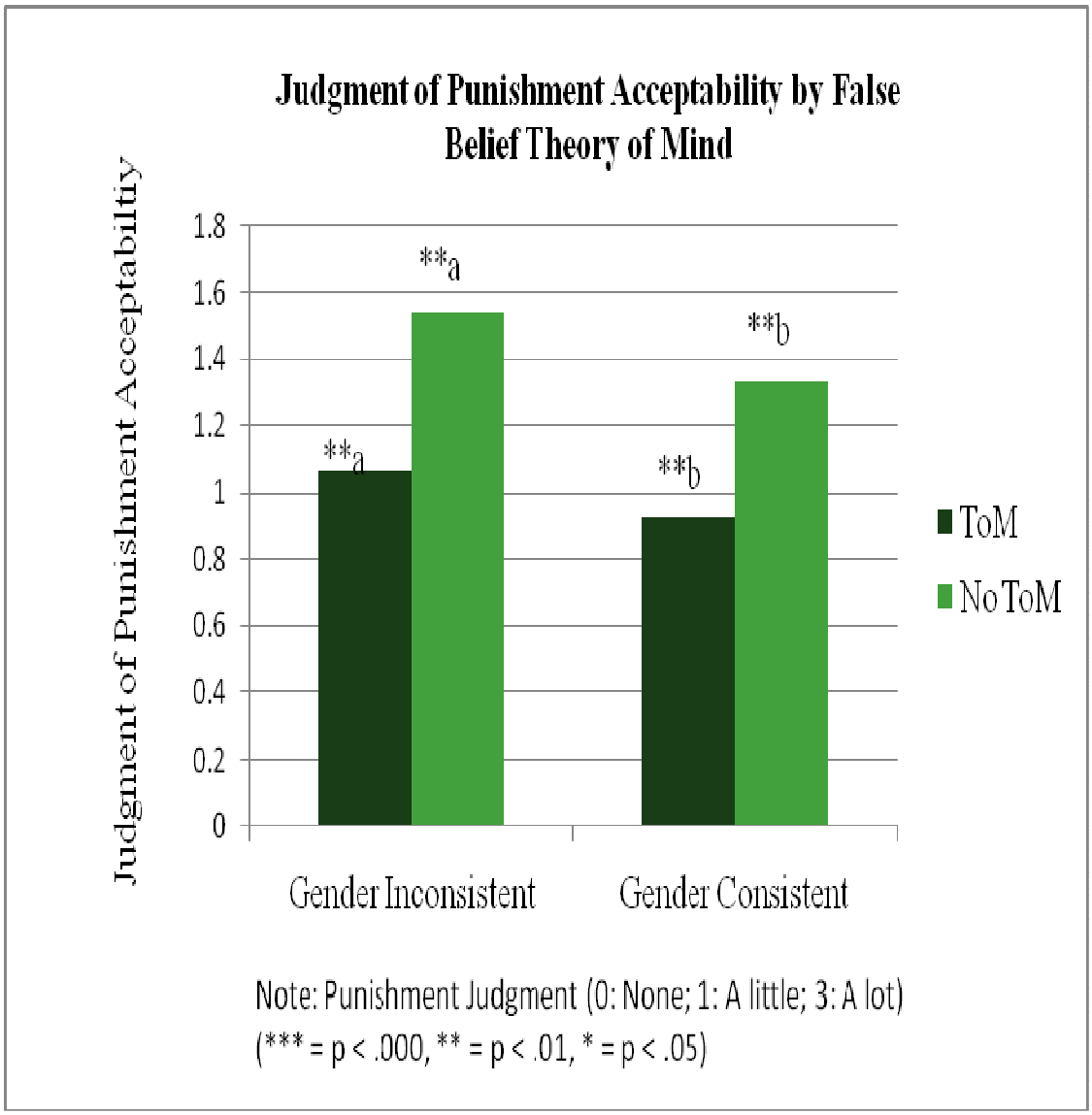


Figure 6: Judgment of Punishment Acceptability by Scenario

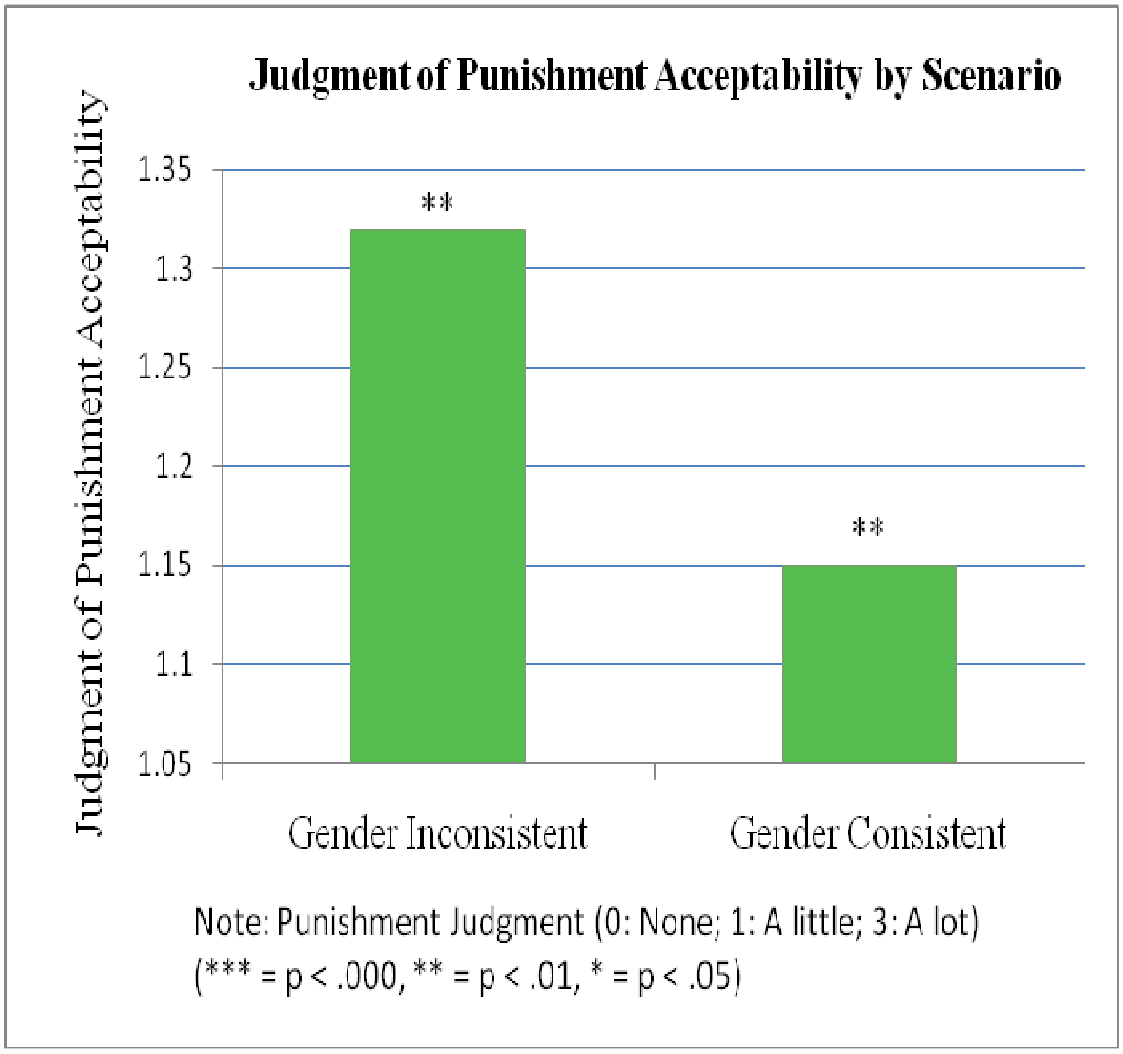
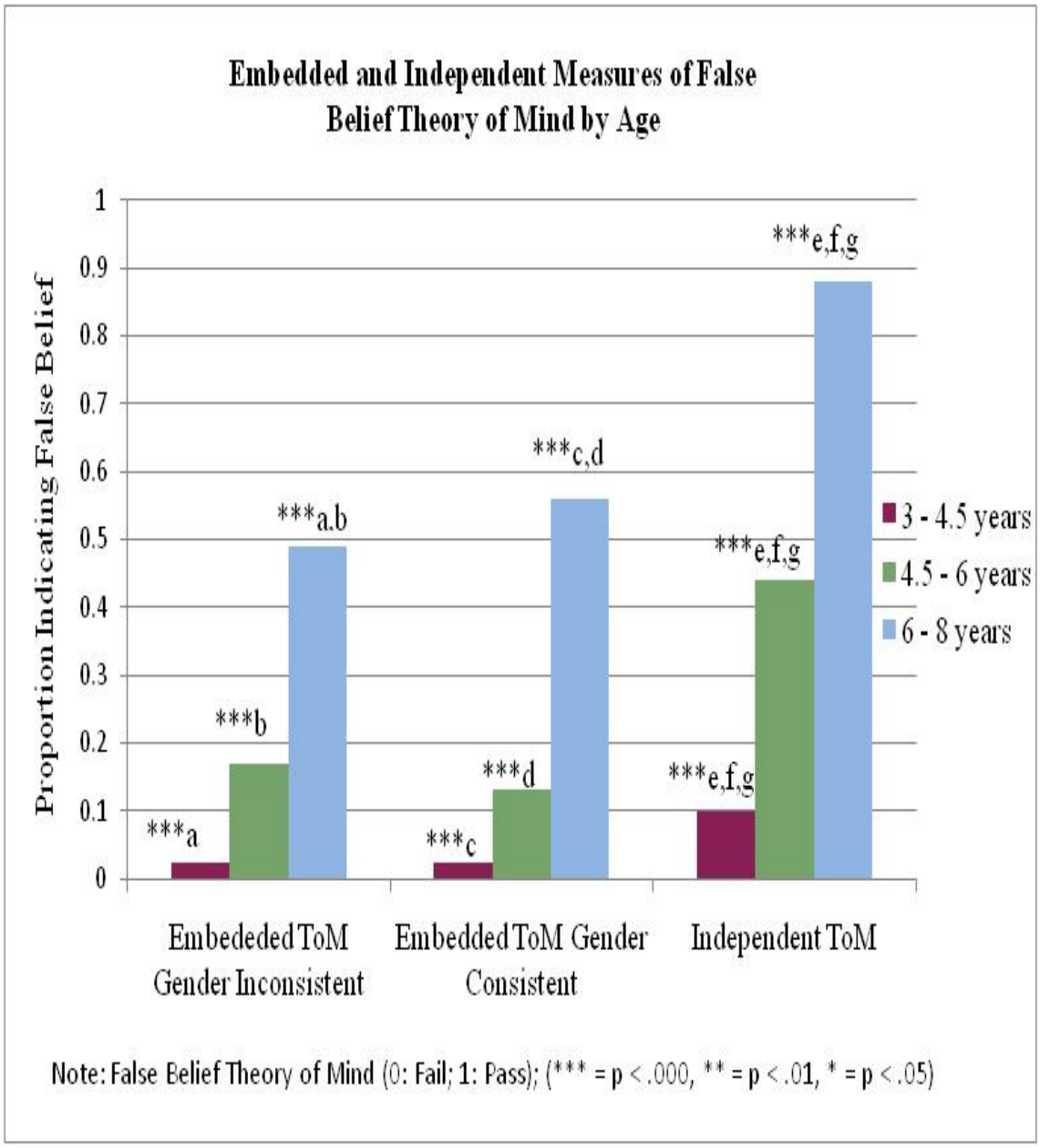


Figure 7: Embedded and Independent Measures of False Belief Theory of Mind by Age



Appendices

Appendix A: Initial IRB Approval

UNIVERSITY OF MARYLAND
INSTITUTIONAL REVIEW BOARD

0101 Lee Building
College Park, Maryland 20742-5125
301.405.4212 TEL 301.314.1475 FAX
irb@deans.umd.edu
www.umresearch.umd.edu/IRB

February 19, 2010

MEMORANDUM

Application Approval Notification

To: Dr. Melanie Killen
Kelly Lynn Mulvey
Human Development

From: Joseph M. Smith, MA, CIM
IRB Manager
University of Maryland, College Park

Re: **IRB Application Number:** 09-0166
Project Title: "Children's Attributions of Intentionality"

Approval Date: February 19, 2010

Expiration Date: February 19, 2011

Type of Application: Renewal

Type of Research: Non-Exempt

Type of Review for Application: Expedited

The University of Maryland, College Park Institutional Review Board (IRB) approved your IRB application. The research was approved in accordance with the University IRB policies and procedures and 45 CFR 46, the Federal Policy for the Protection of Human Subjects. Please include the above-cited IRB application number in any future

communications with our office regarding this research.

Recruitment/Consent: For research requiring written informed consent, the IRB-approved and stamped informed consent document is enclosed. The expiration date for IRB approval has been stamped on the informed consent document. Please keep copies of the consent forms used for this research for three years after the completion of the research.

Continuing Review: If you intend to continue to collect data from human subjects or to analyze private, identifiable data collected from human subjects, after the expiration date for this approval (indicated above), you must submit a renewal application to the IRB Office at least 45 days before the approval expiration date. If IRB approval of your project expires, all human subject research activities including the enrollment of new subjects, data collection, and analysis of identifiable private information must stop until the renewal application is approved by the IRB.

Modifications: Any changes to the approved protocol must be approved by the IRB before the change is implemented, except when a change is necessary to eliminate apparent immediate hazards to the subjects. If you would like to modify the approved protocol, please submit an addendum request to the IRB Office. The instructions for submitting a request are posted on the IRB web site at : http://www.umresearch.umd.edu/IRB/addendum app.htm.

Unanticipated Problems Involving Risks: You must promptly report any unanticipated problems involving risks to subjects or others to the IRB Manager at 301-405-0678 or jsmith@umresearch.umd.edu.

Student Researchers: Unless otherwise requested, this IRB approval document was sent to the Principal Investigator (PI). The PI should pass on the approval document or a copy to the student researchers. This IRB approval document may be a requirement for student researchers applying for graduation. The IRB may not be able to provide copies of the approval documents if several years have passed since the date of the original approval.

Additional Information: Please contact the IRB Office at 301-405-4212 if you have any IRB-related questions or concerns or email at irb@umd.edu.

PARENTAL CONSENT FORM

Project Title	Children's Attributions of Intentions
Why is this research being done?	The purpose of the research is to understand how children attribute intentions in situations involving ownership of toys.
What will your child be asked to do?	The procedure involves a one-time, audiotape-recorded individually administered interview session, lasting approximately 30 minutes. Your child will be interviewed in a quiet setting by a trained research assistant from the University of Maryland. A few short stories, developed by the researcher, about everyday peer encounters and social rule transgressions along with illustrated picture cards will be presented to your child and simple, straightforward questions evaluating the situation will be asked. In addition, your child will be asked questions to evaluate in what circumstances they can take other people's perspectives. Example questions include: Where did the teacher put the markers? Does Sally know where the teacher put the markers? Why or why not? They will also be asked questions about who can and should play with everyday toys (like dolls and trucks).
What about confidentiality?	We will do our best to keep your personal information confidential. We will not share your answers with anyone, including your teachers, principal, or parents. If we write a report or article about this research project, your name will not appear in it.
What are the risks of this research?	There are no known risks associated with this research.
What are the benefits of this research?	This research is not designed to help your child personally. Instead, research is obtained about age-related patterns regarding stereotypes and social rule transgressions. The results will help us learn more about what kids think about social relationships. Educators, counselors, and school professionals will incorporate the findings into their curriculum and guidance programs through reports made available by us to the participating schools
Do I have to be in this research? **May I stop participating at any time?**	Your child's participation is strictly voluntary. You can ask any questions at any time, or withdraw your child from participation at any time. Your child may decide to stop participating at any time and will not be penalized or lose any benefits. Participation is not a school or class requirement. Participation will not affect your child's grades or performance evaluation.
What if I have questions?	This research is being conducted by **Dr. Melanie Killen**, a professor in the Department of Human Development at the University of Maryland, College Park. If you have any questions about the research study itself, please contact Dr. Killen at: **Department of Human Development, 3304 Benjamin Building, College Park, MD 20742-1131; (telephone) 301-405-3176.** If you have questions about your rights as a research subject or wish to report a research-related injury, please contact: **Institutional Review Board (IRB) Office, University of Maryland, College Park, Maryland, 20742; (e-mail) irb@deans.umd.edu; (telephone) 301-405-0678.** This research has been reviewed according to the University of Maryland, College Park IRB procedures for research involving human subjects.
Consent	Your signature indicates that: the research has been explained to you; your questions have been fully answered; and that you allow your child to participate in this research project.
Signature and Date	**Child's Name**
	Parental Signature
	Date

Child's Class: _____

For all your queries and additional information contact: **irb@umd.edu**

IRB APPROVED
EXPIRES ON
FEB 19 2011
UNIVERSITY OF MARYLAND

Appendix B: Parental Consent Form

UNIVERSITY OF MARYLAND

Institute for Child Study/Department of Human Development
3304 Benjamin Building
College Park, MD 20742-1131

Melanie Killen, Ph.D.
Office: 301.405.3176
Email: mkillen@umd.edu

Dear Parents/ Guardians:

We are conducting a project on children's perspective taking about property ownership. This is a follow-up study to one that we recently conducted in which we found that children, under about 7 years of age, had difficulty determining another child's intentions in situations in which children were asked to take the perspective of another child. We would like to ask your permission to interview your son or daughter for this new project in which we are varying the stories to have more information regarding ownership of property.

Children are told simple short vignettes about typical exchanges between children in the classroom (such as taking a toy from someone else) and then asked to respond to a series of simple open-ended questions. For example, one child "mistakenly" puts a doll or truck owned by another child into their own backpack. Participants are asked whether the child intended to keep the toy, and what his or her intentions will be regarding their next action. In addition to administering the "intentionality" story, we will ask children about ownership and interest of toys (e.g., who likes to play with trucks?). There are no right or wrong answers. We are interested in documenting the ways in which children's judgments and reasoning are related to the development of perspective taking.

All information is confidential and anonymous. Please read the description of the project on the reverse side of this letter. **If you are willing to have your child participate in the project, please fill out the information and return it to your child's teacher.**

The results of this study will help teachers, counselors, and educators understand children's intentions and social development. This research has been approved by the Institutional Review Board at the University of Maryland.

We thank you, in advance, for reading this letter, and for your willingness to allow your child to participate. We have found that children enjoy the interviews. They enjoy hearing the stories, and they appreciate the chance to express their viewpoints to an interested adult.

Thank you,
Melanie Killen, Ph.D., Professor
Megan Clark Kelly, Doctoral Candidate

PARENTAL CONSENT FORM

Project Title	Children's Attributions of Intentions
Why is this research being done?	The purpose of the research is to understand how children attribute intentions in situations involving ownership of toys.
What will your child be asked to do?	The procedure involves a one-time, audiotape-recorded individually administered interview session, lasting approximately 30 minutes. Your child will be interviewed in a quiet setting by a trained research assistant from the University of Maryland. A few short stories, developed by the researcher, about everyday peer encounters and social rule transgressions along with illustrated picture cards will be presented to your child and simple, straightforward questions evaluating the situation will be asked. In addition, your child will be asked questions to evaluate in what circumstances they can take other people's perspectives. Example questions include: Where did the teacher put the markers? Does Sally know where the teacher put the markers? Why or why not? They will also be asked questions about who can and should play with everyday toys (like dolls and trucks).
What about confidentiality?	We will do our best to keep your personal information confidential. We will not share your answers with anyone, including your teachers, principal, or parents. If we write a report or article about this research project, your name will not appear in it.
What are the risks of this research?	There are no known risks associated with this research.
What are the benefits of this research?	This research is not designed to help your child personally. Instead, research is obtained about age-related patterns regarding stereotypes and social rule transgressions. The results will help us learn more about what kids think about social relationships. Educators, counselors, and school professionals will incorporate the findings into their curriculum and guidance programs through reports made available by us to the participating schools
Do I have to be in this research? **May I stop participating at any time?**	Your child's participation is strictly voluntary. You can ask any questions at any time, or withdraw your child from participation at any time. Your child may decide to stop participating at any time and will not be penalized or lose any benefits. Participation is not a school or class requirement. Participation will not affect your child's grades or performance evaluation.
What if I have questions?	This research is being conducted by **Dr. Melanie Killen**, a professor in the Department of Human Development at the University of Maryland, College Park. If you have any questions about the research study itself, please contact Dr. Killen at: **Department of Human Development, 3304 Benjamin Building, College Park, MD 20742-1131; (telephone) 301-405-3176.** If you have questions about your rights as a research subject or wish to report a research-related injury, please contact: **Institutional Review Board (IRB) Office, University of Maryland, College Park, Maryland, 20742; (e-mail) irb@deans.umd.edu; (telephone) 301-405-0678.**This research has been reviewed according to the University of Maryland, College Park IRB procedures for research involving human subjects.
Consent	Your signature indicates that: the research has been explained to you; your questions have been fully answered; and that you allow your child to participate in this research project.
Signature and Date	**Child's Name** **Child's Classroom** **Child's Birth Date**
	Parental Signature
	Date

Appendix C: Complete Interview

University of Maryland
Attributions of Intentions in Early Childhood

Date of Interview:_____

Interviewer's Initials:_____

Participant Number:_____

Participant Name:_____

Date of Birth:_____

Gender: M (F)

Ethnicity: B W H A other_____

School:_____

Teacher:_____

INTRODUCTION:

I am going to show you pictures of some kids and tell you a little bit about them. Then I want to ask you some questions about these kids. After you look at the pictures, I am going to show you some cards of kids doing different things and then I will ask you some questions about the kids in the cards. I am interested in finding out what children your age think about things kids do. There are no right or wrong answers. This is not a test. No one will see your answers. So just tell me what you think. Do you have any questions?

We are going to tape-record this interview to help us remember what we talked about. So, before we start, let's make sure this tape-recorder works.

Rewind and check tape-recording:

Tape-Recorder Check: "This is (Name of Interviewer) and I'm talking with (Name of Interviewee). (Interviewee's name's) birth date is _____. Today's date is _____.

Introduce Likert Scale: Have you ever seen one of these before?

This is a way we show how much we think something is good or bad (pointing to correct sides of the scale). Can you show me how much you like pizza? (Wait until the child shows you). I like it this much (use the opposite side of the scale). Can you show me how much you like playing outside? Do you think you understand how to use this? (Check for understanding and continue with examples if needed).

* Possible counter probes are noted with [*brackets*]

Story 1:

This is Tina and this is Mark (show pictures). They are children in this classroom who like playing with dolls and trucks and balls. Tina and Mark have backpacks that look the same, except Tina's has a T on the front and Mark's has an M on the front. This is Tina's (show T backpack) and this is Mark's (show M backpack). They are playing with their toys at school when the teacher asks them to get their backpacks ready to go home before they go outside for recess. Tina picks up her doll and puts it next to the pile of backpacks near the door, where her backpack and Mark's backpack are sitting. Then, she goes outside for recess. Mark stays inside to help the teacher clean up the classroom. Mark sees Tina's doll and puts it into this backpack (show M backpack). Can you show me which backpack the doll is in now? (If incorrect, retell/clarify).

Q1. Whose backpack did Mark **think** he was putting the doll into? (Intent and check)

MARK'S TINA'S

Q2. When Mark put the doll into the backpack, did **he** think he was doing something that was alright or not alright?

NOT ALRIGHT ALRIGHT

Q3. How alright/not alright (READ WHICHEVER THEY PICKED IN THE LAST QUESTION) did Mark think he was for doing that?

1 2 3 4
NOT ALRIGHT ALRIGHT

Q4.Why?
[What makes it good or bad?]

_____

_____

Q5.When Mark put the doll into the backpack, do **you** think he was doing something that was alright or not alright?

NOT ALRIGHT ALRIGHT

Q6. How alright/not alright (READ WHICHEVER THEY PICKED IN THE LAST QUESTION) do you think he was for doing that?

1 2 3 4
NOT ALRIGHT ALRIGHT

Q7.Why?
[What makes it good or bad?]

_____
_____

Q8. How much do you think Mark and Tina like each other?

1 2 3 4
NOT AT ALL A LOT

Q9.Why?

_____
_____

Now, Tina comes back inside to get ready to go home.

Q10. Where will Tina look for her doll?

ON THE FLOOR IN MARK'S BACKPACK

IN TINA'S BACKPACK

Q11. Why will Tina look there?

_____
_____

Q12. When Tina finds out her doll is not where she left it, how will Tina feel?

GOOD BAD

Q13. Should Mark get in trouble for putting the doll into this backpack (point to M backpack)?

YES NO

Q14. (If Yes) How much trouble should Mark get in?

A LITTLE A LOT

Q15. Why or why not?

_____
_____

Q16. What did Tina think that Mark thought he was trying to do?

PUT TINA'S DOLL INTO MARK'S BACKPACK

PUT TINA'S DOLL INTO TINA'S BACKPACK

Q17. Why?

_____

_____

Q18. How will Tina feel about Mark now that she knows Mark put the doll into this backpack (point to backpack M)?

GOOD BAD NEUTRAL

Q19. How good/bad (READ WHICHEVER THEY PICKED IN THE LAST QUESTION) will Tina feel about Mark?

1 2 3 4
BAD GOOD

Q20. Why?

_____

_____

Story 2:
This is Tara and this is Michelle (show pictures). They are children in this classroom who like playing with dolls and trucks and balls. Tara and Michelle have backpacks that look the same, except Tara's has a T on the front and Michelle's has an M on the front. This is Tara's (show T backpack) and this is Michelle's (show M backpack). They are playing with their toys at school when the teacher asks them to get their backpacks ready to go home before they go outside for recess. Tara picks up her doll and puts it next to the pile of backpacks near the door, where her backpack and Michelle's backpack are sitting. Then, she goes outside for recess. Michelle stays inside to help the teacher clean up the classroom. Michelle sees Tara's doll and puts it into this backpack (show M backpack). Can you show me which backpack the doll is in now? (If incorrect, retell/clarify).

Q1. Whose backpack did Michelle **<u>think</u>** she was putting the doll into?

MICHELLE'S TARA'S

Q2. When Michelle put the doll into the backpack, did **<u>she</u>** think she was doing something that was alright or not alright?

NOT ALRIGHT ALRIGHT

Q3. How alright/not alright (READ WHICHEVER THEY PICKED IN THE LAST QUESTION) did Michelle think she was for doing that?

1 2 3 4
NOT ALRIGHT ALRIGHT

Q4.Why?
[What makes it good or bad?]

_____
_____

Q5.When Michelle put the doll into the backpack, do **<u>you</u>** think she was doing something that was alright or not alright?

NOT ALRIGHT ALRIGHT

Q6. How alright/not alright (READ WHICHEVER THEY PICKED IN THE LAST QUESTION) do you think she was for doing that?

1 2 3 4
NOT ALRIGHT ALRIGHT

Q7.Why?
[What makes it good or bad?]

_____
_____

Q8. How much do you think Michelle and Tara like each other?

1	2	3	4
NOT AT ALL			A LOT

Q9.Why?

_____
_____

Now, Tara comes back inside to get ready to go home.

Q10. Where will Tara look for her doll?

ON THE FLOOR IN MICHELLE'S BACKPACK

IN TARA'S BACKPACK

Q11. Why will Tara look there?

_____
_____

Q12. When Tara finds out her doll is not where she left it, how will Tara feel?

GOOD BAD

Q13. Should Michelle get in trouble for putting the doll into this backpack (point to M backpack)?

YES NO

Q14. (If Yes) How much trouble should Michelle get in?

A LITTLE A LOT

Q15. Why or why not?

_____
_____

Q16. What did Tara think that Michelle thought she was trying to do?

PUT TARA'S DOLL INTO MICHELLE'S BACKPACK

PUT TARA'S DOLL INTO TARA'S BACKPACK

Q17. Why?

_____

_____

Q18. How will Tara feel about Michelle now that she knows Michelle put the doll into this backpack (point to backpack M)?

GOOD BAD NEUTRAL

Q19. How good/bad (READ WHICHEVER THEY PICKED IN THE LAST QUESTION) will Tara feel about Michelle?

1 2 3 4
BAD GOOD

Q20. Why?

_____

_____

Story 3:
See this box (pointing to a crayon box)? This is a crayon box. Now here is Sarah. She is cleaning up the classroom and puts some crackers in the empty crayon box.

Q1. When the other children come back in from playing outside, what will they think is in the crayon box?

CRAYONS CRACKERS

Q2. Did the children who were playing outside see Sarah put the crackers in the box?

YES NO

Q3. What is really in the crayon box?

CRAYONS CRACKERS

Story 4:
Lenny is using the markers before recess over at the art table. Lenny goes outside to play and the teacher, Mr. Jones puts the markers in the cabinet.

Q1. When Lenny comes back inside from recess, where will he look for the markers?

[IF THE PARTICIPANT SAYS "EVERYWHERE," ASK THEM WHERE LENNY WILL LOOK FIRST]

ART TABLE CABINET

Q2. Did Lenny see where Mr. Jones put the markers?

YES NO

Q3. Where are the markers really located?

ON THE ART TABLE IN THE CABINET

Knowledge and Ownership of Toys

Procedure. A picture of a toy is placed in front of the child and the interviewer administers 3 questions per toy. Randomly choose a toy and fill in which toy is presented.

Now I'm going to show you pictures of different toys. I'll ask you a few questions about them and I want you to tell me what you think.

TOY 1: _____

Q1. Who usually plays with this toy?

BOY GIRL BOTH

Q2. Who can play with this toy?

BOY GIRL BOTH

Q3. How much would you like it if a boy wanted to play with this toy?

1	2	3	4
NOT AT ALL			A LOT

TOY 2: _____

Q1. Who usually plays with this toy?

BOY GIRL BOTH

Q2. Who can play with this toy?

BOY GIRL BOTH

Q3. How much would you like it if a boy wanted to play with this toy?

1	2	3	4
NOT AT ALL			A LOT

TOY 3: _____

Q1. Who usually plays with this toy?

BOY GIRL BOTH

Q2. Who can play with this toy?

BOY GIRL BOTH

Q3. How much would you like it if a boy wanted to play with this toy?

1 2 3 4
NOT AT ALL A LOT

TOY 4: _____

Q1. Who usually plays with this toy?

BOY GIRL BOTH

Q2. Who can play with this toy?

BOY GIRL BOTH

Q3. How much would you like it if a boy wanted to play with this toy?

1 2 3 4
NOT AT ALL A LOT

TOY 5: _____

Q1. Who usually plays with this toy?

BOY GIRL BOTH

Q2. Who can play with this toy?

BOY GIRL BOTH

Q3. How much would you like it if a boy wanted to play with this toy?

1 2 3 4
NOT AT ALL A LOT

TOY 6: _____

Q1. Who usually plays with this toy?

BOY GIRL BOTH

Q2. Who can play with this toy?

BOY GIRL BOTH

Q3. How much would you like it if a boy wanted to play with this toy?

1 2 3 4
NOT AT ALL A LOT

TOY 7: _____

Q1. Who usually plays with this toy?

BOY GIRL BOTH

Q2. Who can play with this toy?

BOY GIRL BOTH

Q3. How much would you like it if a boy wanted to play with this toy?

1	2	3	4
NOT AT ALL			A LOT

TOY 8: _____

Q1. Who usually plays with this toy?

BOY GIRL BOTH

Q2. Who can play with this toy?

BOY GIRL BOTH

Q3. How much would you like it if a boy wanted to play with this toy?

1	2	3	4
NOT AT ALL			A LOT

TOY 9: _____

Q1. Who usually plays with this toy?

BOY GIRL BOTH

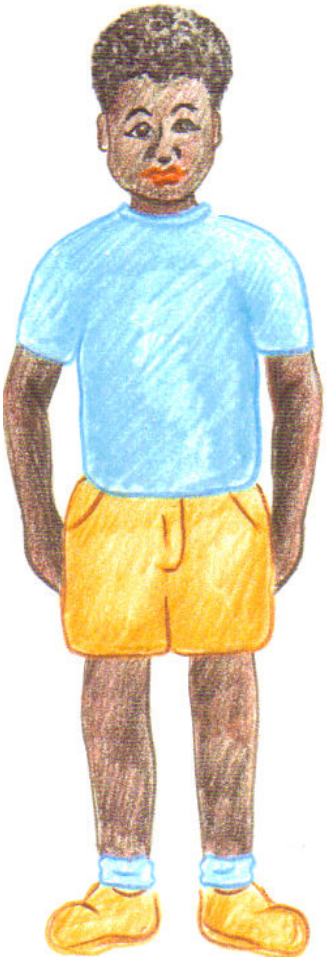
Q2. Who can play with this toy?

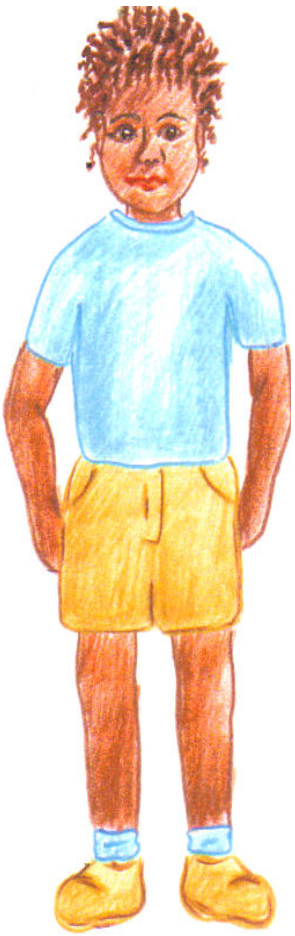
BOY GIRL BOTH

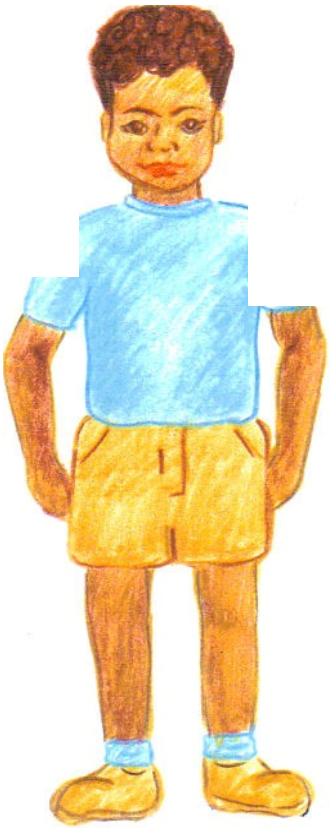
Q3. How much would you like it if a boy wanted to play with this toy?

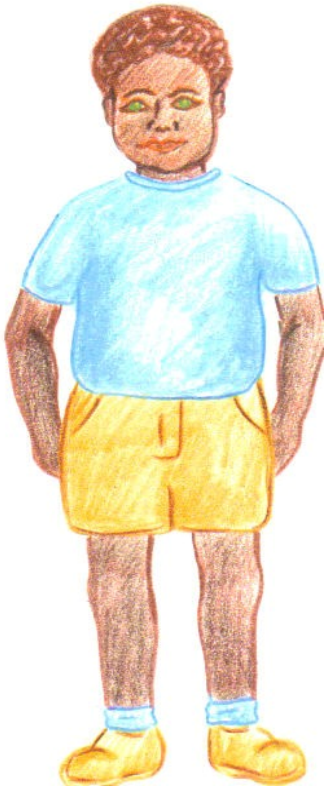
1	2	3	4
NOT AT ALL			A LOT

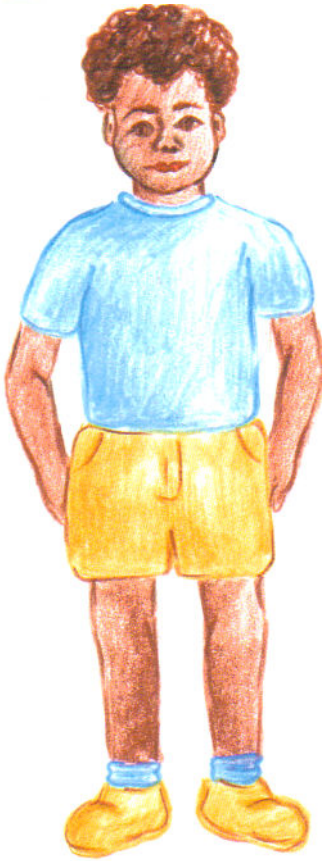
Appendix D: Drawings for use in the Interview

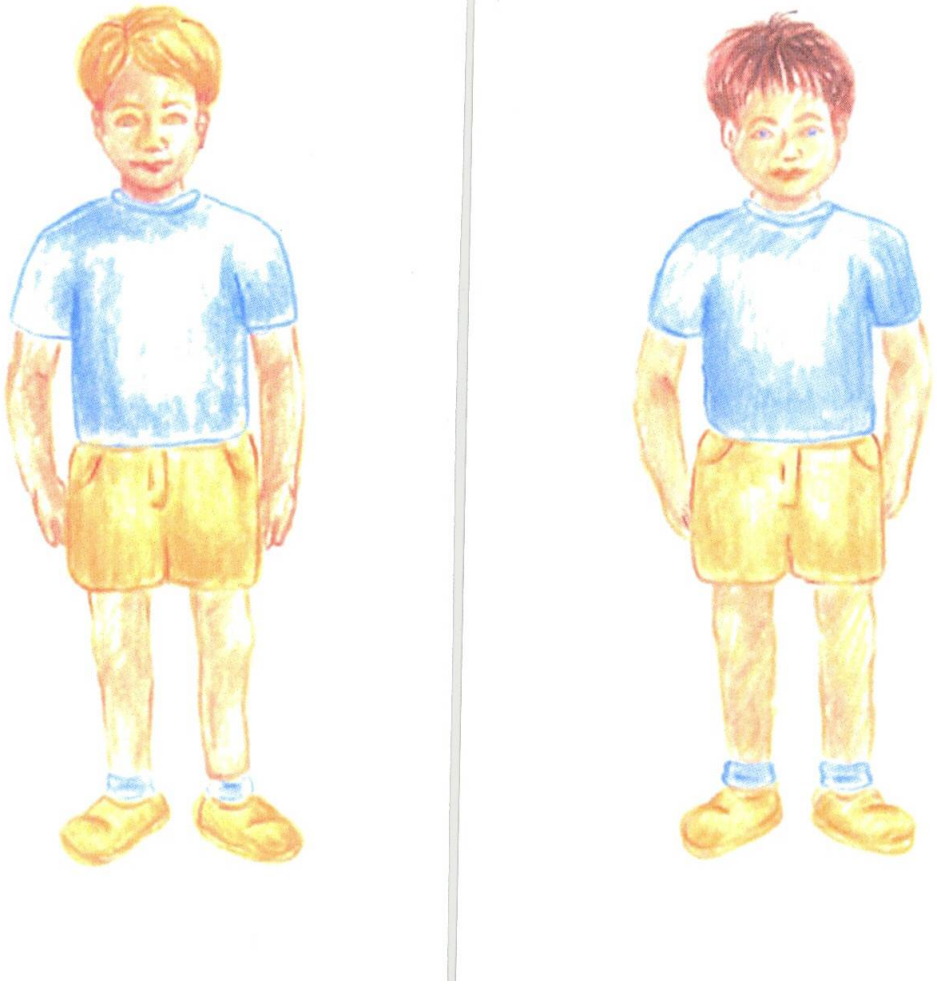


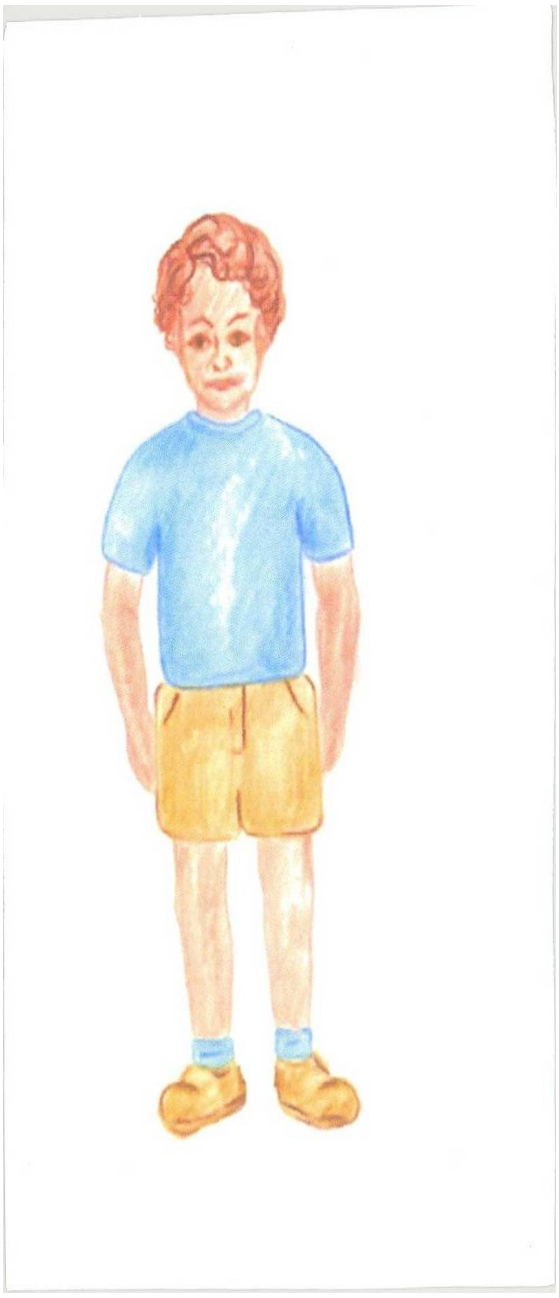




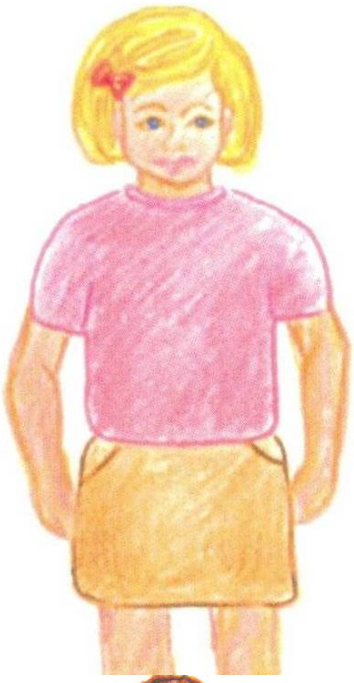


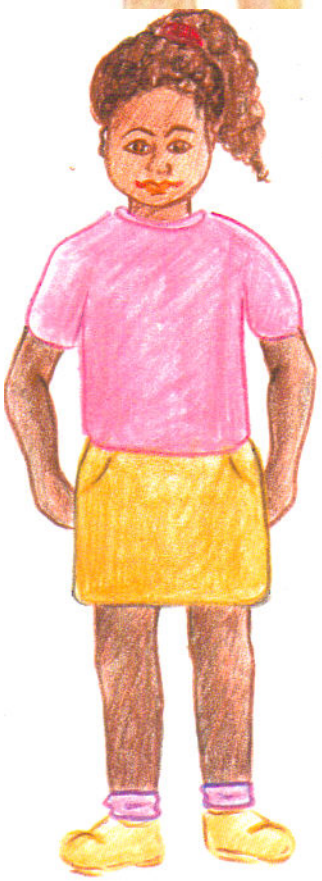


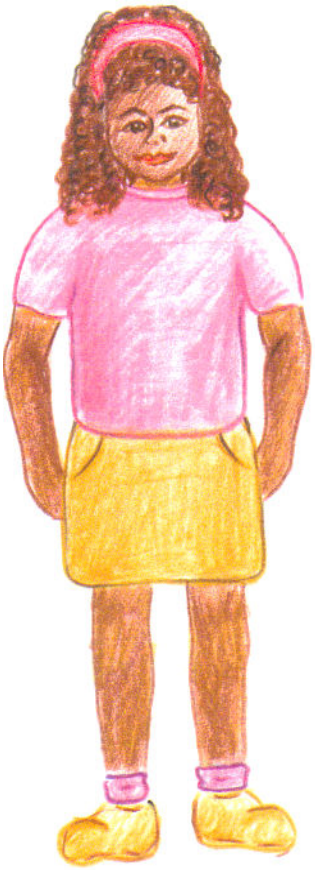




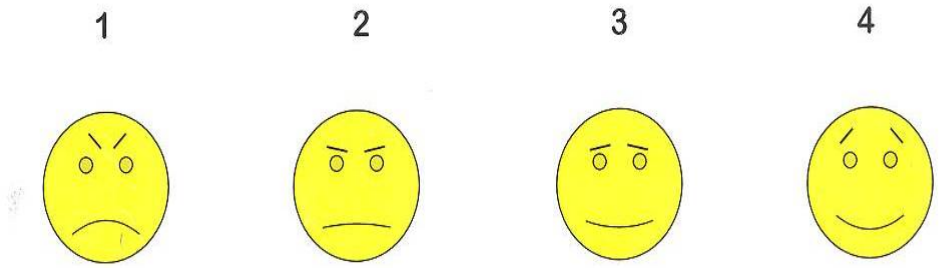
Female characters (randomly chosen by The research assistant for use throughout the interview)





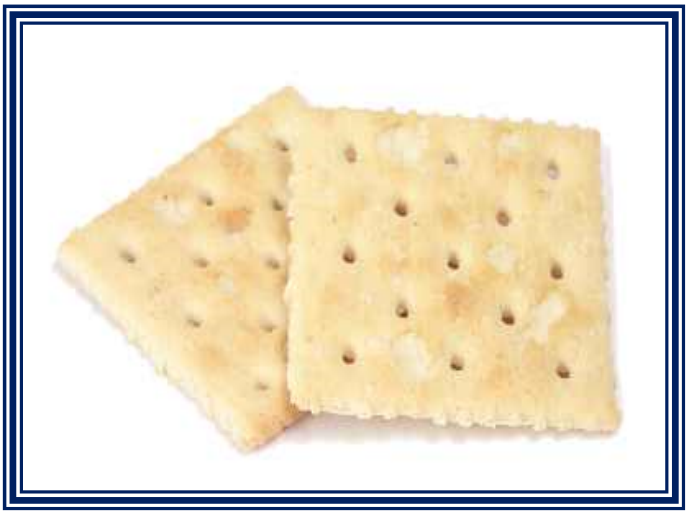


Likert Scale for Interview



Doll, Truck, and Backpacks (Intergroup Attitudes Attribution of Intentions Tasks, story 1 and 2)

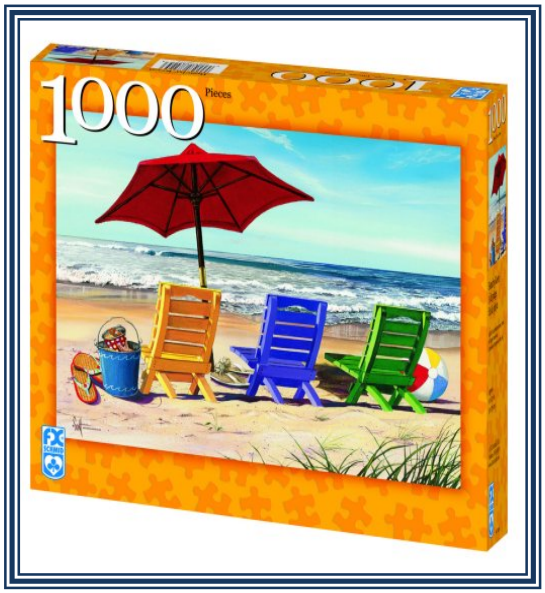
Crayons and Crackers (Theory of Mind Task 1, false belief, false contents)



Drawing table, markers, and cabinet (Theory of Mind Task 2, false belief location change)

Toys for Gender Stereotype Task (order randomly chosen for each interview)
Female Gender-Stereotyped Toys

Gender Neutral Toys



Male Gender-Stereotyped Toys

References

Antill, J. K., Goodnow, J., Russell, J., & Cotton, J. (1996). The influence of parents and family context on children's involvement in household tasks. [empirical]. *Sex Roles, 34*, 215-236.

Arthur, A. E., Bigler, R. S., Liben, L. S., Gelman, S. A., & Ruble, D. N. (2008). Gender stereotyping and prejudice in young children: A developmental intergroup pespective. In S. Levy & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 66-86). Oxford, U.K.: Oxford University Press.

Ashmore, R. D., & Del Boca, R. K. (1979). Sex stereotypes and implicit personality theory: Toward a cognitive-social psychological conceptualization. *Sex Roles, 5*, 219-248.

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

Baird, J. A., & Astington, J. W. (2004). The role of mental state understanding in the development of moral cognition and moral action. *New Directions for Child and Adolescent Development, 2004*(103), 37-49.

Banaji, M. R., Hardin, C., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology, 65*(2), 272-281. doi: 10.1037/0022-3514.65.2.272

Bellmore, A. D., Witkow, M. R., Graham, S., & Juvonen, J. (2004). Beyond the Individual: The Impact of Ethnic Context and Classroom Behavioral Norms

on Victims' Adjustment. *Developmental Psychology, 40*(6), 1159-1172. doi: 10.1037/0012-1649.40.6.1159

Bigler, R. S., Brown, C. S., & Markell, M. (2001). When groups are not created equal: Effects of group status on the formation of intergroup attitudes in children. *Child Development, 72*, 1151-1162.

Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology, 70*(6), 1142-1163. doi: 10.1037/0022-3514.70.6.1142

Boldizar, J. P., & Messick, D. M. (1988). Intergroup fairness biases: Is ours the fairer sex? *Social Justice Research, 2*, 95 - 111.

Booth, A. E., & Waxman, S. R. (2003). Mapping Words to the World in Infancy: Infants' Expectations for Count Nouns and Adjectives. [Article]. *Journal of Cognition & Development, 4*(3), 357-381. doi: No_doi

Brehl, B., & Wainryb, C. (2005). *Children's explanations of harmful behavior: Naive realism and interpretatioin*. University of Utah.

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86*, 307-324.

Brewer, M. B., & Silver, M. (1978). Ingroup bias a function of task characteristics. [Article]. *European Journal of Social Psychology, 8*(3), 393-400.

Brownell, C. A., Ramani, G. B., & Zerwas, S. (2006). Becoming a Social Partner With Peers: Cooperation and Social Understanding in One- and Two-Year-Olds. *Child Development, 77*(4), 803-821. doi: 10.1111/j.1467-8624.2006.00904.x

Brownell, C. A., Svetlova, M., & Nichols, S. (2009). To share or not to share: When do toddlers respond to another's needs? *Infancy, 14*(1), 117-130. doi: 10.1080/15250000802569868

Camerer, C. F. (2003). Strategizing in the Brain. [Article]. *Science, 300*(5626), 1673.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behaviour & Development, 21,* 315-330.

Carter, D. B., & Patterson, C. J. (1982). Sex roles as social conventions: The development of children's conceptions of sex-role stereotypes. *Developmental Psychology, 18,* 812-824.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*(1), 74-101. doi: 10.1037/0033-2909.115.1.74

Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms in reactive and proactive aggression. *Child Development 67,* 993-1002.

Diesendruck, G., Gelman, S. A., & Lebowitz, K. (1998). Conceptual and Linguistic Biases in Children's Word Learning. *Developmental Psychology, 34*(5), 823-839.

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology, 37*(5), 630-641. doi: 10.1037/0012-1649.37.5.630

Eichstedt, J. A., Serbin, L. A., Poulin-Dubois, D., & Sen, M. G. (2002). Of bears and men: Infants' knowledge of conventional and metaphorical gender stereotypes. [Article]. *Infant Behavior & Development, 25*(3), 296.

Foddy, M., Platow, M. J., & Yamagishi, T. (2009). Group-based trust in strangers: The role of stereotypes and expectations. *Psychological Science, 20*(4), 419-422. doi: 10.1111/j.1467-9280.2009.02312.x

Freeman, N. (2007). Preschoolers' Perceptions of Gender Appropriate Toys and their Parents' Beliefs About Genderized Behaviors: Miscommunication, Mixed Messages, or Hidden Truths? [Article]. *Early Childhood Education Journal, 34*(5), 357-366. doi: 10.1007/s10643-006-0123-x

Gelman, S. A., Taylor, M. G., Nguyen, S. P., Leaper, C., & Bigler, R. S. (2004). Mother-child conversations about gender: Understanding the acquisition of essentialist beliefs: VI. General discussion. *Monographs of the Society for Research in Child Development, 69*(1), 93-118. doi: 10.1111/j.0037-976X.2004.00279.x

Gopnik, A., & Astington, J. W. (1988). Children's Understanding of Representational Change and Its Relation to the Understanding of False Belief and the Appearance-Reality Distinction. [Article]. *Child Development, 59*(1), 26-37. doi: 10.1111/1467-8624.ep8860955

Graham, S. (2006). Peer victimization in school: Exploring the ethnic context. *Current Directions in Psychological Science, 15*(6), 317-321.

Harris, P. L., & Leevers, H. J. (2000). Reasoning from false premises. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind.* (pp. 67-86). Hove England: Psychology Press/Taylor & Francis (UK).

Harris, P. L., & Nunez, M. (1996). Understanding of Permission Rules by Preschool Children. [Article]. *Child Development, 67*(4), 1572-1591. doi: 10.1111/1467-8624.ep9704041034

Helwig, C. C., Tisak, M.S., & Turiel, E. (1990). Children's social reasoning and context. *Child Development, 61*, 2060-2078.

Helwig, C. C., Zelazo, P. D., & Wilson, M. (2001). Children's judgments of psychological harm in normal and noncanonical situations. *Child Development, 72*(1), 66-81.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., et al. (2005). 'Economic man' in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences, 28*(6), 795-855.

Heyman, G. D., & Gelman, S. A. (1999). The Use of Trait Labels in Making Psychological Inferences. [Article]. *Child Development, 70*(3), 604.

Heyman, G. D., & Gelman, S. A. (2000). Preschool children's use of novel predicates to make inductive inferences about people. *Cognitive Development, 15*(3), 263-280. doi: 10.1016/s0885-2014(00)00028-9

Horn, S. (2003). Adolescents' reasoning about exclusion from social groups. *Developmental Psychology, 39*, 71-84.

Horn, S., Killen, M., & Stangor, C. (1999). The influence of group stereotypes on adolescents' moral reasoning. *Journal of Early Adolescence, 19*, 98-113.

Inhelder, B., & Piaget, J. (1958). [The growth of logical thinking: From childhood to adolescence].

Jones, E. F., & Nelson-Le Gall, S. (1995). The Influence of Personal Effort Cues on Children's Judgments of Morality and Disposition. *Merrill-Palmer Quarterly, 41*(1), 53-69.

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system justification and the production of false consciousness. *British Journal of Social Psychology, 33*, 1-27.

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A Decade of System Justification Theory: Accumulated Evidence of Conscious and Unconscious Bolstering of the Status Quo. *Political Psychology, 25*(6), 881-920. doi: 10.1111/j.1467-9221.2004.00402.x

Juvonen, J., Graham, S., & Schuster, M. A. (2003). Bullying among young adolescents: The strong, the weak, and the troubled. *Pediatrics, 112*, 1231-1237.

Juvonen, J., Nishina, A., & Graham, S. (2006). Ethnic diversity and perceptions of safety in urban middle schools. *Psychological Science, 17*, 394-401.

Kalish, C. W. (2006). Integrating Normative and Psychological Knowledge: What Should We Be Thinking About? [Article]. *Journal of Cognition & Culture, 6*(1/2), 191-208. doi: 10.1163/156853706776931277

Karniol, R. (1987). Not all failures are alike: Self attribution and perception of teachers' attributions for failing tests in liked versus disliked subjects. *British Journal of Educational Psychology, 57*(1), 21-25.

Karniol, R., & Aida, A. (1997). Judging toy breakers: Gender stereotypes have devious effects on children. *Sex Roles, 36*(3), 195-205.

Karpinski, A. C., & Scullin, M. H. (2009). Suggestibility under pressure: Theory of mind, executive function, and suggestibility in preschoolers. *Journal of Applied Developmental Psychology, 30*(6), 749-763. doi: 10.1016/j.appdev.2009.05.004

Killen, M., Henning, A., Kelly, M. C., Crystal, D., & Ruck, M. (2007). Evaluations of interracial peer encounters by majority and minority U.S. children and adolescents. *International Journal of Behavioral Development, 31*, 491-500.

Killen, M., Kelly, M., Richardson, C., Crystal, D., & Ruck, M. (2009). *European-American children's and adolescents' evaluations of interracial exclusion* University of Maryland.

Killen, M., Kelly, M. C., Richardson, C., & Jampol, N. (in press). Attributions of intentions and fairness judgments regarding interracial peer encounters. *Developmental Psychology*.

Killen, M., Lee-Kim, J., McGlothlin, H., & Stangor, C. (2002). How children and adolescents evaluate gender and racial exclusion *Monographs of the Society for Research in Child Development* (Vol. 67, No. 4). Oxford, England: Blackwell Publishers.

Killen, M., Margie, N. G., & Sinno, S. (2006). Morality in the context of intergroup relationships. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 155-183). Mahwah, NJ: Lawrence Erlbaum Associates.

Killen, M., McGlothlin, H., & Lee-Kim, J. (2002a). Between individuals and culture: Individuals' evaluations of exclusion from social groups. In H. Keller, Y. Poortinga & A. Schoelmerich (Eds.), *Between biology and culture: Perspectives on ontogenetic development* (pp. 159-190). Cambridge, England: Cambridge University Press.

Killen, M., McGlothlin, H., & Lee-Kim, J. (2002b). *Exclusion on the basis of gender and race: A personal choice or a moral imperative?* Paper presented at the at the Annual Symposium of the Jean Piaget Society, Philadelphia, PA.

Killen, M., Mulvey, K. L., Richardson, C., Jampol, N., & Woodward, A. (under review). The accidental transgressor: Testing theory of mind and morality knowledge in young children.

Killen, M., Pisacane, K., Lee-Kim, J., & Ardila-Rey, A. (2001). Fairness or stereotypes? Young children's priorities when evaluating group exclusion and inclusion. *Developmental Psychology, 37,* 587-596.

Killen, M., Richardson, C., & Kelly, M. C. (2010). Developmental intergroup attitudes: Stereotyping, exclusion, fairness, and justice. In J. F. Dovidio, M. Hewstone, P. Glick & V. M. Esses (Eds.), *Handbook of Prejudice and Discrimination*. Thousand Oaks, CA: Sage Publishers.

Killen, M., Sinno, S., & Margie, N. G. (2007). Children's experiences and judgments about group exclusion and inclusion. In R. V. Kail (Ed.), *Advances in Child Development and Behavior* (pp. 173-218). New York: Elseveir.

Killen, M., & Smetana, J. G. (1999). Social Interactions in Preschool Classrooms and the Development of Young Children's Conceptions. [Article]. *Child Development, 70*(2), 486.

Killen, M., & Stangor, C. (2001). Children's social reasoning about inclusion and exclusion in gender and race peer group contexts. *Child Development, 72*, 174-186.

Kiyonari, T. (2002). Expectations of a generalized exchange system and ingroup favoritism: An experimental study of bounded reciprocity. *Japanese Journal of Psychology, 73*(1), 1-9.

Lagattuta, K. H. (2005). When You Shouldn't Do What You Want to Do: Young Children's Understanding of Desires, Rules, and Emotions. *Child Development, 76*(3), 713-733.

Leinbach, M. D., & Fagot, B. I. (1986). Acquisition of Gender Labels: A Test for Toddlers. [Article]. *Sex Roles, 15*(11/12), 655-666.

Leslie, A., Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect: Theory of mind and moral judgment. *Psychological Science, 17*, 421-427.

Leslie, A. M. (2000). *'Theory of mind' as a mechanism of selective attention* (2nd ed.). Cambridge, MA: MIT Press.

Liben, L. S., & Bigler, R. S. (2002). The developmental course of gender differentiation: Conceptualizing, measuring and evaluating constructs and pathways. *Monographs of the Society for Research in Child Development*(Serial No. 269, Vol. 67(2)).

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

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12- and 18-Month-Olds Point to Provide Information for Others. [Article]. *Journal of Cognition & Development, 7*(2), 173-187. doi: 10.1207/s15327647jcd0702_2

Malle, B. F., Moses, L. J., & Baldwin, D. A. (2001). *Intentions and intentionality: Foundations of social cognition*. Cambridge, MA US: The MIT Press.

Martin, C. L. (1989). Children's use of gender-related information in making social judgments. *Developmental Psychology, 25*(1), 80-88. doi: 10.1037/0012-1649.25.1.80

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

Martin, C. L., Fabes, R. A., Evans, S. M., & Wyman, H. (1999). Social cognition on the playground: Children's beliefs about playing with girls versus boys and their relations to sex segregated play. [Article]. *Journal of Social & Personal Relationships, 16*(6), 751.

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

Martin, C. L., Wood, C. H., & Little, J. K. (1990). The development of gender stereotype components. [Article]. *Child Development, 61*(6), 1891. doi: 10.1111/1467-8624.ep9103040647

McAllister, D. J. (1995). Affect - and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal, 38*, 24 - 59.

McGlothlin, H., & Killen, M. (2006). *Intergroup attitudes of children attending ethnically heterogeneous and homogeneous schools.* manuscript, University of Maryland.

Nelson-le Gall, S. A. (1985). Motive–outcome matching and outcome foreseeability: Effects on attribution of intentionality and moral judgments. *Developmental Psychology, 21*(2), 332-337. doi: 10.1037/0012-1649.21.2.332

Nichols, S. R., Svetlova, M., & Brownell, C. A. (2010). Toddlers' Understanding of Peers' Emotions. [Article]. *Journal of Genetic Psychology, 171*(1), 35-53.

Nucci, L., & Weber, E. K. (1995). Social Interactions in the Home and the Development of Young Children's Conceptions of the Personal. *Child Development, 66*(5), 1438-1452.

Nucci, L. P. (2001). *Education in the moral domain*. Cambridge, England: Cambridge University Press.

Nucci, L. P., & Turiel, E. (1978). Social interactions in the context of moral and conventional transgressions. *Child Development, 49*, 400-407.

Núñez, M., & Harris, P. L. (1998). Psychological and Deontic Concepts: Separate Domains or Intimate Connection? [Article]. *Mind & Language, 13*(2), 153.

Patterson, M. M., & Bigler, R. S. (2006). Preschool children's attention to environmental messages about groups: Social categorization and the origins of intergroup bias. *Child Development, 77*, 847 - 860.

Peets, K., Hodges, R., Kikas, E., & Salmivalli, C. (2007). Hostile attributions and behavioral strategies in children: Does relationship type matter? *Developmental Psychology, 43*, 889-900.

Phillips, W., Baron-Cohen, S., & Rutter, M. (1998). Understanding intention in normal development and in autism. *British Journal of Developmental Psychology, 16*(3), 337-348.

Piaget, J. (1932/1965). *The moral judgment of the child*. New York: Free Press.

Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.

Pillow, B. H., & Mash, C. (1999). Young children's understanding of interpretation, expectation and direct perception as sources of false belief. *British Journal of Developmental Psychology, 17*(2), 263-276. doi: 10.1348/026151099165267

Platow, M. J., McClintock, C. G., & Liebrand, W. B. (1990). Predicting intergroup fairness and ingroup bias in the minimal group paradigm. *European Journal of Social Psychology, 20*(3), 221-239.

Poulin, D. D., Serbin, L. A., & Derbyshire, A. (1998). Toddlers' intermodal and verbal knowledge about gender. *Merrill-Palmer Quarterly, 44*(3), 338-354.

Powlishta, K. K. (1995). Gender bias in children's perceptions of personality traits. [Article]. *Sex Roles, 32*(1/2), 17-28.

Rowley, S., Kurtz-Costes, B., Mistry, R., & Leagans, L. (2007). Social status as a predictor of race and gender stereotypes in childhood and early adolescence. *Social Development, 16*, 151 - 168.

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

Rudman, L. A., & Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin, 26*(11), 1315-1328. doi: 10.1177/0146167200263001

Schult, C. A. (2002). Children's Understanding of the Distinction between Intentions and Desires. [Article]. *Child Development, 73*(6), 1727-1747. doi: 10.1111/1467-8624.t01-1-00502

Shepard, W. O., & Hess, D. T. (1975). Attitudes in Four Age Groups Toward Sex Role Division in Adult Occupations and Activities. *Journal of Vocational Behavior*.

Sherif, M. (1966). *Group conflict and cooperation: Their social psychology*. London: Routledge.

Shultz, T. R., Wright, K., & Schleifer, M. (1986). Assignment of Moral Responsibility and Punishment. [Article]. *Child Development, 57*(1), 177-184. doi: 10.1111/1467-8624.ep7251035

Siegal, M., & Peterson, C. C. (1998). Preschoolers' Understanding of Lies and Innocent and Negligent Mistakes. *Developmental Psychology, 34*(2), 332-341.

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*(2), 147-183. doi: 10.1006/drev.1993.1007

Smetana, J. G. (1981). Preschool children's conceptions of moral and social rules. *Child Development, 52*, 1333-1336.

Smetana, J. G. (1983). Social-cognitive development: Domain distinctions and coordinations. *Developmental Review, 3*, 131-147.

Smetana, J. G. (1986). Preschool children's conceptions of sex-role transgressions. *Child Development, 57*, 832-871.

Smetana, J. G. (1993). Understanding of social rules. In M. Bennett (Ed.), *The development of social cognition: The child as psychologist*. New York: Guilford.

Smetana, J. G. (1995). Morality in context: Abstractions, ambiguities, and applications. In R. Vasta (Ed.), *Annals of child development* (Vol. 10, pp. 83-130). London: Jessica Kinglsey Publishers.

Smetana, J. G. (2006). Social-cognitive domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 119-154). Mahwah, NJ: Lawrence Erlbaum Associates.

Smetana, J. G., Schlagman, N., & Adams, P. (1993). Preschoolers' judgments about hypothetical and actual transgressions. *Child Development, 64*, 202-214.

Stangor, C., & Schaller, M. (1996). Stereotypes as individual and collective representations. In C. N. Macrae, M. Hewstone & C. Stangor (Eds.), *Foundations of stereotypes and stereotyping* (pp. 3-37). New York: Guilford Press.

Stoddart, T., & Turiel, E. (1985). Children's concepts of cross-gender activities. *Child Development, 56*, 1241-1252.

Tajfel, H. (1981). *Human groups and social categories.* Cambridge, UK: Cambridge University Press.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology, 1*(2), 149-178. doi: 10.1002/ejsp.2420010202

Tajfel, H., & Turner, J. (2001). An integrative theory of intergroup conflict. In M. A. Hogg & D. Abrams (Eds.), *Intergroup relations: Essential readings.* (pp. 94-109). New York, NY US: Psychology Press.

Takagishi, H., Kameshima, S., Schug, J., Koizumi, M., & Yamagishi, T. (2010). Theory of Mind Enhances Preference for Fairness. *Journal of Experimental Child Psychology, 105*(1-2), 130-137.

Terwogt, M. M., & Rieffe, C. (2003). Stereotyped beliefs about desirability: Implications for characterizing the child's theory of mind. *New Ideas in Psychology, 21*(1), 69-84. doi: 10.1016/s0732-118x(02)00031-4

Theimer, C. E., Killen, M., & Stangor, C. (2001). Young children's evaluations of exclusion in gender-stereotypic peer contexts. *Developmental Psychology, 37*, 18-27.

Tisak, M. S. (1995). Domains of social reasoning and beyond. In R. Vasta (Ed.), *Annals of child development* (Vol. 11, pp. 95-130). London: Jessica Kingsley Publishers.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The ontogeny and phylogeny of cultural cognition. *Behavioural & Brain Sciences, 28*, 675-691.

Trautner, H. M., Ruble, D. N., Cyphers, L., Kirsten, B., Behrendt, R., & Hartmann, P. (2005). Rigidity and flexibility of gender stereotypes in childhood: developmental or differential? [Article]. *Infant & Child Development, 14*(4), 365-381. doi: 10.1002/icd.399

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, England: Cambridge University Press.

Turiel, E. (1998). The development of morality. In W. Damon (Ed.), *Handbook of child psychology* (5th ed., Vol. 3: Social, emotional, and personality development, pp. 863-932). New York: Wiley.

Turiel, E., Killen, M., & Helwig, C. C. (1987). Morality: Its structure, functions, and vagaries. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children* (pp. 155-243). Chicago: University of Chicago Press.

Wainryb, C. (2000). Values and truths: The making and judging of moral decisions. In M. Laupa (Ed.), *New directions for child and adolescent development: Rights and wrongs: How children and young adults evaluate the world, No. 89.* (pp. 33-46). San Francisco, CA US: Jossey-Bass.

Wainryb, C. (2004). "Is" and "ought": Moral judgments about the world as understood. *New Directions for Child and Adolescent Development, 2004*(103), 3-18.

Wainryb, C., & Brehl, B. A. (2006). I thought she knew that would hurt my feelings: Developing psychological knowledge and moral thinking. In R. V. Kail (Ed.), *Advances in child development and behavior (Vol 34).* (pp. 131-171). San Diego, CA US: Elsevier Academic Press.

Wainryb, C., & Ford, S. (1998). Young Children's Evaluations of Acts Based on Beliefs Different from Their Own. *Merrill-Palmer Quarterly, 44*(4), 484-503.

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative Activities in Young Children and Chimpanzees. [Article]. *Child Development, 77*(3), 640-663. doi: 10.1111/j.1467-8624.2006.00895.x

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science, 311*, 1301-1303.

Warneken, F., & Tomasello, M. (2009). The roots of human altrusim. *British Journal of Psychology, 100*, 455-471.

Weber, J. (1993). Exploring the Relationship Between Personal Values and Moral Reasoning. [Article]. *Human Relations, 46*(4), 435-464.

Weisner, T. S., & Wilson-Mitchell, J. E. (1990). Nonconventional family life-styles and sex typing in six-year-olds. *Child Development, 61*, 1915-1933.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief. [Article]. *Child Development, 72*(3), 655.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*, 502-517.

Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition, 35*(3), 245-275. doi: 10.1016/0010-0277(90)90024-e

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. [Article]. *Cognition, 13*(1), 103-128.

Woodward, A., & Needham, A. (2009). *Learning and the infant mind*. New York, NY: Oxford University Press.

Yamagishi, T., Jin, N., & Kiyonari, T. (1999). Bounded generalized reciprocity: In-group boasting and in-group favoritism. *Advances in Group Processes, 16,* 161 - 197.

Yamagishi, T., & Kiyonari, T. (2000). The group as the container of generalized reciprocity. *Social Psychology Quarterly, 63*(2), 116-132. doi: 10.2307/2695887

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology, 28*(1), 126-136. doi: 10.1037/0012-1649.28.1.126

Zelazo, P. D., Helwig, C. C., & Lau, A. (1996). Intention, act, and outcome in behavioral prediction and moral judgment. *Child Development, 67*, 2478-2492.