

ABSTRACT

Title of Dissertation: RESISTING STEREOTYPES: EVALUATIONS OF PEER
GROUP EXCLUSION

Kelly Lynn Sara Mulvey, Doctor of Philosophy, 2013

Dissertation directed by: Professor Melanie Killen, Department of Human Development
and Quantitative Methodology

Gender stereotypes permeate peer groups, often functioning as the norms, or conventions, which contribute to group identity. Little is known about the conditions under which children will resist the norms of their peer groups, including norms that reflect stereotypic expectations. This study investigated this issue by measuring how children responded to members of their gender groups who disagreed with the group about gender stereotyped aggressive behaviors (female stereotype: relational aggression, male stereotype: physical aggression) as well as about social activities (football and ballet). Social domain theory as well as social identity theory provided the basis for formulating the design and the hypotheses. It was expected that children and adolescents would expect their peers to challenge the group, but that they would be concerned about the consequences of challenging the group in terms of social exclusion. Participants ($N = 292$, 9-10 and 13-14 years of age) assessed members of same-gender peer groups who disagreed with their group. The findings revealed that children and adolescents generally expected that their peers will resist the group's gender stereotypic norms surrounding

aggression, but that this may be more difficult for boys when voicing their counter-stereotypic opinions. Further, participants themselves believed that they would be less influenced by gender stereotypes than would their peers. They asserted that they would, individually, be more likely to resist the group than they expect a peer would be. This research also revealed important barriers to resisting the group. Specifically, children and adolescents expected that group members who dissent from or resist the group are likely to be excluded from the group for voicing their dissent. This stands in sharp contrast to much research which indicates that children are not accepting of exclusion. Rejecting the behavior of one's peer group, especially when that behavior has negative intrinsic consequences for others, is a key step towards changing the culture of peer groups more broadly. However, the findings indicate that, while children and adolescents are optimistic about their peers challenging the peer group, they also see exclusion as a very real possibility and consequence for such resistance.

RESISTING STEREOTYPES: EVALUATIONS OF PEER GROUP EXCLUSION

by

Kelly Lynn Sara Mulvey

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

2013

Advisory Committee:

Professor Melanie Killen, Chair

Professor Geetha Ramani

Professor Kenneth Rubin

Professor Meredith Rowe

Professor Jude Cassidy, Dean's Representative

© Copyright by
Kelly Lynn Sara Mulvey
2013

Dedication

I dedicate this dissertation to my wonderful two children, Adalynn and Theodore, who have always been a source of inspiration and joy.

Acknowledgements

This research has come to fruition with the help and support of many people. First, the support of my amazing husband, Adam Hartstone-Rose, has been invaluable during this process. His consistent encouragement and guidance are, of course, wonderful, but I have to acknowledge as well his amazing artistic skills and thank him for the many hours he put into designing images for my studies. Finally, I must thank him for his Daddy-day Wednesdays at Chipotle with the kids, which enabled me to fully engage in our lab meetings, even if it meant not getting home until after the children were asleep for the night! Of course I must also acknowledge my ever willing in-home participants, my children, Adalynn and Theodore, who have been so supportive that they taught all their friends at school to play “research” instead of “house.” I also am ever grateful for the support of my parents, Brian and Joanne Mulvey, whose support for all things academic throughout my life helped me to get here and who frequently and without hesitation offered to help whenever needed. I also appreciate the support of my siblings, Kerrie Anne Loyd, Brian Mulvey and Brandon Mulvey, who always provide great conversation, advice, and diversions from the world of research.

This dissertation would not have been fathomable without the amazing support of my advisor, Dr. Melanie Killen. Melanie’s mentorship, guidance and faith in me have helped to ensure my success. I thank her, first, for taking a chance on a former English major who wanted to study children’s moral development. But moreover, I thank her for

the countless ways in which she has provided opportunities, support and encouragement throughout my doctoral career. I am so very lucky to have found a mentor who so respects her graduate students and dedicates so much energy to training them to be strong developmental scientists. I look forward to many years of future collaborations and I am very proud to have been able to work with and learn from Melanie.

I must also thank my fellow graduate students and scholars in the Social and Moral Development Laboratory including Alaina Brenick, Cameron Richardson, Megan Clark Kelly, Aline Hitti, Shelby Cooley, Laura Elenbaas, Michael Rizzo and Jeeyoung Noh. I am especially grateful to Aline who traversed graduate school, research, writing and commuting long distances to UMD side by side with me.

I also acknowledge the support and guidance of my dissertation committee, Drs. Jude Cassidy, Geetha Ramani, Meredith Rowe and Kenneth Rubin as well as other academic mentors here at the University of Maryland including Dr. Allan Wigfield and Dr. Judith Torney-Purta. Additionally, I would like to thank Dr. Adam Rutland and Dr. Dominic Abrams, who welcomed me into their collaborative projects with Melanie, even as a young graduate student, and taught me so much. My dissertation truly developed directly from the projects I conducted with Melanie, Aline, Adam and Dom.

Finally, I must acknowledge the students and my colleagues at Durham School of the Arts, whose amazing lives first inspired me to learn more about children's social and moral development. I am especially grateful for Sarah Corey, and Gerry and Ray Larson. In each project that I develop, I think of conversations and experiences I had while at DSA. Last, I must acknowledge Ro Thorne and Alan Teasley, my mentors in Duke's

Masters of Arts in Teaching program who helped sharpen my focus on social justice issues in children’s lives.

Table of Contents

Dedication	ii
Acknowledgements	iii
List of Figures	ix
List of Tables	x
List of Appendices	xi
Chapter I.....	1
Introduction.....	1
Theoretical Rationale	6
Social Domain Theory	7
Developmental Subjective Group Dynamics.....	9
Study Rationale.....	10
Challenging the Group: Morality and Group Identity	10
Types of Gender Stereotypes	12
Evaluating Resistance to Peer Group Behavior	16
Repercussions for Resistance: Exclusion?.....	18
Social-Cognitive Skills	21
Current Study: Design and Hypotheses	22
Study Design.....	22
Study Hypotheses.....	24
Chapter II	30

Literature Review.....	30
Overview of the goals of the paper.....	33
Research on gender exclusion in childhood.....	33
Evaluating exclusion: The role of social-cognitive abilities.....	35
A Framework for Studying Exclusion: Integrating Social Domain Theory, Social Identity Theory and Examinations of Theory of Mind.....	36
Social Domain Theory	37
Social Identity Theory and Subjective Group Dynamics	39
Intentionality, Theory of Mind and the Psychological States of Others.....	40
Gender Stereotypes	42
Stereotypes and Exclusion	44
Children’s and Adolescents’ Reasoning About Exclusion Because of Group Membership	45
Is Exclusion Always About Group Membership?	52
Research on Conflating Personality Traits with Group Membership.....	57
The Role of Group Membership, Norms and Identity in Exclusion Decisions	59
Judging Others’ Intentions.....	65
Intentionality, Theory of Mind and Social Relationships.....	68
Conclusions.....	76
Chapter III.....	79
Methodology	79
Participants.....	79
Design	79
Procedure	81
Measures	81

Gender Exclusion Task.....	82
Interpretative Theory of Mind Task.....	85
Gender Stereotype Task.....	88
Plan for Analysis.....	88
Results.....	89
Gender Stereotypic Activities.....	89
Stereotype Measure.....	89
Individual Resistance to Group Norms about Gender Stereotypic Activities	91
Relationship between Peer and Individual Resistance to Group Norms about Gender Stereotypic Activities.....	93
Group Favorability Towards the Deviant Group Member	94
Justifications for Group Favorability Towards the Deviant Member.....	97
Individual Favorability Towards the Deviant Member	100
Justifications for Individual Favorability Towards the Deviant Member.....	101
Relation between Group and Individual Favorability Towards Deviant Members in the Context of Gender Stereotypic Activities.....	104
Likelihood of Exclusion of the Deviant Member	104
Justifications for Likelihood of Exclusion of the Deviant Member	106
Inclusion Decisions.....	108
Gendered Forms of Aggression	108
Stereotype Measure.....	108
Relation between Peer and Individual Resistance to Group Norms about Aggression	109
Group Favorability Towards the Deviant Group Member	110
Justifications for Group Favorability Towards the Deviant Member.....	112
Individual Favorability Towards the Deviant Member	114

Justifications for Individual Favorability Towards the Deviant Member.....	117
Group Favorability Versus Individual Favorability.....	119
Likelihood of Exclusion.....	120
Justifications for Likelihood of Exclusion.....	122
Likelihood of Exclusion by Group Favorability Evaluation.....	124
Inclusion Choice	125
Justifications for Inclusion Choice.....	128
Interpretative Theory of Mind Descriptive Statistics.....	130
Interpretative Theory of Mind and Individual and Group Favorability.....	130
Discussion.....	132
Gender Stereotypic Activities	132
Gendered Forms of Aggression	138
Conclusions.....	146
Contributions to the Literature.....	146
Limitations and Future Directions	148
Implications.....	150
References.....	204

List of Figures

Figure 1: Task Design.....	151
Figure 2: Likelihood of Exclusion of the Deviant Member.....	156
Figure 3: Deviant who Wants to do Ballet: Reasoning by Exclusion Likelihood.....	157
Figure 4: Likelihood of Resistance Versus Individual Likelihood of Resistance to Group Norms about Aggression.....	159
Figure 5 Group Favorability Versus Individual Favorability	162
Figure 6 Likelihood of Exclusion	163
Figure 7 Likelihood of Exclusion by Age Group	164
Figure 8: Likelihood of Exclusion by Dichotomous Group Favorability Evaluation	166
Figure 9: Inclusion of Outgroup Member Who Matches the Group Norm	166

List of Tables

Table 1: Summary of Hypotheses and Analyses	152
Table 2: Percentage of Participants Affirming Stereotypes Regarding Who Usually Engages in Each Type of Activity or Behavior	155
Table 3: Means and Standard Deviations for Proportion of Justifications Used for Evaluations of Group Favorability Toward the Deviant Member	160
Table 4: Means and Standard Deviations for Proportion of Justifications Used for Evaluations of Individual Favorability Toward the Deviant Member	161
Table 5: Means and Standard Deviations for Proportion of Justifications Used for Evaluations of Likelihood of Exclusion of the Deviant Member	165
Table 6: Means and Standard Deviations for Proportion of Justifications Used for Inclusion Choice	168

List of Appendices

Appendix A: Institutional Review Board Approval and Consent Forms	169
Appendix B: Sample Tasks.....	181
Appendix C: Coding System	201

Chapter I

Introduction

Children form groups from an early age, and decisions about group membership often create conflict among friends. One way to better understand children's conflicts is to study the decision-making process behind their group interactions. When do children think it is okay to be different from the group? When do children believe that it is okay to let the group know that they want to change the activity of the group? The overall focus of this research is on how children evaluate peer group decisions to include and exclude others. The contexts examined for this study were those associated with gender stereotypic expectations about aggressive behavior. Specifically, the research question focused on the legitimacy of excluding someone from a group based on gender, and when children expected groups to condone different forms of aggressive behavior due to the gender stereotypes surrounding different types of aggression. Gender-associated aggressive behavior refers to behavior that is associated (perhaps only through stereotypes) with boys, such as physical aggression, and with girls, such as relational aggression (e.g., gossiping, teasing, with negative intentions) (see Card, Stucky, Sawalani, & Little, 2008; Crick & Grotpeter, 1995; Ostrov & Godleski, 2010 for a summary of research on gender differences in aggression). The empirical question centered on when children view these forms of aggression to be condoned in the context

of peer group dynamics, and when children viewed it as legitimate to challenge the group norms that perpetuate aggressive behavior (relational or physical).

A further consideration is how children will respond to members of their gender group who disagree with or dissent from the group based on gender stereotypic expectations. Gender-associated forms of aggression and gender-stereotypic activities were included for analysis in this project and were: 1) gendered forms of aggression, physical (male) and relational (female) ; and 2) gendered types of activities, football (male) and ballet (female).

It is important to note that the research literature has recently shown that, across many cultural contexts and in many studies, in fact, boys engage in more physical aggression than do girls, but that there are no differences in mean levels of relational aggression between boys and girls (Card et al., 2008; Lansford et al., 2012). However, the stereotype indicating that girls engage in more relational aggression and boys engage in more physical aggression persists. For instance, even the most recent research on gender and aggression often relies upon these stereotypes as a foundation behind the research design (Kochel, Miller, Updegraff, Ladd, & Kochenderfer-Ladd, 2012). This continued reliance upon gender stereotypic conceptions of aggression is problematic and this research aims to examine if children actually do perceive these behaviors as stereotypic and if they use such stereotypes in evaluating their peers.

Thus, the aim of this study is to examine how children react to changes within the group, specifically when group members challenge the stereotypic expectations of their group. The findings from this project will help to provide information that can guide teachers and counselors in establishing classroom and school environments where

children feel comfortable in their group interactions, even when this involves challenging or resisting stereotypes. This is a critically important research area, as confidence in challenging group behaviors, which are unacceptable because they are founded upon stereotypes or exhibit prejudicial or biased attitudes, or group behaviors, which cause harm to others, is the first step towards child-led forms of social justice where children influence their peers to change group behaviors.

This study draws from theories and methodologies in social development which have investigated children's and adolescents' evaluations of social exclusion (Killen & Rutland, 2011), specifically social domain theory (Killen, 2007; Turiel, 2006) and developmental subjective group dynamics (Abrams & Rutland, 2008). This study will extend previous research by examining resistance to norms which also involve widely held stereotypes. Previously, researchers have shown that children view deviating from a group as acceptable when the group distributes resources unequally or breaks social conventions about dress codes (Killen, Rutland, Abrams, Mulvey, & Hitti, 2012), but no research has examined what children think about the costs of resisting group norms, or whether it is okay to challenge a group when the norms are about gender appropriate or gender stereotypic expectations. The general aim of this research is to examine whether children consider resistance to gender-stereotypic behaviors to be grounds for exclusion from a group. The goals are to: 1) to examine evaluations of exclusion from groups which hold norms that are related to the gender identity of the group, and 2) to examine the role of social-cognitive skills, such as theory of mind, in evaluations of resisting stereotypes and gender exclusion.

In order to understand how children respond to resistance to peer group norms, or behaviors, this study employed hypothetical scenarios involving groups that adhere to and resist gender stereotypes. Participants made judgments about members of these groups who disagree with their group in the context of physical aggression, relational aggression, and gender stereotypic activities (football and ballet). As an example, do children think it is okay for a girl to tell her girls' group that gossiping (relational aggression) is not okay when the group has been known to gossip? Or a boy to tell his boys' group that pushing and shoving (physical aggression) is not okay when the group has been known to act rough? These questions are important because children encounter these types of situations, and how children construe the norms of groups in these contexts is not known. To compare participants' evaluations of different "gendered" forms of aggression with other activities associated with gender, evaluations of resistance regarding general social activities associated with gender were also studied in this project (e.g., football and ballet). Thus, gender stereotypic activities surround choices about activities for social interactions, while physical and relational aggression involve psychological decisions about behaviors that are potentially harmful to others. While some research has examined children's reactions to peers who engage in physical aggression, relational aggression and social conventional violations (cheering for a different team), this research only assessed participants' favorability towards these peers and did so in gender neutral contexts (Atkin & Gummerum, 2012).

What factors contribute to children's judgments about resisting group norms, particularly in contexts in which the norms may be wrong from a moral perspective (e.g., inflicting harm on others)? One aspect of children's evaluations which has received

attention recently has to do with theory of mind, which is a form of perspective taking. In fact, research indicates that theory of mind abilities are related to exclusion judgments (Abrams, Rutland, Pelletier, & Ferrell, 2009), as well as to understanding of gender discrimination (Brown, Bigler, & Chu, 2010) and moral transgressions (Killen, Mulvey, Richardson, Jampol, & Woodward, 2011). Additionally, some research indicates that holding gender stereotypes may impact one's ability to take others' perspectives (Kelly, Mulvey, Hitti, Moin, & Killen, May 2011; Terwogt & Rieffe, 2003)

Understanding the relation between theory of mind competence and evaluations of peer group interactions is important, as theory of mind is developing at the same time as children are becoming more experienced in group contexts. Little is known, however, about how theory of mind may impact how children perceive peers who challenge group norms. Are children with greater or more sophisticated theory of mind skills better able to understand the perspective of the child who disagrees with the group? Can they also understand the group's perspective? If children, themselves, hold gender stereotypic expectations will this impair their ability to understand others' perspectives?

Understanding more about the relation between theory of mind competence and evaluations of peers who challenge stereotypic group norms will provide insight into how teachers, parents and group leaders can better guide children towards successful peer relationships. Thus, assessments of theory of mind will be compared to participants' exclusion judgments in order to determine the relation between understanding group dynamics surrounding gender stereotypes and social-cognitive skills.

In sum, the first aim of this research is to better understand evaluations of social exclusion in the context of complex peer group interactions where stereotypes come into

play, group norms are clearly delineated, and resistance to these norms is experimentally manipulated. Additionally, the second aim of this research is to identify relations between social-cognitive judgments in peer relationships and social-cognitive skills.

Theoretical Rationale

As children engage with peer groups, they frequently encounter situations in which they must balance information about their social worlds and make moral decisions. Very early in childhood, concepts of fairness and equality emerge, enabling even children as young as 2 ½ to 3 years of age to make moral judgments in social encounters. At the same time, children are able to distinguish issues which involve rights, justice, welfare and harm from those which involve conventions, traditions or customs, and those which involve personal preferences (Nucci, 1981; Smetana, 2006; Turiel, 1983, 2006). Children's differentiated understanding of social issues is reflected in their reasoning about moral and social decisions which they make.

While children are developing concepts of fairness, they also develop a sense of group identity, and stereotypes and prejudicial attitudes about others unlike themselves (Rutland, Killen, & Abrams, 2010). Quite early, children begin to understand group dynamics, and the complex interplay between distinct social groups (Abrams & Rutland, 2008). From as early as 6 years of age, children can and do differentiate between ingroup and outgroup members who either adhere to or resist the norms, or conventions, practices and beliefs of their groups. Finally, during this time period, children's social cognitive skills, such as theory of mind (Wellman & Liu, 2004), are increasing, as is their reliance upon stereotypes, including those about gender (Liben & Bigler, 2002; Mulvey, Hitti, & Killen, 2010).

Two distinct, yet complementary theoretical frameworks drive much of our understanding of how children balance their complex social worlds: social domain theory (SDT), which distinguishes between the societal, moral, and psychological domains of knowledge used in making judgments (Turiel, 1983), and developmental subjective group dynamics (DSGD), which argues that children differentiate distinctly between ingroup and outgroup members who express deviant and normative views (Abrams, Rutland, & Cameron, 2003). Both frameworks include a recognition of the growing social-cognitive competencies of children, as well as the stereotypes and biases which they may bring to social interactions.

Social Domain Theory

Social domain theory provides a key framework for examining children's social interactions and moral decision-making. Specifically, domain theory identifies three domains children use in making judgments: moral (fairness, justice, rights and welfare), societal (customs, conventions, group functioning, traditions) and psychological (personal preference, autonomy, intentions and understanding mental states) (Turiel, 1983). Over 30 years of research indicates that, in many situations, individuals clearly and systematically identify a single domain in their reasoning (for a review, see Smetana, 2006). These are considered straightforward or prototypic scenarios. As an example, children overwhelmingly identify hitting another person as wrong and justify this evaluation by referencing the intrinsic negative consequences of hitting (harm to another), classifying the event as a moral transgression.

Whereas many events which children encounter are prototypic, and thus, easily identified as falling within one particular domain, increasingly, children also encounter

multifaceted situations, where children must weigh competing concerns from multiple domains in making an evaluation or judgment. For instance, stereotypes are an example of a multifaceted issue: children may recognize the harmful or unfair nature of holding a stereotype about someone else (moral domain), while also acknowledging that some stereotypes may enable children to distinguish their group from others and to define their group's customs (societal domain). When individuals evaluate complex, multifaceted acts and issues, they must balance and weigh different considerations, reasoning, at times, using multiple domains.

As early as age three, children clearly differentiate between moral acts, as those involving harm to others; societal acts, as those surrounding customs and traditions; and psychological acts as those which should not be regulated by others (Nucci, 1981; Smetana, 1983; Turiel, 1983). Early in development, children clearly use moral reasoning when addressing issues involving physical harm. Soon after, they also use moral reasoning for psychological harm. And, by middle childhood or early adolescence, issues of group functioning begin to play an important role in how children conceptualize acts (Horn, 2003). While initial research using social domain theory focused on prototypic issues, in the past decade researchers have begun to examine more complex social interactions, including those involving intergroup components (Killen, Margie, & Sinno, 2006).

The recent focus on intergroup relations by social domain researchers has involved, in particular, children's assessments of inclusion and exclusion decisions (Killen, 2007). This research provides a particularly effective means of understanding the intersection between children's conceptions of intergroup relations, stereotypes and

biases, and the moral, societal and psychological domains. Research on social exclusion has occurred around the world, finding that children, at times, view exclusion as acceptable, especially because of societal reasons such as traditions or group functioning, and at times, unacceptable, particularly when exclusion is based on group membership, such as gender or ethnicity (Hitti, Mulvey, & Killen, 2011).

Developmental Subjective Group Dynamics

Developmental subjective group dynamics (DSGD) focuses on group identity and the relation between judgments about members of one's ingroup and judgments about ingroups and outgroups generally (Abrams, Rutland, & Cameron, 2003; Abrams, Rutland, Cameron, & Marques, 2003). Specifically, DSGD examines children's evaluations of ingroup and outgroup members who express deviant and normative attitudes towards the group (Abrams & Rutland, 2008). For example research in this field has examined group norms involving loyalty to the group, such as cheering for one's ingroup sports team. Research has found that children prefer individuals who express normative, or loyal, ideas about the ingroup, regardless of their group affiliation and likewise express greater dislike for individuals who deviate from the group norms (Abrams, Rutland, Cameron, & Ferrell, 2007; Abrams, Rutland, Cameron, et al., 2003; Abrams, Rutland, Ferrell, & Pelletier, 2008; Abrams et al., 2009). Additionally, research on DSGD has shown that when evaluating moral or immoral peers versus loyal or disloyal peers, children use intergroup bias for group based evaluations (about loyalty), but not morality based evaluations (Abrams, Rutland, et al., 2008). DSGD has also revealed that older children adhere more to group norms than do younger children and rely more on group-functioning in decision-making than do younger children (Abrams et

al., 2007). This has led researchers to examine children's Theory of Social Mind (ToSM), or their understanding of how groups function (Abrams et al., 2009). ToSM predicts that support for exclusion is greater in those who better understand group functioning and that ToSM increases with age.

Thus, DSGD brings to the study of exclusion a focus on the salience of group identity and an understanding of how individuals evaluate ingroup and outgroup members who are either loyal or deviant towards group norms. SDT, additionally, brings a clear system for delineating different domains of reasoning, and a focus on the complexity of children's social interactions to the study of exclusion. Both of these frameworks provided the foundation for the current study.

Study Rationale

Challenging the Group: Morality and Group Identity

A recent study was designed to focus on gender-based groups, assessing how children respond to deviance from group norms (Killen, Rutland, et al., 2012). In this study, children and adolescents (4th and 8th graders) were asked to evaluate peer groups which either held conventional norms about dress or moral norms about distribution of resources. Additionally, in some conditions, the group norm adhered to a generic norm, which is one that reflected societal conventions or moral principles (Abrams, Rutland, et al., 2008), and some adhered to a group-specific norm, which was counter to the generic norm.

Thus, in the conventional domain, participants assessed a group that wanted to wear a club tee-shirt to a school event (traditional: conventional domain, generic norm), and a group that did not want to wear a club tee-shirt to a school event (non-traditional:

conventional domain, group-specific norm). In the moral domain, participants assessed a group that wanted to distribute resources evenly between their group and another group (equal: moral domain, generic norm), and a group that wanted to keep more resources for their own group than they gave to another group (unequal: moral domain, group-specific norm). In each condition, participants assessed members of the group who disagreed with or deviated from the group, and instead adhered to the opposite norm. For instance, if the group held an equal norm, the deviant member would espouse an unequal norm, suggesting that the group should keep more resources for themselves.

In contrast to findings from DSGD, which has previously only examined generic norms, children and adolescents were supportive of some forms of deviance. Specifically, while they asserted that groups would negatively evaluate all members who deviate from group norms, participants identified the equal and traditional deviant acts positively and asserted that they would individually like the equal and traditional deviant members (Mulvey, Hitti, Rutland, Abrams, & Killen, under review). Further, participants did not see any of these forms of deviance to be grounds for exclusion from the group.

Additionally, little ingroup bias was found. When participants were asked to choose between including someone who shared the gender identity of the group, but resisted the group's norm, or someone who did not share the gender identity of the group, but adhered to the group's norm, they paid much greater attention to the group norm and the valance of that norm, than to gender identity. Specifically, they were especially willing to include in the group a peer who did not share the gender of the group, but did share the group's norm (for instance including an equal girl in an equal boys group), when the group norm was also a generic norm (equal and traditional). Further, when they

did choose to include an ingroup member in a group, this was often when the group held a group-specific norm (unequal and non-traditional). In these cases, they recognized, not the gender-identity of the individual they chose to include, but rather the generic nature of that individual's norm (saying for instance that including an equal girl into an unequal girls group was the best choice because she was being fair).

This study introduces a number of questions, however, regarding children's willingness to accept deviance from group norms. Whereas children and adolescents were willing to accept deviance if it adhered to generic norms, they were less willing to accept deviance that resisted generic norms. Moreover, they still asserted that groups would not like deviant members, regardless of the valance of the deviance. This study, further, was conducted in the context of gender groups, but the norms of the groups were free from gender stereotypes or gender-specific information. In children's everyday lives, however, gender stereotypes are abundant and can shape children's thinking about their peer groups (Liben & Bigler, 2002).

Types of Gender Stereotypes

Thus, the current study was designed to extend this previous research by examining group norms which explicitly also involve gender stereotypes. Understanding how children perceive resistance to gender stereotypes will provide key information that can be used to help children combat harmful stereotypes as well as to design environments where children will feel comfortable expressing resistance to gender stereotypes. Challenging gender stereotypes can, in many situations, be vitally important. For instance, gender stereotypes about appropriate activities can lead to prejudice, bias and discrimination that extend through adulthood. Further, gender stereotypes condoning

aggression directly involve harm to others. Such stereotypes may perpetuate societal condoned differences in academics, sports and the workplace, leading to unequal opportunities and differential expectations for each gender.

This research was designed, then, to examine different types of gender stereotypes. Though stereotypes regarding conventions (such as choice of activities or clothing) are most commonly studied, not all gender stereotypes are conventional, or about traditions, customs and practices. In fact, some gender stereotypes surround moral behaviors. Specifically, aggression has been identified as male stereotype (labeled a “trait” stereotype by much of the literature, see Liben and Bigler (2002)). Additionally, outside of the stereotyping literature, foundational findings from peer relations researchers have identified physical aggression as a male-typed form of aggression and relational, or social, aggression as a female-typed form of aggression (Crick & Grotpeter, 1995). Physical aggression is identified as behaviors such as hitting, and fighting, while relational aggression involves gossiping, excluding others, and saying mean things to others. More recent research findings are mixed, with some research indicating that boys engage in more of both types of aggression, and girls engage in more relational aggression (Ostrov & Godleski, 2010). Other research indicates that the differences between genders in forms of aggression are minor, and that whereas boys do engage in more physical aggression than do girls, that there are no differences for relational aggression between genders (Lansford et al., 2012). Thus, it appears that associations between each gender and particular forms of aggression, are, in fact, stereotypes. There is also some evidence that children who participate in gender non-conforming types of aggression are more at risk for externalizing problems (girls who are physically

aggressive and boys who are relationally aggressive) (Crick, Ostrov, & Werner, 2006). Further, recent research indicates that girls who are aggressive (physical or relational) are more likely to be rejected and excluded than are boys who are aggressive (Kochel et al., 2012). Thus, though recent research findings are mixed regarding actual mean differences between the genders regarding different forms of aggression, messages about gender-appropriate forms of aggression are condoned by society and there are implications for non-conformity in terms of aggression. Moreover, however, both physical and relational aggression have intrinsic negative consequences for the recipients of these forms of aggression (Buhs, Ladd, & Herald, 2006), and, thus, should be considered moral transgressions. Finally, research indicates that aggressive behavior (physical) is very rare, but that it peaks by 27 months of age, and following this peak, children follow distinct and different trajectories, with some children showing persistent high levels of aggression throughout childhood and into adolescence (Brame, Nagin, & Tremblay, 2001; Broidy et al., 2003; Hartup, 2005). Further, some children who are highly aggressive are also quite popular, while others are socially rejected (Estell, Cairns, Farmer, & Cairns, 2002).

Whereas researchers examining gender stereotypes have less commonly examined forms of aggression, the peer relations findings indicate that children do perceive these aggressive behaviors to be linked to gender (Crick & Grotpeter, 1995). Physical aggression as a male stereotyped behavior and relational aggression as a female stereotyped behavior have primarily been examined within the peer relations literature, however it is important to examine these forms of aggression as involving gender stereotypes. Examining these stereotypes about aggression will enable us to better

understand how children perceive these behaviors and, specifically, under what circumstances children will be willing to reject these behaviors as unfair or unjust.

Further, more needs to be understood about how children respond to conventional gender stereotypes, such as those regarding after-school activities, both as a comparison to stereotypes involving morally relevant behaviors and norms, and in light of the significant societal messages condoning gender segregated play and activities from childhood through adulthood. Research has also established that strong gender stereotypes are held about choices of afterschool activities, and that deviating from gender stereotypic activities is perceived negatively (Horn, 2007).

Thus, new research is needed, which extends the Killen, Rutland, et al. (2012) study, and examines how children and adolescents view gender stereotyped group norms, both moral (about physical and relational aggression) and conventional (about activities, such as football and ballet). Extensive research indicates that groups do hold group norms, which are the conventions and practices of the group, including norms about how groups treat other individuals (Nesdale, 2008). No research, however, to date has compared how children respond to moral and conventional gender stereotypic group norms. This will clarify if, as found in the previous study, children attend more to the nature of the group norm than to group identity (gender), even in the face of stereotypic group norms. Will children reject both forms of aggression from both boys and girls, or will gender stereotypes which condone different forms of aggression impact their evaluations? This study will provide more information about if children and adolescents perceive group norms which involve gender stereotypes as salient and what pressures they experience to adhere to these norms.

Evaluating Resistance to Peer Group Behavior

It is also important that new research identify exactly how likely children and adolescents think group members who disagree with their gender stereotypic group norms are to express their resistance to their groups. Previous research has not assessed how likely children think resistance to group norms is, but, rather, has always established that resistance or deviance has occurred, and then asked children for their evaluations of the deviant group member (Abrams & Rutland, 2008). Research examining peer influence, though, indicates how strong an influence peer groups are, particularly for young adolescents (Brechwald & Prinstein, 2011). Additionally, research indicates that, with age, adolescents are better able to resist peer influence (Sumter, Bokhorst, Steinberg, & Westenberg, 2009). Further, it has been documented that children and adolescents conform to a range of peer group norms, including those condoning or promoting aggressive behavior (Brechwald & Prinstein, 2011). Research is needed, then, which clarifies under what conditions children and adolescents feel comfortable resisting their group. This is important as challenges to gender stereotypes and aggressive behavior, which come from within the peer group may be an effective way to combat harmful and resilient gender stereotypes as well as aggression. Thus, it is important to understand if, and in what circumstances, children will speak up, and work as the agents of change to resist gender stereotypes.

Research has also identified a bystander effect, whereby individuals may not go to the aid victims (Latane & Darley, 1970). Findings with children indicate the importance of bystanders, documenting that higher instances of defending victims is related to lower levels of aggression and bullying in classrooms (Salmivalli, Voeten, & Poskiparta, 2011).

Additionally, research has documented that higher classroom attitudes about collective-efficacy (beliefs about a group's ability to achieve a particular outcome) and, for girls, higher levels of empathy, are related to more instances of intervening in peer aggression (Barchia & Bussey, 2011). Further, research on children's evaluations of victim responses to unfair treatment reveals that, by age 7, children evaluate resisting victimization positively (Shaw & Wainryb, 2006). It is important to note that research which actually includes observations of bystander behavior is quite rare; much of this work instead documents self report of bystander behavior. Further, most of this research involves true bystanders, those who observe, but are not directly involved in a situation, or the victims, themselves. It is often the case, though, that children and adolescents may directly be part of groups which engage in aggressive behavior or perpetuation of stereotypes. Further, research documents that if adolescents are part of groups which engage in antisocial behaviors, like aggression, talking about these behaviors actually increases instances of these antisocial behaviors (Piehler & Dishion, 2007). This indicates that something like "deviancy training" occurs among adolescents as part of their conversations in their peer groups (Piehler & Dishion, 2007). Less is known, however, about if peers can influence members of their own peer group in positive ways. Thus, new research is needed which examines if children view resistance to the norms of one's own group as likely, or not. This is particularly important given the pervasive influence of peers (Brechwald & Prinstein, 2011), on the one hand, and the positive impact that resisting aggression can have on future instances of aggression (Salmivalli et al., 2011).

Whereas the peer influence literature and developmental subjective group dynamics, generally, show how powerful the influence of peer group norms can be on children and adolescents, there is also research that shows little relation between the prejudices and stereotypes held by adolescents who are friends (Ritchey & Fishbein, 2001). This indicates that, while a peer group may hold a stereotypic norm, there may be individual members of that group who do not subscribe to that stereotypic belief.

Further, research indicates that children do recognize the importance of personal opinion and personal choice: When asked to decide between a personal preference and a friend's preference, younger children prioritized the friend's preference across conditions, whereas adolescents asserted that fulfilling one's personal prerogative was an acceptable decision (Komolova & Wainryb, 2011). In these situations, however, neither gender stereotypes nor norms about aggression were involved. Additionally, these were dyadic exchanges, where asserting one's own preference or belief may be easier than in a group scenario. This research does indicate, however, the importance of personal choice (Nucci, 1981) and autonomy for children, which suggests that children may, especially in the context of social-conventional norms involving gender stereotypes (such as which activity to participate in, football or ballet), see a role for autonomy and personal choice. Across all scenarios, the question remains, though, do children think that these dissenting members can and will voice their resistance? Further, how will they, personally, evaluate such resistance, and how do they think groups will evaluate it?

Repercussions for Resistance: Exclusion?

It may be that one of the key factors which determines if children think that resisting stereotypic group norms is likely will be their sense of what the repercussions,

or consequences, for resisting the group will be. Children may be unwilling to express resistance when they are part of the group that is perpetuating stereotypes or aggression because of fear that changing the status quo may negatively impact their position in the group (Juvonen & Galván, 2008). Research with adults indicates that adults dislike peers who do not conform to negative group behaviors (“moral rebels”) (Monin, Sawyer, & Marquez, 2008), but research with children indicates that children like peers who resist the group in order to assert behaviors that align with moral principles (Killen, Rutland, et al., 2012). It may be though, that while children, individually evaluate such forms of deviance positively, they also think that deviating from group norms will have consequences, making such deviance less likely.

One potential, and particularly harmful, consequence for resisting the peer group is exclusion from that peer group. Research has documented that social-exclusion can have significant impacts on children’s academic motivation and success in school, as well as on their mental health and well-being (Buhs et al., 2006). Additionally, children who fail to adhere to social group norms are at risk for exclusion from those groups and rejection by their peers (Abrams & Rutland, 2008; Juvonen & Galván, 2008). Finally, extensive research with children and adolescents indicates that children reject exclusion as unacceptable (Killen, Sinno, & Margie, 2007), but may use information about stereotypes (Killen, Pisacane, Lee-Kim, & Ardila-Rey, 2001), personality traits (Park & Killen, 2010), or group norms (Killen, Rutland, et al., 2012) to justify exclusion. Further, while much research has focused on the social deficits of the target of exclusion or rejection from the peer group, exclusion can also occur because of one’s group membership. A recent proposal urged researchers to consider the role that both

interpersonal and intergroup components could play in exclusion decisions (Killen, Mulvey, & Hitti, 2012). In the context of challenging one's peer group about gender stereotypes regarding aggression, little is known about if peers are concerned that they will be rejected from the group because they advocate for gender non-conforming behaviors (for instance, a boy urging his group to play nicely instead of rough). One may, in fact, be excluded from a group with whom one shares group membership for not exhibiting normative (stereotypic) group behavior (in terms of gender, in this context).

Thus, this research aims to identify whether exclusion is viewed as a likely consequence for resisting the group's stereotypic and aggressive norms, and whether a relation exists between children's judgments about the likelihood that exclusion will be a consequence for resistance and their judgments about the likelihood of resistance, itself. Previous research has assessed if participants would support exclusion as a consequence for a particular behavior, for instance resisting the group norm; no research to date, however, has assessed if participants think that groups would be likely to exclude someone because they challenge the group norm. The difference here is that while children generally reject exclusion, they may, in fact, believe that *groups* would condone exclusion. Even if children evaluate exclusion as morally unacceptable, they may think that groups will use exclusion as a consequence for resisting the group. Thus, it is important to assess if children and adolescents think that exclusion will be a consequence for disagreeing with the group. If children do perceive exclusion to be a potential consequence for resisting gender stereotypic norms, will they be less likely to express their disagreement to the group? Understanding this relation will provide key new information to aid teachers, policy-makers and group leaders in fostering peer group

relationships where resisting negative peer behavior is seen as a possibility and where children do not fear repercussions such as exclusion.

Social-Cognitive Skills

The current study, thus, examines children's evaluations of the likelihood of resistance, and of exclusion, as well as individual and group responses to resistance, in the context of group norms about social-conventional and moral gender stereotypes. While understanding children's decision-making will prove insightful, it may be the case that underlying social-cognitive skills may impact children's evaluations (Mulvey, Hitti, & Killen, 2013). Specifically, research has identified the relation between theory of mind, the ability to recognize other's intentions, beliefs and desires (Wellman & Liu, 2004), and moral judgments (Killen et al., 2011). Further, research indicates that a relation may exist between perspective-taking and shared experiences (Chandler & Helm, 1984), gender stereotypes and theory of mind (Kelly, Mulvey, Hitti, & Moin, 2011; Terwogt & Rieffe, 2003), between theory of mind and understanding of gender discrimination (Brown et al., 2010), and between theory of mind and evaluations of exclusion (Abrams et al., 2009). Additionally, poor theory of mind skills have been linked to problematic outcomes: young children with poor theory of mind skills predict the likelihood that one becomes a victim, a bully or a bully-victim in later childhood (Shakoor et al., 2012) Thus, one aim of this study is to identify if children who exhibit theory of mind competence differ from children who do not exhibit theory of mind competence in their evaluations of group members who challenge gender stereotypic group norms. This research will examine more complex forms of theory of mind. While much research with theory of mind uses simple laboratory-based assessments, such as the

traditional false-belief tasks (Wellman & Liu, 2004), recent research indicates that, when assessed in complex peer interactions, including those involving a potential victim, it may be more challenging to employ theory of mind skills (Killen et al., 2011). Further, this research indicates that employing theory of mind skills may be more challenging in complex social scenarios. This will clarify if potential age-related differences are driven by social-cognitive development. Specifically, it may be the case that possessing theory of mind skills will enable participants to better take into account both the perspective of the group and the perspective of the individual member, in evaluating the likelihood of resistance as well as the consequences for resistance in terms of exclusion.

In sum, this research addresses under what conditions individuals view stereotypic expectations as legitimate or unfair and how children conceptualize the costs to challenging gender stereotypes, and the social-cognitive requirements for making such judgments. How do children and adolescents weigh the negative moral nature of these gender-stereotyped aggressive behaviors with the prevailing social acceptance of these behaviors for each gender? Can and do they recognize the value of breaking these gender stereotypes and how do they think their peer groups will respond to such resistance?

Current Study: Design and Hypotheses

Study Design

Two age-groups (9-10 year olds and 13-14 year olds) were chosen for this study as these groups span development from childhood to adolescence. Additionally, as research indicates both that children become more flexible in their thinking about stereotypes with age (Arthur, Bigler, Liben, Gelman, & Ruble, 2008), and that they are

still quite unwilling to accept gender non-conformity through adolescence (Horn, 2007), examining both children and adolescents is theoretically of interest.

Participants assessed scenarios about groups that conform to and resist stereotypes about: a) physical aggression, b) relational aggression, and c) gender stereotypic activities. Finally, participants completed an *Interpretive Theory of Mind Task* (Carpendale & Chandler, 1996) and a *Stereotype Awareness Task*, which was modified from Signorella, Bigler, and Liben (1993). See Appendix B, for an example protocol.

Participants evaluated how likely they think resistance to the group norm is, as well as how likely it is that they, personally, would resist the group norm. While research has examined how children respond to deviance from groups (Abrams & Rutland, 2008), less is known about if children and adolescents think that resistance is likely.

Understanding more about likelihood of resistance will fill an important gap in our knowledge of children's evaluations of social relationships. Further, research indicates that children, individually, do not always agree with the decisions of group (Mulvey et al., under review), but less is known about if children will agree with an individual group member's decision to resist a group norm in different contexts.

Additionally, participants made favorability judgments, assessing their expectations of how much the group would like a member who challenges the group and rating how much they would individually like such a member. These measures allowed for an examination of if participants are able to distinguish their own perspective from the group's perspective and to assess instances in which children may recognize that the group may hold a belief which they do not share.

Additionally, two measures assessed potential repercussions for resisting group norms involving gender stereotypes. While previous research has most often assessed acceptability of exclusion, finding that children and adolescents are not willing to exclude others (Killen & Rutland, 2011), exclusion is still pervasive among children. This study extended previous measures by assessing likelihood of exclusion. This item measured not if participants approve of exclusion, but rather, if participants think a dissenting member *will be* excluded for expressing disagreement. Thus, in order to better understand this discrepancy, this study included an assessment of if participants think that exclusion will be a repercussion for challenging a group's norms.

In many situations, children must choose between peers and only include one person. The last measure required that participants make a forced choice inclusion decision. This measure also assessed repercussions for resisting the group: do participants think that a target *should be* denied entry into a group because he or she disagrees with the group? Additionally, this measure assessed the relative weight that participants put on group identity (gender) and group norms in assessing who should be part of different social groups.

For each of these measures (except for the first two), reasoning was also assessed, drawing on social domain theory. Measuring reasoning is particularly important, as it helps to identify why children and adolescents make the decisions they do and what is driving such decisions. Little research which has examined children's understanding of group dynamics has also assessed reasoning, yet reasoning can provide essential insight into precisely why children condone or reject certain behaviors in group contexts.

Study Hypotheses

See Table 1 for an overview of hypotheses and analyses that were conducted. Overall, it was expected that participants would rely less on gender stereotypes for the moral conditions than for the conventional conditions. Additionally, it was expected that participants will be more favorable to targets who resist aggressive behavior. There may be a shifting standard (Biernat & Manis, 1994), whereby gender non-conformity (resisting gender stereotypes) will be less acceptable for boys than for girls, in both the moral and conventional domains.

Likelihood of resistance and individual likelihood of resistance. It was expected that participants will be more likely to expect targets will resist the group when the group is stereotype non-conforming than when the group conforms to stereotypes. Further, resistance will be less likely for the gender stereotypic activities when the group adheres to stereotypes than when they do not. It was expected that participants would least expect resistance from a boy who wants to do ballet when his group wants to do football, given the research on the shifting standard (Biernat & Manis, 1994).

In terms of age group, it was expected that children would see deviance as more likely than will adolescents due to greater concern by adolescents for group functioning (Horn, 2003) and concern over peer influence (Brechwald & Prinstein, 2011). Finally, it was an open question if participants would differ in their responses depending on if they are assessing their ingroup or their outgroup.

It was expected that participants, themselves, would be more attuned to challenging aggressive behavior than they would expect their peers to be. Thus, it was expected that participants would rate their own likelihood of resistance to aggressive norms to be higher than the ratings they provide for their peers.

Group favorability toward dissenting member. Participants would likely expect that groups will not like dissenting members, but this would vary if the dissenter is resisting or condoning aggression. Further, groups would be more favorable to dissenting members who resist aggression, as aggression would likely be viewed as a moral transgression. There may be a shifting standard, with participants asserting that groups would find deviance towards non-conforming behaviors as less acceptable, especially for boys who want to do ballet. There would likely be differences in reasoning for this assessment as well, with participants citing group functioning and other societal justifications when asserting that the group would not like dissenting members who resist stereotypes. Adolescents would be more likely to reference group functioning than would children. Additionally, some participants would likely justify their responses in terms of stereotypic concerns, particularly in the conventional context.

Individual favorability toward dissenting member. While it was expected that group favorability judgments would be influenced by stereotypic expectations, it was expected that individual favorability toward the dissenting member would be more strongly influenced by the norm of the group, with participants supporting dissenting members who adhere to generic moral principles and resist aggression, regardless of their gender. Status differences may emerge for the gender stereotypic activities, with non-conformity being seen as more positive for girls, than for boys, because participants may believe that boys should not move down the status ladder to act in stereotypically female ways.

There would likely also be differences by theory of mind ability, with participants with more developed theory of mind skills more likely to positively evaluate members who challenge a group's aggressive norms.

There would likely be differences in reasoning used as well, with participants referencing moral reasons when supporting the dissenting members who avoid aggression, and referencing harm, in particular, when evaluating the dissenting members who engage in aggression more negatively. Adolescents would likely focus on group functioning, particular in the context of gender stereotypic activities (football and ballet) and physical aggression during a sports game. Finally, with age, participants would reference autonomy, saying that it is up to the dissenting member.

Likelihood of exclusion. It was expected that, unlike prior research which has found that children view exclusion as unacceptable in most cases, children would identify exclusion as a likely repercussion for resisting the group. Generally, children would be more likely to expect exclusion, as their social groups may be less fixed than adolescents' and, thus, more malleable. There may be a shifting standard, with participants expecting exclusion of the boy who wants to play ballet when the group wants to play football. Exclusion would be seen as more acceptable for the dissenting members who advocate for aggression (going against moral principles by gossiping and playing rough) than for those who resist aggression.

Reasoning would likely vary based condition. Finally, it was expected that participants may make more use of multiple domains in their reasoning for this question, recognizing the unfairness of exclusion by using the moral domain, but also noting the

likelihood of exclusion by referencing conventional forms of reasoning such as group goals and group functioning.

Inclusion Choice. Dissenting member who go against moral principles (by gossiping and playing rough) were less likely to be included. On the other hand, participants would choose to include someone who does not share their group's gender identity, but does share their group's non-aggressive norms, for instance playing nice and being impartial.

Reasoning would likely focus on moral reasons when participants choose to include the non-aggressive target, regardless of if the target is conforming to or resisting stereotypes. Participants with a higher awareness of stereotypes were expected to use conventional reasoning supporting stereotypes, including appeals to maintain the gender identity of the group. Finally, participants may reference social justice and resisting stereotypes when choosing to include the non-stereotypic child in the group, regardless of the group norm.

Implications

Overall, this research will reveal new information about how children evaluate resistance to gender stereotypes, including those condoning different forms of aggression. Importantly, it will also reveal children's views about the repercussions for challenging stereotypes, specifically in terms of exclusion. This research will have important implications for understanding how and under what conditions children choose to resist gender stereotypes, particularly about aggressive behavior. Challenging groups about gender stereotyped aggressive behavior may not be easy for children, because of concerns about repercussion from the group in terms of exclusion.

Examining dissent from gender norms will provide insight into important social issues. Specifically, this study has implications for discrimination based on gender norms and expectations, and will provide insight into the developmental origins of moral judgments as well as prejudicial attitudes and behaviors. The findings will help to improve intergroup relations by informing educators, and parents of best practices for teaching tolerance, diversity, and acceptance of others, and for reducing stereotyping. This study will provide a greater understanding of how children and adolescents respond to gender stereotypes which can lead to interventions and curriculum to be used with children to encourage acceptance, and reduce gender discrimination.

Chapter II

Literature Review

As children and adolescents form social relationships and become part of social peer groups, they experience exclusion as well as opportunities to include or exclude others (Killen, Lee-Kim, McGlothlin, & Stangor, 2002). Researchers in developmental and social psychology have studied social exclusion in a wide range of contexts, from situations in which exclusion occurs between dyads, individuals and groups. The focus of this review will be on gender exclusion, specifically when exclusion from peer groups occurs based on gender identity, as well as on the social-cognitive skills children bring to evaluations of gender exclusion.

In the peer relationships literature, exclusion has often been studied in terms of individual differences, where peer rejection is identified as being an outcome of social competence deficits (e.g., aggression; social anxiety) (Rubin, Bukowski, & Parker, 2006). Exclusion does not always occur because of individual social deficits on the part of the excluded child, however. Rather, exclusion is often based on stereotypes, or biases about a particular group, including stereotypes about gender (Killen, Mulvey, et al., 2012). While exclusion based on gender has existed in cultures for millennia, only in recent decades has there been an explicit research focus in developmental psychology on gender exclusion, that is, how children and adolescents evaluate contexts in which girls or boys are excluded from a group based solely on gender. Research conducted on gender exclusion with children and adolescents in the U.S., Korea, Denmark, and Switzerland (to date) has revealed that gender exclusion is viewed as more acceptable than exclusion based on race and ethnicity (for a review see Hitti et al., 2011). Understanding exactly why and under

what conditions exclusion based on gender is still seen as legitimate will provide insight into the developmental origins of the societal conventions regarding education, occupation and gender roles that contribute to exclusionary decisions and outcomes. Gender exclusion differs from many other forms of exclusion due to the societal affirmation of gender exclusion in sports, schools, and social contexts. Yet, how individuals evaluate it and the social-cognitive processes involved in this evaluation remain less well understood.

Further, researchers have pointed out that more attention should be focused on how gender roles inhibit both girls' and boys' academic, athletic, and professional aspirations (Eccles, Roeser, Vida, Fredricks, & Wigfield, 2006; Ruble, Martin, & Berenbaum, 2006; Sinno & Killen, 2009). For instance, boys are often type-cast as aggressive, tough and sports-oriented, with little focus on their academic or artistic potential. Girls, on the other hand, are seen as fragile and non-athletic and continue to be excluded from athletic opportunities, despite Title IX legislation. In fact, the gap between athletic opportunities for girls and boys has been widening in recent years (NCWGE, 2007). Finally, gender-segregated youth programming (for instance Girl Scouts and Cub Scouts) continues to be offered to children and research indicates that even into adolescence children most frequently socialize in single-sex groups (Maccoby, 2002). Further, socially-condoned messages about gender roles for girls and boys impact their self-esteem, as well as their motivation in a range of academic and social domains (Liben & Bigler, 2002; Ruble et al., 2006). Gender based exclusion is complex, as it often begins in childhood in socially-condoned, yet discriminatory, forms, and carries into adulthood in pernicious ways.

One reason for the complexity of gender-based exclusion is that these forms of exclusion are not often explicitly labeled as based on one's gender or adherence to gender

identity norms. Rather, gender-based exclusion is often rather ambiguous. Research in the peer-rejection literature suggests that there is a significant amount of variation in how children respond to ambiguous instances of potential exclusion. Specifically, Downey et al. (1998) found that some children are more sensitive to rejection than are others and that some children interpret many more scenarios as involving exclusion and rejection than do others. This suggests that at least some children may have difficulty in anticipating, interpreting and understanding potential exclusion experiences based on gender.

Children may differ in their responses to exclusion messages because of differences in their social-cognitive abilities due to their ages. What it is not yet known is how children bring their cognitive skills and psychological knowledge into play when making evaluations about the likelihood of exclusion. Specifically, exclusion decisions, particularly those centered on gender issues, often involve ambiguous situations which include stereotypes and require that children make a variety of judgments about others' psychological states. They need to balance a range of information in making exclusion decisions, including: a) the intentions of the target and the group that is potentially excluding the target; and b) the perspective of the target and the group that is potentially excluding (i.e. what do they think is happening). Some children may not be able to attend to the multiple competing perspectives in making such exclusion decisions, however.

This may be particularly challenging when considering gender-based exclusion as children may struggle to interpret the perspective and intentions of a group or individual who is of the opposite gender. Additionally, some research suggests that stereotypes regarding behavior and activities for a particular gender may cloud social-cognitive

abilities (Kelly, Mulvey, Hitti, & Killen, May 2010; Terwogt & Rieffe, 2003).

Understanding the varied perspectives and intentions of those involved and balancing this understanding in making judgments may rely upon social-cognitive skills, such as theory of mind (ToM), the ability to understand others' intentions, beliefs and desires (Wellman & Liu, 2004). Thus, more research is needed which unravels how children and adolescents respond to instances of gender-based exclusion, and, in particular, what cognitive skills they bring to this task. Better understanding of the cognitive processes surrounding decisions to include or exclude based on gender will provide insight into how gender-based exclusion should be addressed to reduce prejudice and discrimination.

Overview of the goals of the paper

In this review, current research on social exclusion and the development of social-cognitive abilities will be discussed with the goal of pointing to new directions for future research which can address gaps in the literature. Two areas of research will be addressed: a) social exclusion and gender identity; and b) the potential links between socially-relevant forms of theory of mind and evaluations of social exclusion. This review will first briefly introduce the two focal areas, examine theoretical perspectives which influence research on evaluations of exclusion, review key studies examining children's evaluations of exclusion, as well as research on children's social-cognitive abilities, then conclude with an analysis of the current gaps in the literature, which can be filled with the current study.

Research on gender exclusion in childhood

Social exclusion from groups is multi-faceted. Sometimes exclusion is legitimate for group-functioning reasons and other times it is wrong because it is based solely on

group membership, including gender (Rutland et al., 2010). For instance, while it may be legitimate to exclude a girl from a soccer team in an all-boys league, it might be viewed as wrong by some individuals to exclude a boy from a boys' team because he did not adhere to male gender-stereotypes. While many forms of exclusion are acceptable for conventional or prudential reasons, exclusion can also be based on group membership alone. This form of exclusion is problematic because it can be related to bias and prejudice (Killen et al., 2002). Exclusion can have negative consequences on the mental health and future social interactions (Murray-Close & Ostrov, 2009; Rubin et al., 2006) and academic motivation and success (Buhs et al., 2006; Buhs, McGinley, & Toland, 2010; Eccles et al., 2006) of those who are excluded. Moreover, as discussed above, research reveals that children and adolescents are more accepting of social exclusion based on gender than they are of other forms of social exclusion (Killen et al., 2007), which indicates the importance of continued focus on gender exclusion in particular. Further, the peer relations literature indicates that children are often excluded because of aggressive personality traits (Rubin et al., 2006), and that children may hold stereotypes condoning physical aggression for boys and relational aggression for girls (Crick & Grotpeter, 1995), even though this may not reflect actual behaviors for girls and boys (Card et al., 2008; Lansford et al., 2012). Thus examining exclusion decisions in light of behavior which either aligns with or resists these gender stereotypes is important.

Social exclusion often occurs in intergroup contexts, when children are interacting with others who do not share membership in the same gender, ethnicity or religious groups. Thus, when children evaluate instances of exclusion their identification with their group and their sense of the group's goals and norms (practices or conventions)

often come into play. Additionally, in some instances, children may bring stereotypic knowledge of other groups to their intergroup interactions. Research from a social reasoning developmental perspective within the field of developmental intergroup relations has examined the interplay between group identity, morality and social decision-making (Rutland et al., 2010). Research reveals that children often express a bias towards their ingroup and express derogation of the outgroup, and strive to uphold the norms of their ingroup. Understanding how children interpret conflicts between conceptions of what is fair and their loyalty to their ingroup will be particularly important in studying gender based forms of exclusion as children do strongly identify with their gender ingroup (Liben & Bigler, 2002).

The specific focus of this literature review will be to describe research on gender exclusion, gender group norms, and gender stereotypes, especially those related to aggression. Research on these dimensions of gender exclusion in childhood will provide insight into how children balance information about stereotypes, their understanding of what is fair and just, and their notions of group dynamics. Additionally, this research has the potential to clarify why it is that gender exclusion is still seen as acceptable in many instances, and in particular, may clarify the continuing reliance upon gender roles in shaping academic, social and occupational trajectories.

Evaluating exclusion: The role of social-cognitive abilities

Evaluations of exclusion have rarely been examined in light of children's social-cognitive capacities and the variety of forms of psychological knowledge that come into play in instances of social exclusion. Traditional measures of social-cognitive abilities to evaluate intentions and others' mental states are often laboratory-based and removed

from the rich social context of children's lives. Children's ability to interpret the intentions, goals and emotions of those involved in exclusion scenarios involving gender may, however, hinge upon their social-cognitive development. Recently, research has identified ways to examine social-cognitive skills, particularly theory of mind, in more contextualized and authentic ways. Yet, these new means of measuring more socially-relevant forms of theory of mind are rarely applied to children's social-decision making processes and their evaluations of exclusion. This research, however, does indicate that studying how this might play out within the context of gender stereotypes will provide insight into how children resolve tensions between their understanding of intentions, their own allegiance to their group and its goals, and stereotypes which they may hold regarding gender. Thus, research examining social-cognition, and theory of mind, in particular, in socially-relevant ways, as well as research that suggests potential conflicts between social-cognition and gender stereotypes will be described.

This review, then, will examine research on gender exclusion with a new focus the role of the development of social-cognitive abilities and how these are related to social exclusion and moral judgments.

A Framework for Studying Exclusion: Integrating Social Domain Theory, Social Identity Theory and Examinations of Theory of Mind

While social exclusion research has been approached within many different contexts, two different, yet complementary frameworks have been drawn upon when studying inclusion and exclusion: Social Domain Theory, which distinguishes between the societal, moral and psychological domains of knowledge used in making judgments (Turiel, 1983), and Social Identity Theory (SIT), which argues that individuals strive to

maintain their ingroup identity by viewing their own social group more positively than other social groups, and that individuals identify with social groups with positive or higher social status (Tajfel & Turner, 1979). Recent research has begun drawing upon both of these frameworks, in order to examine both moral judgments about social exclusion as well as to recognize the importance children place on group identity and group norms in peer interactions (Killen, Mulvey, et al., 2012; Killen & Rutland, 2011; Rutland et al., 2010). Children simultaneously develop moral beliefs about issues such as welfare, fairness and justice and a sense of group identity (Rutland et al., 2010). In some instances, children, when balancing their developing moral beliefs and their sense of group identity, may turn to ingroup bias and prejudicial attitudes in making decisions to exclude others based on group membership.

In making these decisions, children will necessarily have to weigh their understanding of the intentions, goals and desires of the potential target for exclusion, the group that may exclude, and individual members of this group (Mulvey et al., 2013). To effectively balance these perspectives they may rely upon complex theory of mind skills, in recognizing others' mental states (Wellman & Liu, 2004). Thus, in understanding instances of exclusion when group identity and norms as well as issues of fairness are at play, research should draw upon theory of mind research, Social Domain Theory and Social Identity Theory. The following section will provide a brief overview of these three frameworks for understanding social-cognition.

Social Domain Theory

Extensive research within Social Domain Theory (Smetana, 2006; Turiel, 1983) has shown that children from a very young age (3 years) differentiate their experiences

within the social world, including experiences of exclusion, based on different domains:

1) the moral domain, which includes concerns with welfare, fairness, justice and rights;
2) the societal domain, which includes situations contingent on socially agreed upon rules that are alterable and that, if not present, will cause no direct harm to be inflicted on another; and 3) the psychological domain, which includes personal preferences and choices (Smetana, 2006). Social Domain Theory has guided research revealing the importance of examining children's reasoning when studying exclusion decisions and has provided support for the recognition that children distinguish, from very early ages, between decisions which inherently center on issues such as fairness and justice and those which, drawing on conventions, promote group functioning and social interactions (for a review see Killen et al., 2007).

Social Domain Theory has been a primary framework for the study of exclusion, first in the United States, but also increasingly in other areas of the world, such as the United Kingdom, Spain, The Netherlands, Germany, Korea, Switzerland, China, and Japan (see Hitti et al., 2011, for a review). As will be described in greater detail below, exclusion based solely on group membership, including gender, is generally evaluated as unfair and judged as wrong by most children (Killen et al., 2002), however in complex or ambiguous situations children will often make exclusion judgments based on stereotypes about group identity (Killen et al., 2001). Additionally, children evaluate peer group exclusion based on gender as wrong, however many children still rely upon stereotypes when they reason that exclusion might be acceptable (Killen & Stangor, 2001). Thus, Social Domain Theory has revealed the complexity involved in children's reasoning about social exclusion based on group membership.

Social Identity Theory and Subjective Group Dynamics

Exclusion has also been studied by drawing on Social Identity Theory (Tajfel & Turner, 1979), particularly in Europe and Australia. Attempts to see the ingroup in increasingly positive ways can lead to prejudice towards members of outgroups (Tajfel & Turner, 1979). Even very early research examining intergroup relations revealed the manifestation of intergroup bias in social interactions (Sherif, Harvey, White, Hood, & Sherif, 1961). Self-categorization theory (SCT; Turner, et al., 1987) suggests that, based on cognitive grouping, individuals develop a sense of how to identify one's self (Nesdale, 2004). Thus, people place themselves in a group that they view as most similar to themselves based on some classification label, which is cognitively contrasted with another classification. Such self-categorization emphasizes positive similarities between individuals of the ingroup, thus promoting ingroup bias, while also focusing on the negative differences of the outgroup, which may lead to outgroup prejudice. This process, then, creates opportunities for the development of stereotypes and acts of exclusion based on group membership (Abrams & Rutland, 2008).

One extension of Social Identity Theory, Developmental Subjective Group Dynamics (DSGD), proposes that individuals differentially evaluate and include others based on their adherence to or deviance from group norms (Abrams & Rutland, 2008). As will be described in more detail below, research drawing on DSGD finds that children prefer outgroup members who deviate from their group norms and, thus, espouse the child's ingroup norms more than they like ingroup children who deviate from ingroup norms (Abrams, Rutland, & Cameron, 2003; Abrams et al., 2009). Thus, the framework of Social Identity Theory, and Developmental Subjective Group Dynamics, in particular,

has emerged as a way to focus exclusion research on attending to group norms and group identity.

Intentionality, Theory of Mind and the Psychological States of Others

In addition to social-cognitive theories of development, theories concerning the cognitive development of the ability to understand the psychological states of others will be important to consider when studying social exclusion. This is in part because decisions about social exclusion involve a heavy cognitive load: children must balance information about all of the characters involved in the scenario, while also attending to these characters' intentions, beliefs and desires (theory of mind) and to the group dynamics involved, including the role group identity might play, the norms or conventions held by the group, and stereotypes or assumptions they may hold about a particular group. Two different theoretical approaches have been taken to understanding children's developing theory of mind abilities.

Some researchers suggest an "early competence" model, often called the Theory of Mind Mechanism/Selection Processing Model, which suggests that humans possess an innate capacity for theory of mind abilities which are triggered by environmental factors (Scholl & Leslie, 2001). Essentially, this theory suggests that during the preschool years, children begin to selectively attend to mental states of others through a domain-specific mechanism, which enables them to begin understanding mental representations (Leslie, Friedman, & German, 2004; Scholl & Leslie, 2001). While this theory has some support from empirical findings (Scholl & Leslie, 2001), others argue that it is either an unnecessary theoretical proposition (Stone & Gerrans, 2006) or that the empirical evidence does not support this theory because there does appear to be such a strong age-

related pattern associated with the development of theory of mind (Wellman, Cross, & Watson, 2001).

Gopnik and her colleagues have proposed that conceptual change during the preschool years enables children to understand that not all representations which individuals hold about the world are necessarily accurate (Gopnik & Wellman, 1994; Wellman et al., 2001). This form of conceptual change is thought to occur, according to Theory-Theory (Gopnik & Wellman, 1994), the primary framework which advocates for conceptual change as a driving force behind development of understanding of things like others' mental states, because children approach the world like scientists, testing hypotheses, observing evidence, and refining their understanding based on what they discover (Gopnik & Meltzoff, 1997). Evidence within the theory of mind research provides support for the conceptual change model as research suggests that children do show age-related improvements in theory of mind ability and can successfully complete different tasks at different ages, which may reflect improvement of their conceptual understanding of others' lives with age (Wellman & Cross, 2001; Wellman et al., 2001). What has not fully been addressed by theory of mind researchers are the mechanisms of change. Carpendale and Lewis(2006) have proposed that peer interactions facilitate change in the ability to understand mental states. More research is needed to examine this aspect of developmental social cognition.

While both theories continue to be debated in the research literature, what remains uncontested is that children develop the ability to understand others' mental lives. These abilities are most commonly studied in laboratory experiments that are fairly removed from children's authentic lives. This is despite early research which indicates that

improvement in perspective-taking abilities occurs in situations in which children have opportunities for rich social interaction with their peers, in particular, through shared experience (Chandler & Helm, 1984). Theory of mind abilities, though, have significant implications for children's ability to make judgments about others' desires, intentions and motivations, all of which play an important role in evaluating social exclusion. There has been a call in the research literature for more work which examines the real-world implications of the development of theory of mind competence (Astington, 2001). Within this review, research which has taken on that challenge will be examined and proposals for new ways in which theory of mind can be applied to studies of social exclusion, and gender based exclusion in particular, will be explored.

Gender Stereotypes

In popular culture, gender stereotypes regarding behavior abound, with physical aggression by boys excused as "boys will be boys," and books and television shows like "Gossip Girl" reinforcing and even glorifying the idea that gossip is appropriate for girls. From a very early age, children develop stereotypes about gender (Liben & Bigler, 2002; Mulvey et al., 2010). At the same time, even very young children are capable of making moral decisions (Turiel, 1983). Gender stereotypes about behavior permeate peer groups, often functioning as the norms, or conventions, which shape groups.

A large body of research on gender stereotypes exists, and identifies the early emergence of gender stereotypes, as well as the pervasive influence of gender stereotypes on children's behavior and self-understanding. There are a number of different theoretical models which address how children begin to understand gender, and the relationship between understanding gender and gender stereotyping. As an example,

Martin and Halverson (1981) propose a gender scheme theory, indicating that children learn gender and gender-appropriate activities, behaviors, and objects through interaction with their social worlds: they hear peers, adults and the media indicate what behaviors and activities are appropriate for each gender and begin to act in ways which reflect those societal messages. More recently, Arthur et al. (2008) have proposed a developmental intergroup theory for explaining the emergence of stereotyping in young children. They argue that children are able to perceptually distinguish males from females, even as early as infancy, that differences in the proportions of each gender taking part in specific activities (for instance few boys attend dance lessons) heightens their understanding of males and females as different, that explicit labeling of gender (for instance by teachers who ask all the girls to line up and then all the boys to line up) reinforces stereotypes, and that implicit differentiation by adults (for instance through segregation of boys and girls during some activities) further enhances children's understanding of gender stereotypes.

Trautner et al. (2005) found that gender stereotypes increase during early childhood, with children peaking in their rigidity by approximately age 7, but that stereotypes continue to play a strongly influential role for children beyond age 10, even though most children are much more flexible in their thinking about gender stereotypes by this time. Further, research on gender stereotypes has examined stereotypes in a range of unique domains, including activities, occupations, and traits (Liben & Bigler, 2002).

One particular area in which gender stereotypes abound is in regards to aggressive behavior. Research on peer relationships has extensively examined gender differences in frequency of physical and relational aggression for boys and girls, at first concluding that boys engaged in more physical aggression and girls in more relational aggression (Crick

& Grotper, 1995), and more recently determining that, in fact, boys engage in more physical aggression but that there are not differences, or there are scant differences, in mean rates of relational aggression between boys and girls (Card et al., 2008; Lansford et al., 2012). What has not yet been examined closely is if and how children use these stereotypes about gender and aggression when evaluating their peers and making social decisions in groups.

Stereotypes and Exclusion

One result of children's rigid adherence to stereotypes is denial of opportunity. Specifically, if children hold rigidly to stereotypes, they may deny children who do not adhere to such stereotypes with the opportunity to engage in gender non-conforming activities. For instance, research indicates that preschool children may rely upon stereotypes about activities (playing with dolls and trucks), and willingly exclude someone who does not match a gender stereotype (Theimer, Killen, & Stangor, 2001), but that older children may be more attuned to issues of equity, and reject straightforward exclusion of gender non-conforming children because of moral reasons (Killen & Stangor, 2001). This is important as research indicates that as children spend more time playing in single-gender groups, they become increasingly invested in playing with gender conforming toys (Martin & Fabes, 2001). In more complex scenarios, however, when children are asked to weigh information about skill and gender stereotypes, children are more likely to exclude a non-conforming target, justifying their decision through references to group functioning (Killen & Stangor, 2001). Across these studies, however, exclusion based on gender was seen as more acceptable than exclusion based on race, thus children may attend to societal messages condoning gender stereotypes. For

instance, research indicates that adolescents do not accept gender non-conformity in behavior and activities (Horn, 2007). Further, there is some indication that willingness to exclude others is related to social-cognitive abilities, in particular, theory of mind (Abrams et al., 2009). Thus, research on these different types of stereotypes has firmly established that children hold such stereotypes and use them in their everyday lives, less is known about when children are willing to resist such stereotypes and the social-cognitive abilities that they may bring to such evaluations.

Children's and Adolescents' Reasoning About Exclusion Because of Group Membership

Returning to the foundational manner in which social exclusion has been studied will provide insight into what is already known about children's reasoning about exclusion as well as suggest potential areas for future research. Much of the foundational literature on exclusion explored children's judgments of exclusion because of group membership, including gender. Children can be excluded from groups or activities because of their gender, ethnicity, religion, school affiliation, or nationality (Killen, Mulvey, et al., 2012). The research reviewed here will focus on exclusion due to gender and ethnicity, because the majority of research conducted thus far focuses on these forms of group membership.

Theimer, Killen and Stangor (2001) analyzed European-American preschool children's ($N= 50$) evaluation of inclusion in an activity based on gender-stereotypic expectations of peer activities (e.g., doll-playing, truck-playing) and peer roles (e.g., deciding who will be the teacher and firefighter). Participants evaluated straightforward exclusion scenarios based on gender (for instance a boy wants to join a group of girls

playing with dolls), as well as multifaceted exclusion scenarios, where participants were given information about the experience of the child who wants to join the group. In the equal experience context, participants were asked to choose between a gender stereotype consistent child (e.g., a boy wanting to join a group of boys playing with a truck) and a gender stereotype inconsistent child (e.g., a girl wanting to join a group of boys playing with a truck) when both children were experienced with playing with the target toy or performing the target role. In the unequal experience context, participants were asked to choose between a gender stereotype consistent and a gender stereotype inconsistent child when the gender stereotype inconsistent child has no experience playing with the target toy or performing the target role.

Results indicated that participants were generally un-accepting of exclusion in the straightforward context and used moral reasoning to justify these decisions. In the multifaceted context, children chose to include the child who did not fit the stereotype more often in the unequal experience context. In the equal experience context, more than in the unequal experience context, children did use stereotypes to justify exclusion (e.g., “girls don’t like to play with trucks”). Even in this case, the majority of children did not use stereotypes in making exclusion judgments. This research reveals the complexity of children’s reasoning about gender exclusion and suggests the important role of stereotypes in inclusion and exclusion decisions.

In a follow-up study assessing 1st, 4th and 7th graders, Killen and Stangor (2001), found a slightly different pattern. In this study, which looked at both gender and ethnicity, the children generally chose the non-stereotyped child in the equal skill context multifaceted situation, which suggests active inclusivity. These findings are different

from the Theimer, Killen and Stangor (2001) study, where the pre-school aged children primarily chose the stereotyped child in the equal context. Thus, older children may be more sensitive to issues of equity. The reasoning used in choosing the non-stereotyped child was moral reasoning, focused on issues of fairness and equal opportunity, suggesting that with age children may be more attuned issues of diversity and more willing to reject stereotypes in favor of equal access.

This body of research provided important information about the use of gender-stereotypes in making exclusion decisions in everyday contexts. Additionally, the researchers coded stereotypes when they were mentioned, but did not prompt children to consider gender stereotypes in their responses. Thus, although many children did not use stereotypes, a surprising number did spontaneously turn to stereotypic information when making these decisions. In particular, it appears that in the absence of other information, and when forced to make a decision, children are particularly likely to rely upon stereotypes about which activities are appropriate for each gender. This suggests that more research needs to be conducted on use of gender stereotypes in understanding groups. While this research suggests that children do hold stereotypes about gender-appropriate activities, it does not reveal under what circumstances children see those stereotypes as inflexible.

Additionally, in this paradigm, the group was described as participating in a gender-appropriate activity, and potential new members who were gender stereotype consistent or inconsistent asked to join the group. What is not yet known is how children will respond if a child who is already a member of the group rejects the gender stereotypic activity of the group and asserts that the group should not conform to the

stereotype. Understanding deviant behavior (rejecting a norm that your group adheres to) from within the group will provide greater insight into the range of exclusion experiences surrounding gender group membership. Additionally, what is also not yet known is how children will react if the group itself rejects the gender stereotype and engages in gender non-conformist behavior or activities. Understanding groups which both adhere to and resist gender stereotypes will improve our knowledge not just about gender stereotypes, but will have significant implications for understanding group dynamics and norms.

Further, the type of gender stereotype must be considered. This research has examined gender stereotypes about traditional social (play) activities, but there are also firm gender norms in place about a range of other types of activities and behaviors, for both girls and boys. For instance, there are strong gender stereotypes associated with types of aggressive behaviors, with early research suggesting that girls engage in more relational aggression and boys engage in more physical aggression (Crick & Grotpeter, 1995). While meta-analyses have indicated that these patterns are not as strong as once believed (Card et al., 2008) and examinations across a range of different cultural contexts have confirmed this (Lansford et al., 2012), Ostrov and Godleski (2010) recently proposed a gender-linked model of aggression sub-types, which suggests that girls and boys do still associate relational aggression with girls and physical aggression with boys. Moreover, their model suggests that children would prefer to express gender-normative forms of aggression and that they would also see gender-consistent forms of aggression as more wrong than gender-inconsistent forms. Additionally, they call for research which examines children's thinking about gender-linked forms of aggression using a variety of methodologies.

One area about which little is known, but which the current study addresses, is whether children think that one should be excluded for advocating gender-consistent or gender-inconsistent aggression. Research has shown that when considering relational aggression, in particular, adolescents reference relationship maintenance as an important consideration (Goldstein & Tisak, 2010), suggesting that further research exploring the consequences in terms of group acceptance for aggressive behavior should be explored. This is particularly important considering research that shows continuity in aggressive behavior (especially physical aggression) from childhood through adolescence (Broidy et al., 2003). Finally, this type of research could have significant impacts on how parents and educators talk with children about gender stereotypes and the rigidity of gender roles.

An additional area of concern, though, is the role of peer-influence on children's exclusionary decision-making. In social scenarios, children must weigh their own interpretation of the scenario with their expectations about the view and norms of the group. Thus, while some children did reference stereotypes in the study by Theimer and colleagues (2001), it is unclear if they would continue to apply those stereotypes when faced with peers who disaffirm stereotypes or vice versa. Thus, in order to address these additional questions, Killen, Pisacane, Lee-Kim and Ardila-Rey (2001) counter-probed children about their inclusion decisions by offering them an alternative reason to include the child they did not pick, one that reflected a peer's view about the inclusion decision. This technique tested children's conviction in their decision and investigated age differences (3 to 5 year olds, $N=72$) in assessing children's use of stereotypes in reasoning about similar inclusion scenarios as those used in Theimer, Killen and Stangor (2001).

Most children who initially used stereotypic information to make their decision changed their decisions to focus on moral judgment when given the opportunity (after a probe by the interviewer in which the fairness of turn-taking was mentioned), while those who initially focused on the moral aspects of the situation did not change their decisions to consider the stereotypic argument as often (after a probe by the interviewer in which the stereotype about who plays with toys was mentioned). The study also showed that younger children were more likely than older children to choose the stereotypic child prior to probing, and were more likely to base their judgment on stereotypic expectations.

The strong impact which probing had on reducing use of stereotypes suggests the very important influence of peers on children and the potential deficits that very young children may have in judging the beliefs and desires of those around them. This is confirmed by research on bullying, which indicates that bystanders can play a vital role in reducing aggressive behavior of their peers (Salmivalli et al., 2011). Less is known, however, about how peers who are actually part of a group that engages in stereotypic or aggressive behavior might influence the group if they resist the group's norms. Further, also assessing theory of mind ability in these studies would have been insightful as it may be the case that children who have theory of mind were better able to anticipate and take into account the potential response of the peer group. This may be especially important when studying gender-based stereotypes, as children do encounter socially-sanctioned gender stereotypes in a range of contexts.

Understanding the peer group's perspective and intentions when evaluating exclusion decisions may, thus, impact the exclusion evaluation that is made, either because children will want to align themselves with the peer group or because they may

be more skilled at recognizing that this perspective may not be the best or the right perspective at times (for instance, if it is based on stereotypes) and want to reject this peer group perspective. Thus, research on group membership and social exclusion suggests that future studies should include information about the intentions of those involved in the scenario as well as measures of ToM.

This body of research on exclusion because of group membership, including gender, reveals the complexity of children's social reasoning about exclusion. While children are, in general, unwilling to accept exclusion and advocate for inclusivity using moral reasoning, children do view some forms of exclusion as acceptable. In particular, many of the studies reviewed thus far reveal that children are more accepting of exclusion based on gender than exclusion based on race or ethnicity (Killen et al., 2007). When children view exclusion as acceptable, this may be due to a reliance upon stereotypes or assumptions about appropriate activities for members of different groups. For instance, a child might support exclusion of a girl from a football team by arguing that girls should not play certain sports (Park, Lee-Kim, Killen, Kim, & Park, 2011). Supporting exclusion of another because of his or her group membership may lead to more serious forms of prejudicial behavior and treatment of others.

While we know that children do at times support exclusion for reasons associated with stereotypes, including gender stereotypes, we do not yet have a clear picture of why they think that this form of exclusion is acceptable. On the one hand, they could be simply relying upon stereotypes and ignoring the potential harm to the victim that may occur because of this exclusion. On the other hand, they may be putting more weight on issues of group functioning and be thinking that including someone from another group

will be disruptive. Research on social exclusion should be designed to unpack these alternative explanations, as gender exclusion may often also involve stereotypes about behavior or appearance. Stereotypes could be invoked by the excluder implicitly or explicitly when the exclusion occurs, or the child who is excluded could infer a reliance upon stereotypes which may or may not be present. For instance, research with ethnic minority students reveals that even when exclusion is not explicitly about race, they are concerned that the target of exclusion may interpret the exclusion as being race-based (Margie, Killen, Sinno, & McGlothlin, 2005). Research should examine if this is also the case for exclusion in situations which invoke gender stereotypes. This is particularly important in light of findings which indicate that children with more sophisticated interpretative theory of mind skills are less likely to infer the presence of gender discrimination than those without such skills (Brown et al., 2010).

Is Exclusion Always About Group Membership?

While exclusion is often based upon group membership, as often, exclusion is based on personal traits, such as individual characteristics or features of one's personality. For instance, a group may exclude a child who is shy from the debate team because they assume that someone shy will not be comfortable in this context. Forms of exclusion not based solely on group membership are a unique area for research, as children may evaluate these forms of exclusion as more legitimate because of conventions or norms of the group. Thus, exclusion based on personality traits may be viewed by the group as legitimate. It is possible, however, that those who are excluded may struggle in interpreting the intentions of a group or individual who excludes them because of a personality trait and over-attribute negative intentions to the excluders.

Specifically, the child who is excluded from the debate team for being shy may interpret this exclusion as malicious (for instance, exclusion of her because of assumptions about girls' ability to engage in such debates) when the team may have simply intended to prevent the excluded child from experiencing what they might have assumed would be an uncomfortable environment for someone shy. Thus, it is important to examine research which has explored both group (or category) and personal reasons for exclusion.

Research has begun to examine if there are differences in how legitimate exclusion based on group membership is versus exclusion based on personality traits, such as shyness or aggression. Park and Killen (2010) conducted a cross-cultural study (Korea and the USA) that assessed exclusion due to group membership as well as individual characteristics. The sample included 10 and 13 year old children ($N = 333$ from the United States of America and $N = 397$ from Korea). The aims of the study included examining if evaluations of exclusion based on personality traits (shy or aggressive) and group characteristics (gender or nationality) vary based on social context and type of exclusion. Specifically, the study included an intergroup, interpersonal, and intrapersonal context. The intergroup context was a peer-exclusion scenario where a child is excluded from a group working on a group project in school, the interpersonal context was a friendship-rejection scenario where one child does not want to be friends with another, and the intrapersonal context involved a victimization scenario where one child is picked on because of a group or personality trait.

All participants received stories presenting all contexts and all personality and group characteristics, with the personality and group characteristics counter-balanced. Participants judged the acceptability of exclusion and provided justifications. Using just

two assessments and choosing two (judgment and justification) which have been highly validated within Social Domain research (Smetana, 2006) provided a simple, clear way to assess both group and personal types of exclusion using the same measures.

The findings included that, across all contexts and both cultures, girls were less accepting of exclusion than were boys. Additionally, the victimization context was seen as the least acceptable form of exclusion by all participants. In terms of personality and group characteristics, it was seen as most acceptable to exclude aggressive peers, and Americans were more willing than Koreans to justify exclusion of an aggressive peer. Generally, participants viewed exclusion based on group characteristics (gender and nationality) as unfair because of moral reasons. Additionally, perhaps due to greater experience with people from different nationalities, Americans were more inclusive of different nationality peers than Koreans. Older children were more likely than were younger children to endorse exclusion because of aggression and reject exclusion based on nationality. Findings suggest different degrees of acceptability by context, as well. Friendship rejection was the most acceptable, followed by group exclusion and then victimization

Thus, this research suggests that children may perceive exclusion because of group membership as less legitimate than exclusion because of personality traits. This raises the additional question, however, of how children will respond to and anticipate exclusion when group membership intersects with stereotypes about personality traits. For instance, how do children evaluate exclusion which occurs because a girl is physically aggressive, essentially engaging in a form of aggression which does not conform to gender stereotypes for girls? Further, this research does not clarify, however,

how those children who are excluded for these different types of reasons will experience the exclusion.

Nesdale and colleagues (2007) were able to test this, however, by having participants undergo a simulated exclusion experience. In this study, exclusion was either because of personal reasons (lack of drawing skill) or group membership reasons (referred to in the study as “category” reasons, i.e. being a member of a particular school). In particular, the study assessed 6 and 8 year old Anglo-Australian children ($N = 160$) using a modification of a minimal-group paradigm (the same methodology as is used in a number of studies including Nesdale, Maass, Durkin, & Griffiths, 2005; Nesdale, Maass, Griffiths, & Durkin, 2003). Participants were first assigned to a group. Then participants in the experimental condition were told that they were rejected from their group because of either personal (drawing skill) or categorical (member of a certain school) reasons. Participants in the control condition were next told that the experimenter remembered that the first group was full. Following this, participants were reassigned to a new group which either held an inclusive or exclusive group norm towards others. Finally, researchers measured participants’ affect (using the State-Trait Anxiety Inventory for Children) and attitudes towards the original, and new groups as well as an outgroup. Attitudes towards the groups were measured using three items: like, trust and desire to play with the group.

Results revealed that across the conditions, younger children expressed more negative affect than did older children and more positive group attitudes. Children in the group with the inclusive norm expressed more positive group attitudes than did children in the group with the exclusive norm. When rejected for a category reason, children’s

attitudes were more negative than when rejected for a personal reason. Thus, not only do children appear to differentiate between exclusion because of group membership and exclusion based on personal characteristics as was shown by Park and Killen (2010), but their future attitudes are also impacted in a similar way, with group membership rejection resulting in a stronger impact on attitudes. This suggests that, in fact, children who are excluded because of group membership reasons do perceive this as less legitimate than exclusion based on personality traits. Additionally, all participants who experienced rejection (as opposed to participants in the control condition), held much more negative attitudes towards the outgroup after rejection. This suggests that rejection can create a cycle of negative intergroup relations (Killen, Mulvey, et al., 2012).

These two studies reveal that personality traits or personal characteristics are seen as more legitimate bases for exclusion than is group membership, and that exclusion based on group membership can have significant impacts on outgroup prejudice and negative attitudes. However, as mentioned above, often personality traits are conflated with group membership. Children may hold stereotypic beliefs about particular groups' personality traits (for instance, children in the Park and Killen (2010) study may have thought that Korean children are also shy based on stereotypes about the reserved nature of Asians or may have thought that boys are more aggressive than girls based on stereotypes about behavior for each gender).

This is particularly concerning when examining issues of gender-conformity (Horn, 2007). While a boy may express personality traits which are traditionally considered more feminine, children may conflate this with group membership and exclude this boy based on negative stereotypes that they hold toward homosexuals (see

Horn, 2008). This would be problematic on three counts: 1) the boy may not actually identify as homosexual; 2) the group is relying on stereotypes about a particular group (homosexuals); and 3) the group is expressing prejudice towards this group. On the other hand, when a girl wishes to engage in a stereotypically male activity (for instance playing football), she may be excluded because of stereotypes about girls' strength, aggressiveness or skill in athletics. In both of these instances, children may justify exclusion based on a personality trait which is assumed to be present because of an underlying stereotype about a particular gender. Research which finds a relation in middle childhood between social acceptance and gender appropriate forms of play, particularly among boys (Moller, Hymel, & Rubin, 1992) supports the fact that underlying stereotypes about gender may play a role in how children justify inclusion and exclusion of peers. Thus, future research should aim to carefully probe children's reasoning about exclusion to attempt to determine whether underlying stereotypes or biases are involved in exclusion decisions which appear to be focused on personality traits.

Research on Conflating Personality Traits with Group Membership

Some research already has examined how children might use external cues, such as behavior or appearance, which would likely be markers of personality traits, in making judgments about group membership. Horn (2007) interviewed 10th and 12th grade students ($N = 264$) about their judgments of how acceptable same-sex peers were who varied in terms of their conformity to gender norms about appearance and activity choices and who varied in terms of their sexual orientation (homosexual or heterosexual). Participants, thus, evaluated gender-conforming and non-conforming peers who were

both straight and gay. For instance, for females, the gender non-conforming activity was football and the conforming activity was volleyball. For males, the gender non-conforming appearance was wearing eyeliner and nail polish and the gender conforming appearance was “acting and dressing like most of the other guys”.

The results indicated that appearance and activity choice had a significant impact on judgments. Both straight and gay or lesbian targets who were gender non-conformist were rated as less acceptable by their peers than were the gender conforming targets. Additionally, the boys rated the straight individual who was non-conformist as least acceptable. Thus, choice of outward appearance and activities, which might be thought of as features of one’s personality, were used by adolescents as legitimate bases for making judgments. This suggests that appearance and choice of activities may at times be the source of powerful stereotypes and that children and adolescents may turn to these stereotypes in making judgments. Additionally, while we know from this study how adolescents respond to gender conformity and non-conformity in making general judgments about others, we do not yet know how this will play out for younger children and in more complex contexts. For instance, assessing gender conformity and non-conformity in a group situation will be particularly insightful, as groups do often hold norms about behavior or appearance that are related explicitly to their identity as a group (for instance a group of girls may always play with their dolls together).

Thus, the current research will examine groups which hold either gender-conforming or non-conforming norms, in terms of appearance or activities, in order to elucidate this relationship between group membership and personality traits. Assessing children’s reasoning about deviance from gender-conformist and gender-nonconformist

groups as well as their judgments about the acceptability of exclusion will clarify if children do conflate personality traits with group membership in stereotypic ways. If so, children might, then, find exclusion of a non-conforming child from a conforming group as acceptable because this child may be viewed as part of a gender-identity outgroup due to appearance or behavior. Likewise, they may be much less accepting of groups which are gender non-conforming because of stereotypes that they may hold. Thus, while research has begun to address the differences in judgments and experiences of exclusion based on group membership or other features, more work needs to be done that examines the interplay between stereotypes about group membership and personality traits. These new lines of research may clarify what stereotypes children do hold about personality traits, norms and behaviors for different groups and how they may implicitly or explicitly rely upon these stereotypes in making exclusion decisions.

The Role of Group Membership, Norms and Identity in Exclusion Decisions

Understanding how children perceive deviance from group norms, in general, will help to explain what expectations children hold for individual members of groups. The research reviewed thus far has examined children who are excluded from groups. A significant body of research drawing on Subjective Group Dynamics (Abrams & Rutland, 2008), has examined evaluations of deviant members of groups and the group reaction to these members. Individuals can deviate (reject or depart from) from many types of group norms (or practices, beliefs or conventions of a group) and can deviate in many ways. For instance, one can deviate from an explicitly stated norm (such as a group rule) or an implicit one (for example a convention that has just arisen over time but is not stated) and can deviate through simply voicing an alternate opinion or by actually acting in a manner

contrary to the group norm. Children, who are developing a sense of self, may, at times, not want to go along with their group because they do not agree with the group practices, norms or opinions. Children commonly experience deviance and opportunities for deviance from social groups, and, in particular, opportunities for deviance from gender-based norms and expectations. Thus, research which examines children's evaluations of these types of deviance will shed light on the conditions under which exclusion because of deviance may be viewed as acceptable. This will be particularly important for examining exclusion based on gender as so many groups (formal and informal) which children are part of are single-sex groups.

In one study, focused on summer school-based groups, which are minimal in that children have not had time to form strong bonds or relationships with their ingroup members, Abrams, Rutland, Cameron and Marques (2003), surveyed children (6-7 year olds and 10-11 year olds, $N= 67$) about normative and deviant members of summer-school groups, assessing ingroup bias and favorability of normative and deviant group members. Children were told about normative group members, who made positive statements about their summer school and deviant group members, who made positive statements about both their summer school and another summer school. Children evaluated normative and deviant members of both an ingroup (their summer school) and an outgroup (another summer school). The deviant members supported the outgroup in both conditions (for instance, a deviant ingroup member expresses support for the other summer school, the outgroup).

They found that participants showed a strong ingroup bias and were more favorable towards normative than deviant group members. Additionally, participants

were more favorable towards deviant outgroup members than deviant ingroup members, especially with age. Older children were even more focused on differentiating their responses to loyal and deviant in- and outgroup members than younger children. Thus, older children seem better able to consider the implications for deviance, which leads to a greater dislike for group members who deviate from their own ingroup norms and greater like for group members who support ingroup norms. Older children, then, in some ways are able to overcome their conceptions of the outgroup, and value individual outgroup members who express deviant views, thus supporting the ingroup.

From a Social Domain Theory viewpoint, this study only assessed general social-conventional group norms (the normative and deviant group members made statements about general favorability towards the group). In social interactions, however, children are faced with a range of types of group norms, both social-conventional group norms about the customs, traditions and rules of a group, as well as moral group norms, about issues of fairness, justice and welfare. Additionally, groups can hold both positive and negative group norms (for instance, as an extreme example, a gang might hold a morally unacceptable group norm about fighting with rival gangs). Children must negotiate their social worlds by making decisions about if they feel that group norms are reasonable and if deviance from group norms, may at times, be acceptable.

Thus, Killen, Rutland, et al. (2012) assessed deviance from both moral and social-conventional group norms when the deviance either aligned with general societal principles or group-specific principles. Additionally, they evaluated exclusion acceptability more directly than had been done before in research using Subjective Group Dynamics. Finally, this study examined gender-based groups, providing insight into how

children use gender in making decisions about deviants. Participants included 4th and 8th grade children ($N= 381$). Participants evaluated vignettes in which groups excluded members who deviated from social conventional (e.g. wearing or not wearing a club shirt) or moral (e.g. equitable or inequitable distribution of funds) norms.

Analyses revealed that children and adolescents differentiate between different types of deviance: deviance that is considered morally unacceptable (advocating unequal distribution of resources) and that is considered conventionally unacceptable (not wearing an assigned group tee-shirt) is judged as wrong and, potentially, as grounds for exclusion from groups, whereas morally acceptable and conventionally acceptable deviance is judged as appropriate and exclusion because of these forms of deviance is judged as inappropriate. Age-related differences were found, with children focusing strictly on issues of fairness, while adolescents recognized a role for both fairness as well as group functioning. Additionally, analyses revealed that participants differentiated between their rating of the group favorability toward the deviant target and their own rating of favorability toward the deviant target. In all conditions except when the deviant was advocating for unequal distribution of resources, participants judged that the group would rate the deviant less positively than how the participants themselves would judge the deviant.

This research reveals the sophistication of children's reasoning about group dynamics and exclusion decisions. Additionally, it suggests that children do not always agree with their own groups and recognize that a group may hold different intentions, beliefs and goals than an individual. Generally, children did not use gender ingroup and outgroup distinctions when evaluating the acceptability of the act or of exclusion,

however, children who decided that a same gender child who does not share the norm of the group should be included in the group over an opposite gender child who shares the norm of the group relied much more frequently upon gender stereotypes than did other participants in reasoning about this decision. Thus, it appears that gender stereotypes do underlie some aspects of children's decision-making about inclusion and exclusion.

This body of research collectively reveals the importance of group norms to children and the large range of norms which influence children's social reasoning and decision-making. Generally, findings reveal that deviance from groups is considered negative: deviant ingroup members are liked less than deviant outgroup members. However, when the forms of deviance are aligned with general societal principles (for instance advocating for fair distribution of resources against a group's norm of unequal distribution of resources), deviance from group norms is not judged as negative.

The study described however, included only norms that did not relate to the group's identity in terms of group membership (i.e. the groups were divided by girls and boys, but the norms did not relate to gender identity). Research has also been conducted where the norms of the group have to do with the national identity of the group: for instance, a group of English children cheers for the English soccer team (Abrams, Rutland, & Cameron, 2003). This has not yet been done with group norms related to the gender identity of the groups.

For instance, how will deviance from a female group which has a norm that they play ballet together be viewed if the deviant wants to play football? Further, how might deviance in terms of moral behaviors which are gender-specific be seen? For instance, girls are often associated with relational aggression and boys with physical aggression

(Crick & Grotpeter, 1995). How might children respond when a girl advocates for engaging in physical aggression or a boy in relational aggression? What if a girls' group advocates for physical aggression and one member dissents and pushes for adherence to the stereotype (girls do not engage in physical aggression)? These forms of deviance from norms related to gender group membership may be viewed as positive or negative. Children who recognize the negative impacts of stereotypes may view these forms of deviance as positive as they may see the group norms as being based on stereotypes which may be holding the group back. On the other hand, children may view these forms of deviance as negative, if they hold tightly to stereotypic expectations about groups. Further, the moral principles may matter, with children supporting behaviors which are morally acceptable (resistance of physical and relational aggression) regardless of gender. Examining deviance involving norms related to forms of aggression which are gender-typed is particularly important, as research indicates that children who participate in gender non-conforming types of aggression are at risk for externalizing problems (Crick et al., 2006). Evaluations of these forms of deviance, however, are necessarily complex.

An important component of how children will evaluate these forms of deviance and exclusion will be their own social-cognitive abilities. Social relationships require balancing information about justice, group identity, group norms and societal expectations and stereotypes, as well as interpreting the social cues of those around you (Rutland et al., 2010). As children develop the ability to recognize that others may not have access to the same knowledge or that they may think and judge situations differently, they will use this information in interpreting social relationships. Thus,

understanding children's developing ability to recognize other's goals, intentions, desires and beliefs will clarify some of the age-related changes in making exclusion decisions.

Judging Others' Intentions

When evaluating an exclusion decision, understanding intentions is a central part of determining whether exclusion is legitimate or wrong. Returning to the idea that some forms of exclusion may be viewed as legitimate and others as illegitimate, intentionality could play an essential role in making different judgments. For instance, if a child has a difficult time judging intentions, he or she may not perceive that some exclusion messages may have neutral intentions and instead perceive all messages as driven by negative intentions. Thus, examining closely at what point children are able to understand intentions accurately will clarify judgments that children make, particularly in exclusion scenarios. This will be particularly important for the study of gender-based exclusion, as children often need to judge the intentions of those who are different (in terms of gender, or gender-norm conformity) than themselves.

Early research on judging intentions suggested that younger children tend to focus on outcomes, while older children can coordinate between outcomes and intentions (Piaget, 1932). Since that point, research has begun to recognize the complex development of children's abilities to understand and appreciate the mental states of others, including their intentions. Young children (between the ages of 3-5) develop theory of mind (ToM), the ability to recognize that others have desires, intentions and beliefs different than one's own (Wellman & Liu, 2004). Children follow a developmental trajectory for theory of mind abilities (Wellman & Liu, 2004), showing skill with increasingly cognitively complex forms of theory of mind with age.

Research has shown that a relation exists between early understanding of intentional actions and theory of mind competence (Wellman, Lopez-Duran, & LaBounty, 2008). Thus, by 3-5 years of age, normatively developing children should have false belief theory of mind competence, and have skills in recognizing the intentions of others. Application of these skills to social situations should enable children to accurately interpret the intentions of other individuals and groups in interactions including exclusion or rejection contexts.

Research within the field of peer rejection sensitivity and social information processing, however, reveals that not all children interpret situations the same way, even in normative populations where children should exhibit theory of mind competence. Social information processing research indicates that there is variation in how children interpret the same scenario, with some children exhibiting a hostile attribution bias, which results in them attributing hostile intentions in ambiguous contexts (Crick & Dodge, 1996; DeWall, Twenge, Gitter, & Baumeister, 2009; Orobio de Castro, Veerman, Koops, Bosch, & Monshouwer, 2002). Further, there is variation in how children interpret social situations and some children are acutely sensitive to rejection (Downey et al., 1998). Downey, et. al (1998) conducted three studies with 5th through 7th graders (Study 1: $N = 382$, Study 2: $N = 76$, Study 3: $N = 228$). In Study 1, participants completed the Children's Rejection Sensitivity Questionnaire (CRSQ), which measures how likely children are to react to ambiguous scenarios as though they were rejected or disliked, how likely they are to overreact to being rejected and how angrily or anxiously they expect rejection. In Study 2, children who had taken the CRSQ as part of Study 1 were subjected to an experimentally manipulated rejection scenario. They were told that

they could invite a friend to join them as part of the interview. Then, they were told that the child they had chosen did not want to join them. Children completed a measure of distress before and after the rejection manipulation. Children who scored as high on angry expectation of rejection were more distressed in an experimentally manipulated rejection scenario. Finally, Study 3 included self and teacher reports of aggression and victimization. This study showed that sensitivity to rejection was related to reports of greater aggression. Additionally, rejection sensitivity was shown to predict problems with teachers and peers.

This research suggests that children who are less skilled at interpreting intentions in ambiguous potential rejection or exclusion scenarios struggle in a variety of ways. These children who are more sensitive to rejection also often experience rejection. Research has also shown that peer rejected children do not differ from non-rejected children, however, in performance on the traditional theory of mind tasks (Badenes, Estevan, & Bacete, 2000). But, these children are showing some type of deficit in interpreting others' intentions and actions. Perhaps the traditional theory of mind tasks are so controlled and removed from social interactions that children who struggle with social-cognitive skills in real-world contexts are still able to succeed on these laboratory-based tasks. Thus, while even very young children begin to interpret actions as intentional, accurately judging intentions as positive or negative is a more complex task than simply knowing that an action was done intentionally or successfully passing a simple task showing that one individual holds different beliefs than another.

Rather, understanding what information children are attending to in social situations is essential. Specifically, in making an exclusion decision or interpreting a

potential exclusionary interaction, children must balance information about the goals and norms of the group, the intentions of the individual who wants to interact with the group, the intentions of the group itself and the individual members within that group, stereotypes, biases or assumptions that individuals involved in the interaction may hold, the legitimacy or illegitimacy of the exclusion decision, and the prior information that children may have about the actors involved in the exchange. Social interactions are rarely simple or straightforward. ToM is usually measured in non-social scenarios, while exclusion and peer rejection occur in rich social environments. What is not yet known is if ToM competence, when measured in a socially relevant manner, is related to ability to interpret exclusion scenarios. Some research on theory of mind has begun to try to fill this gap.

Intentionality, Theory of Mind and Social Relationships

While theory of mind is often measured using simple, lab-based contexts, it is likely that children's social-cognitive abilities do not function in the same way in controlled laboratory setting as they do in the social world, when children have to also balance information about groups, relationships and social expectations. For instance, while a child may show theory of mind competence in traditional tasks, if he or she also holds strong gender stereotypes, these cognitive constructs may create dissonance and interfere with a child's ability to judge intentions. Research is beginning to indicate that when children are making judgments about situations which involve stereotypes, including gender stereotypes, they may use those stereotypes to make attributions of intentions, regardless of their theory of mind abilities (Kelly et al., May 2010).

Additionally, research indicates that when children evaluate the emotions of a child who

desires a gender stereotype nonconforming toy, they often misinterpret the child's emotions when that child either receives the desired toy or the nondesired toy (Terwogt & Rieffe, 2003). Specifically, even when the children exhibit diverse desires theory of mind competence, they misinterpret the emotions of the child, suggesting the child would be happy to receive the stereotype consistent toy and unhappy to receive the stereotype inconsistent toy (even when the child desires the nonconforming toy). Additionally, research indicates that children with interpretative theory of mind are less likely to assume gender discrimination in ambiguous scenarios than are children without interpretative theory of mind (Brown et al., 2010). Thus, the relation between gender stereotypes and theory of mind may be complicated. How theory of mind abilities come into play, then, when gender stereotypes are activated in exclusion scenarios should be studied in more detail.

Further, Shiverick and Moore (2007) showed that children and adults can distinguish their own understanding of a character's intentions and the understanding held by an adult (a teacher) with different access to information about what has happened, and can use this information to inform their moral judgments. Specifically, in this study, participants assessed scenarios in which they had access to information about an actor's intention (positive or negative) and the outcome of an event (neutral or negative). Participants were asked to make their own moral judgments about the actor's action. Additionally, they were told about an authority figure who either had access to information about the intention only, the outcome only or both the intention and the outcome. Participants also assessed the authority figure's moral judgment (i.e. How good or bad does the teacher think the actor is?) and completed measures of second-order

theory of mind. While younger children were more likely to attribute a belief about intentions that did not match the prior information given in the story, both children and adults were shown to use information about intentions and second-order mental states in making moral evaluations.

Effectively judging others' emotions and reacting to instances of exclusion, including those involving gender stereotypes, requires effective interpretation of social situations, including interpretation of intentions. Research has begun to look at the ways in which theory of mind is measured and has tried to frame theory of mind competence within more socially-rich environments than is traditional. Often, theory of mind tasks are devoid of social information. Rather, in the traditional false contents task, for instance, a puppet or doll is described as concealing an unconventional item within a unexpected container (for instance placing Smarties candies inside a crayon box) and children are asked what another puppet or doll will think is in the container. In many cases, though, no information is provided about the puppet's motives or the relationship between the puppets. Additionally, the scenarios are usually dyadic, while many social interactions involve groups of people. Thus, research which has moved away from the traditional false belief tasks while still focusing on assessing intentions and interactions between social beings should provide a more authentic reflection of children's social cognition.

Drawing on Developmental Subjective Group Dynamics (DSGD) and previous findings that children show greater favorability to outgroup members who deviate by endorsing ingroup norms than by ingroup members who deviate from ingroup norms, Abrams et al. (2009) examined if more exclusive children and adolescents had a greater

sense of how groups function and a better ability to take the social perspective of others in intergroup contexts. Their aims were to examine whether children with greater social perspective taking abilities—which they called Theory of Social Mind (ToSM)— were more likely to exclude others, to examine if multiple classification skill (the ability to classify individuals using more than one trait or feature) led to decreased intergroup bias and to examine if greater exposure to a variety of groups led to better understanding of group norms. Finally, they aimed to identify age-related changes in these abilities and in exclusion judgments.

Abrams et al. (2009) designed two studies focused on group identity in competitive groups, namely, groups of soccer fans from Britain and France (Study 1) and two imaginary teams (Study 2). Both studies used primarily White British children (5-11 years old, Study 1: $N = 167$, Study 2: $N = 149$). Children were given scenarios about different ingroups and outgroups (soccer fans and invented Red and Green teams for Studies 1 and 2, respectively) and asked to make decisions about how much they liked each group (intergroup bias), how much they would like and thought an individual and a group would like a group member after they expressed loyal and deviant group norms (intragroup bias and understanding of ingroup bias). The multiple classification skill task asked children to group objects by their traits. The ToSM task asked children to assess a situation involving a false evaluation of another character (how would a character feel about another character who secretly stole from him). Abrams et al. (2009) found that social perspective taking was related to understanding group dynamics and, particularly, understanding of social inclusion and exclusion decisions. Interestingly, greater multiple classification skill was related to decreased intergroup bias. Greater exposure to groups

led to greater understanding of group norms. However, with age, children gain better multiple classification skill and greater ToSM. These stand in contrast to each other, because greater multiple classification skill leads to decreased intergroup bias, but greater ToSM leads to greater adherence to group norms, and thus greater exclusivity based on group norms. This study indicates future work should examine how exactly ToSM as well as multiple classification ability are used by youth. While neither of these studies explicitly examined gender-based exclusion, the findings suggest that applying socially-relevant forms of theory of mind to scenarios involving gender exclusion may provide greater information about why children willingly reject straightforward exclusion based on gender, but condone gender exclusion in more complex contexts. Perhaps a focus on group perspectives and group dynamics leads to a prioritizing of the group goals over the potential harms of exclusion because of gender.

Research examining morally-relevant theory of mind (Killen et al., 2011), which is an adaptation of the false belief tasks using scenarios which involve authentic social interactions between a transgressor and a victim, provides more evidence that theory of mind should be measured in a socially relevant manner and that it would be an important variable for studies of gender exclusion. Specifically, in this study, participants aged 3-8 years ($N = 162$) assessed prototypic and accidental transgressions as well as morally-relevant and traditional theory of mind tasks. In the accidental scenario, participants evaluated a situation in which an accidental transgressor threw out a cupcake, left in a bag by the protagonist (who went outside). In the prototypic scenario, participants evaluated a situation in which a deliberate transgressor pushed a victim off of a swing. The novel findings pertained to significant patterns between moral reasoning and theory

of mind (when age was covaried out). Participants with false belief ToM evaluated the accidental transgressor's intentions as significantly more all right than they themselves evaluated the act. Participants without false belief ToM did not differ in their judgment about the transgressor's intentions and the act itself, judging both as wrong. All children judged the prototypic transgressor's intentions and act as wrong. Though children with false belief ToM evaluated the accidental transgression as wrong, they rated it more acceptable to punish a transgressor in the prototypic scenario than in the accidental scenario. Participants without false belief ToM did not differentiate punishment between the scenarios.

The findings demonstrate that theory of mind is necessary for a comprehensive evaluation of moral transgressions. Without false belief ToM, children are more likely to attribute negative intentions to an "accidental" transgressor than when children have false belief ToM. This error by children may contribute to interpersonal conflict given that misattributing negative intentions to others accounts for a large proportion of peer conflict. Additionally, these more socially attuned measures of theory of mind suggest that there are greater relationships between theory of mind ability and ability to make moral judgments (including those about exclusion) and emotion judgments than previously thought and that when theory of mind tasks are embedded in morally-relevant contexts (involving a transgressor and a victim) theory of mind does not function in the same manner as when theory of mind is measured in a socially-removed manner as it is traditionally measured.

Thus, these studies suggest that theory of mind is an important cognitive skill to measure when studying exclusion. Understanding group goals and intentions aids

children in evaluating social situations. Children with greater theory of social mind may be more exclusive due to their greater focus on group goals and group functioning. Additionally, children may struggle in interpreting the emotions and intentions of different actors in social exchanges which can impact their evaluations of a scenario. What is not yet known is how children interpret collective (group) excluders and balance their interpretation of the intentions and goals of both the excluders and those who are excluded. ToSM, and morally-relevant ToM all move towards more socially-relevant ToM measures, however socially-relevant ToM needs to be tested more systematically, drawing on the strengths of these new ToM measures. Socially-relevant ToM should relate to the ability to interpret exclusionary scenarios and to recognize that some forms of exclusion are warranted while others are unjustifiable. Socially-relevant ToM will be particularly insightful for studies of gender exclusion as exclusion based on gender so often does involve norms about the behavior or appearance of each gender or stereotypes about typical activities or personality traits for each gender. One's ability to take the perspective of another or of a group in situations where complex gender dynamics are at play may be impeded, thus, measuring theory of mind and other social-cognitive abilities within authentic social contexts will provide a better sense of how social cognition plays a role in exclusion decisions surrounding gender.

Yet another way in which research has moved towards assessing a more socially-rich form of theory of mind is through interpretative theory of mind (Carpendale & Chandler, 1996), which requires that one recognize that when given the same information, two people can come to equally likely conclusions. This form of theory of mind has been found to develop later than false belief understanding, typically by 8 years

of age (Carpendale & Chandler, 1996), though some findings suggest even more prolonged development (Mull & Evans, 2010). Interpretative theory of mind requires that one consider social interactions and relationships to a greater degree than false belief theory of mind, however, it is still not fully socially-contextualized. Research using interpretative theory of mind, however, suggests that this may be a particularly appropriate measure to use when examining use of gender stereotypes, as it has been shown that children with interpretative theory of mind are less likely to make accusations of gender discrimination in situations in which discrimination may have occurred (Brown et al., 2010). They may have judged that while it was possible that gender discrimination had occurred, it was inappropriate to jump to conclusions about discrimination if the situation was somewhat ambiguous.

Thus, some research suggests that theory of mind competence aids one in recognizing intentions and using this information to make moral judgments (Killen et al., 2011); some research suggests that having theory of mind competence is related to greater exclusivity due to a better understanding of group dynamics (Abrams et al., 2009); and some suggests that having theory of mind is related to being less willing to perceive gender discrimination in situations where it may have occurred (Brown et al., 2010). This suggests that social-cognitive abilities such as theory of mind are intimately involved in moral decision-making, including situations involving group membership, such as gender. Further, the relation between such judgments and theory of mind is not yet clear. Research should aim, then, to continue unpacking the connections between social-cognitive capacities and moral decision-making, particularly in situations involving exclusion.

Conclusions

The research reviewed here reveals that social cognition about exclusion/inclusion decisions is multi-faceted. Children can reason in complex ways about exclusion, including exclusion based on gender, and children must weigh information about group membership, the goals of the group, the reasons for exclusion and the norms of the group in making exclusion decisions. Recognizing the importance of reasoning as well as interpreting intentions is essential in studying exclusion. This review revealed that, while children do reject many forms of exclusion, they are more willing to accept gender-based exclusion than other forms of group-based exclusion, such as race or ethnicity. Additionally, as reviewed, research indicates that children and adolescents do differentiate between forms of exclusion, viewing some forms of exclusion as acceptable because of social-conventions. Further, they distinguish between different forms of deviance from groups, seeing some forms of deviance positively, but do not condone exclusion of any deviant group members. The bulk of the research that has been conducted on exclusion, however, has assessed acceptability of exclusion. What is still unknown is how *likely* children think that exclusion is in different contexts. Further, what is still not known, however, is how children evaluate excluding someone or the likelihood of excluding someone who is already part of a group when they deviate from the social-conventions or moral practices of that group which surround gender identity.

Thus, future research should continue to examine reasoning about exclusion in complex peer interactions. This will provide greater insight into when and under what circumstances children view exclusion and exclusion messages as legitimate and illegitimate. Understanding when children view exclusion as legitimate will aid

educators who work to resolve conflicts between children because it will provide them with tools for communicating with children when exclusion occurs and will guide them in finding ways to help children to only exclude in circumstances where it will not harm others and where it is based upon legitimate concerns. This will be especially helpful in addressing issues of gender exclusion, where children may receive contradictory messages from society, parents and peers about appropriate behaviors, activities and appearance for each gender.

Additionally, while research has revealed that group norms related to group identity can be important in a nationality context, what is not yet known is how such norms would play out when they relate to the gender group identity of the groups. Thus, research should also examine deviance from group norms that relate to gender group membership, in particular research should focus on gender stereotypes about aggressive behavior, as group responses to deviance in terms of gender stereotypes about aggression is an understudied, but important area. For instance, future research might examine group norms about what it means to be a girl or boy, for instance, and how children respond to deviance from group norms when the norms themselves either conform to traditional stereotypes about gender or move against traditional stereotypes about gender. This is an important new direction for research because of the strong ingroup biases that surround gender as well as the early emergence and pervasiveness of negativity towards those who do not conform to gender stereotypes. Understanding how children respond to deviance from groups which conform to or resist gender stereotypes, especially those involving aggressive behavior, will aid educators and counselors managing the complex conflicts that children have surrounding the development of their gender identity.

Finally, research needs to more fully explore the relation between exclusion decisions and understanding of intentions using more finely tuned measures of socially-relevant theory of mind. This new avenue of research will be particularly insightful as some research is beginning to reveal the complex and surprising ways that theory of mind competence is related to a range of types of moral judgment as well as how knowledge of and use of stereotypes may impact applications of theory of mind competence.

Unraveling the relation between children's social-cognitive abilities and their judgments will provide essential information that can help guide parents and educators as they find ways to encourage their children to take others' perspectives and to fully assess the motives, desires and emotional states of those involved in potential conflicts. This will be particularly important to study within the context of gender conformity and non-conformity and exclusion due to gender because social-cognition within this context involves not only an assessment of others' beliefs, intentions and desires, but also an assessment of group norms regarding gender identity and societal stereotypes about gender.

Thus, within this review, new areas of research have been identified which, together, will help move research towards a more complete understanding of social exclusion, with the aim of better informing intervention programs designed to ameliorate the often detrimental consequences of experiencing social exclusion in school contexts.

Chapter III

Methodology

Participants

Participants ($N = 292$) included 90 9-10 year olds ($M = 9.63$ $SD = 2.99$, $Range = 9.40$ years to 11.61 years), and 202 13-14 year olds ($M = 13.95$ $SD = .43$, $Range = 13.05$ years to 15.88 years) from public elementary and middle schools in the Mid-Atlantic region. Given the varied analyses of interest, and expecting medium effects at best, and with the desire to achieve power levels of .80, the a priori power analysis for a 2 group, 2-tailed ANOVA test indicated an appropriate sample size would include at least 128 participants. Participants were approximately evenly divided by gender (52.4% female), and were ethnically representative of the United States (school demographic information identified approximately 30% ethnic minority students in the schools). Further, school demographic information indicates that participants were from low to low-middle income schools. Only students receiving parental consent (9-10 year olds) and providing student assent (all participants) completed the Tasks (see Appendix A for Institutional Review Board Approval and Consent forms).

Design

The study involved between-subjects and within-subjects factors for an overall design that includes a 2 (Age Group: 9 - 10 and 13 - 14 years) X 2 (Gender: female, male) X 3 (Condition: activities, relational aggression, physical aggression) model with repeated measures on the last factor. Analyses included subsets of these variables, to test specific hypotheses.

Three tasks were administered. The *Gender Exclusion Task* included 2 versions. Each version included 3 conditions (neutral social activities, relational aggression, and physical aggression). Each condition included 2 scenarios: one for a girls group and one for a boys group. The scenarios varied in terms of the conformity norm, depending on the version. Specifically, Version 1, (CRC: Conform, Resist, Conform) included groups which conform to the stereotype for neutral social activities and for physical aggression, but which resist the stereotype for relational aggression. Version 2 (RCR: Resist, Conform, Resist) included groups which resist the stereotype for neutral social activities and physical aggression, but which conform for relational aggression. The six scenarios in each version (12 total) were varied in order to systematically examine differences in reasoning about resisting group norms depending on the type of norm (conforming or not), the domain of the norm (moral or societal), and the condition (physical aggression, relational aggression, neutral social activities) (see Figure 1 for task design). Additionally, as male and female participants at both age groups evaluated both versions, differences based on age, gender, and ingroup or outgroup status were examined.

The neutral social activities were either: football (stereotypic male activity) or ballet (stereotypic female activity). The relational aggression scenarios were: gossip; speaking about non-present other peers (stereotypic female behavior), or impartiality; not speaking about non-present other peers (stereotypic male behavior). The physical aggression scenarios were: rough; pushing and shoving in a soccer game (stereotypic male behavior), or nice; playing nicely in a soccer game (stereotypic female behavior). See Figure 1 for task design.

For each scenario, participants evaluated a member of the group who disagrees with or dissents from the group. This resisting member either adheres to or resists a stereotype, depending on if the group is conforming or non-conforming to the stereotype.

An Interpretative Theory of Mind Task (Carpendale & Chandler, 1996) and a *Gender Stereotype Task* (modified from Signorella et al., 1993) were administered to assess participants' theory of mind competence and their adherence to gender stereotypes.

Procedure

The three tasks were administered by a trained researcher in a quiet room at each school. Participants were told that there are no right or wrong answers and that all responses are anonymous and confidential. Additionally, participants were told that their participation is voluntary and that they may choose to stop the assessment at any time. Participants were also given a warm-up task, which involved practicing using the Likert scale to be used in the survey. For 9 - 10 year old participants, the survey was read aloud by a trained researcher to small groups (3 - 4 participants) of participants of the same gender. For 13 - 14 year old participants, the survey was administered by a trained researcher to larger groups (25 - 30 participants). The necessity to read the survey aloud to the younger participants accounts for the difference in sample size between the 9 - 10 year old participants and the 13 - 14 year old participants. For both age groups, participants recorded their answers. Any questions the participants had were answered by the researcher. The survey took about 40 minutes to complete.

Measures

Participants completed three Tasks (see Appendix B). Participants first completed the *Gender Exclusion Task*, followed by the *Interpretative Theory of Mind Task*, and then

the *Gender Stereotype Task*. The *Gender Exclusion Task* was modified from Killen, Rutland, et al. (2012). The coding categories for the *Gender Exclusion Task* were also adapted from Killen, Rutland, et al. (2012). The *Interpretative Theory of Mind Task* and the coding system were modified from Carpendale and Chandler (1996). The *Gender Stereotype Task* was modified from Signorella, Bigler, and Liben (1993).

Gender Exclusion Task

The *Gender Exclusion Task* consisted of twelve hypothetical scenarios (6 in each of 2 versions) in which a member of a group disagrees with his/her group about the group's norm. Each child heard only six scenarios. There were two scenarios for each of three conditions: social activities, relational aggression, and physical aggression. For each condition, the two scenarios included opposite group norms and one of the scenarios was about a girls group and one was about a boys group. For instance, for the relational aggression scenario, participants might have responded to a girls group with a norm of gossiping and a boys group with a norm of impartiality. As described above, there were two versions. Version 1 included groups which conform to the stereotype for neutral social activities and for physical aggression, but which resist the stereotype for relational aggression. Version 2 included groups which resist the stereotype for neutral social activities and physical aggression, but which conform for relational aggression. Note that this design is premised not upon the belief that there are differences in mean rates of aggression between girls and boys, but rather that there are stereotypes associating girls with relational aggression and boys with physical aggression. Approximately equal numbers of male and female participants from each age group completed each version. Females received surveys with the girls groups labeled as "your group" and the boys

groups labeled as “their group.” Males received surveys with the boys groups labeled as “your group” and the girls groups labeled as “their group.”

Dependent Measures for the Gender Exclusion Task. For each scenario in both versions, the same assessments were given, to allow for a direct comparison between scenarios. The first assessment was: 1) *Likelihood of resistance*: What do you think the dissenter will do? Participants had a dichotomous choice: Go along with the group, or tell them what he/she thinks. This measure assessed if participants think that individuals who disagree with their group will express their disagreement to the group. The second assessment was: 2) *Individual likelihood of resistance*: What would you do? Participants had a dichotomous choice: Go along with the group, or tell them what he/she thinks. This measure assessed if participants themselves would be willing to express disagreement with the group.

After this assessment, participants were told that the dissenting member chose to tell the group his/her thoughts. Participants then rated: 3) *Group favorability, dissenting member*: How okay or not okay will they think what she says is? This assessment was measured with a Likert scale with ratings from 1 = really not okay to 6 = really okay. This measure assessed how favorable participants think the group will be towards someone who is a member of the group, but vocally disagrees with the group. This was followed by an assessment of 4) *Reasoning* in order to determine the reasons why participants believe a group will feel either favorable or not favorable to a dissenting member. Reasoning data were coded using the Coding System, See Appendix C.

Next, in order to compare how participants’ feel about this dissenting member to how participants think groups will feel about this dissenting member, the survey included

a measure which asked participants for their own perspective. Specifically, they assessed: 5) *Individual favorability, dissenting member*: When you hear her, how okay or not okay do you think what she says is? This item was measured with a Likert scale from 1 = really not okay to 6 = really okay. This assessment provided participants' own individual favorability towards the dissenting member. Participants also completed 6) *Reasoning*. Reasoning was assessed in order to determine the reasons why participants feel either favorable or not favorable to a dissenting member. Reasoning data were coded using the Coding System, See Appendix C.

Next, participants were told that the group must decide how to respond to the deviance. Participants then assessed 7) *Intragroup exclusion likelihood, dissenting member*: "Do you think the group will tell her she can't be in the group anymore?" Participants first responded with a dichotomous choice between yes and no, followed by a Likert scale from 1 = very unlikely to 6 = very likely). This assessed perceptions of repercussions for dissent. Specifically, it assessed if participants think that dissenting group members will be excluded from within their group, based on the group member's dissent. Participants also completed 8) *Reasoning*. Reasoning was assessed in order to determine the reasons why participants believe that exclusion is either likely or not likely. Reasoning data were coded using the Coding System, See Appendix C.

Finally, participants were told that the group can invite one more person to join their group and were asked to assess 9) *Intergroup inclusion preference*: "Who should the group include?" They had a dichotomous choice between someone who agrees but is the opposite gender, and someone who disagrees but is the same gender. This assessed whether the participant shows any bias in terms of gender or norm, and what they

determined to be more important for the group: shared gender or shared beliefs.

Additionally, 10) *Reasoning* was assessed. Reasoning data were coded using the Coding System, See

Appendix C.

Coding categories for justifications. A Coding System was established based on extensive pilot testing, and drawing on prior research (Killen et al., 2001; Killen, Rutland, et al., 2012). The Coding System included three broad categories, based on social domain theory: Moral, Societal and Psychological. The subcategory for Moral was: Concern for Other's Welfare (e.g. "If they gossip, they will hurt the feelings of the other kids." or "Pushing and shoving could mean that someone could get badly hurt." or "It is not fair for them to tell her she cannot be in the group just because she does not think that girls should do ballet."). The subcategories for Societal were: Group Functioning (e.g., "Playing nice will mean that they can all work together to win the game."), Inclusion of Diverse Perspectives (e.g., "She can teach them to all play nicely" or "it is better to include someone different") and Gender Group Identity or Stereotypes (e.g., "Girls always gossip." or "Well, it is a boys group, so they should pick another boy."). The subcategories for Psychological were: Autonomy (e.g., "It is up to her what game she wants to play."), and Identification with the Target (e.g., "I like football, too" or "Well, I also wouldn't want to gossip."). Reasoning was identified as Uncodeable if it is undifferentiated (e.g., "It's bad."), inconsistent with the story, or incomplete. Interrater reliability was high, Cohen's $\kappa = .92$.

Interpretative Theory of Mind Task

In order to assess participants' theory of mind skills, the Interpretive Theory of Mind Task (Carpendale & Chandler, 1996) was used. Interpretive theory of mind assesses whether participants can recognize that two people who had access to the same information may come to different conclusions or interpretations of that information. It typically develops by age 7 or 8 (Carpendale & Chandler, 1996), although some recent studies have found prolonged development past age 7 (Mull & Evans, 2010). The task used is modeled upon the Carpendale and Chandler (1996) measures, with two changes. The measure was administered as part of a survey, but originally included puppets and was an interview. Participants first identified if they think that two people can come to different conclusions about the same information, as in the original task. In the original task they were then asked about a third individual. As some of the hypotheses of the current study involve understanding the perspective of a group, participants were not asked about a third individual, but rather about a group of same-gendered peers and their interpretation of the situation. This provided a sense of if participants could accurately interpret the perspective of a group of peers.

The *Interpretive Theory of Mind Task* involved 3 short scenarios. In the first one, participants were told about two same-gendered children playing a game. For instance, in the female version of the story, participants were told that: "Jill and Anna are playing a game. They are supposed to "wait for a ring" before they take the next turn. Jill says they should wait for the telephone to ring. Anna says they should wait for a ring that you wear." In the second scenario, the participants were shown the classic "Duck-Rabbit" Image (Jastrow, 1899). They were told: "Jill and Anna see this picture. Jill says it is a duck. Anna says it is a rabbit." In the third scenario, participants were shown 3 images

of cards. Card 1 has a blue block on it. Card 2 has a red triangle on it. Card 3 has a red block on it. They were told: “Jill and Anna need to find a penny, which is hidden under one of these three cards. The penny is under the card with the block on it. Jill says it is under card 1. Anna says it is under card 3.”

Dependent measures for the theory of mind task. For each of the three scenarios, the same assessments were completed. The first assessment for each scenario was 1) *Explanation*: Is it okay for Jill to say X and Anna to say Y? and 2) *Reasoning*. For the *Explanation* assessment, participants were scored a 0 if they either answered “No” to Question 1 or did not provide reasoning for Question 2 which recognized that one could see either interpretation as legitimate. They received a score of 1 if they responded “Yes” to Question 1 and also provided reasoning for Question 2 which recognized that either interpretation was legitimate. The next assessment was 3) *Group Prediction*: Now a group of girls comes over and sees X. What will they say, X, Y, or would you not know what they would say? And 4) *Reasoning*. For the *Group Prediction* assessment, participants were scored a 0 if they did not answer that they would not know what they would say to Question 3 or did not provide reasoning for Question 4 which recognized that one could see either interpretation as legitimate. They received a score of 1 if they responded that they would not know what the group would say to Question 3 and also provided reasoning for Question 4 which recognized that either interpretation was legitimate. For the scoring of the Interpretative Theory of Mind measure, Interrater reliability was high, Cohen’s $\kappa = .85$. Scores for responses were summed across the 3 scenarios (for a score of 0 = no interpretative theory of mind to 6 = full interpretative theory of mind).

Gender Stereotype Task

The *Gender Stereotype Task* involved assessing gender stereotypes surrounding each behavior or activity in the survey: football, ballet, gossiping, being impartial, playing nice, and playing rough by pushing and shoving. For each activity or behavior, participants completed stereotype measures.

Dependent measures for the gender stereotype task. The first assessment, *Stereotype Awareness*, read 1) “Who usually does X?” with the choices of boy, girl, or both provided. This assessment provided information as to the participant’s awareness of gender stereotypes for each of the behaviors and activities. The second assessment was 2) *Peer group experience*: “How many of your friends do X?” with the choices of none, a few, some, or most. This assessment provided information about how commonly the peer group engages in the behavior or activity. The third assessment, 3) *Personal Experience*, was only measured for the activities (ballet and football) because of the moral valance associated with the behaviors, and the potential for answers reflecting social desirability. This assessment read 3) “Do you do X?” with a dichotomous choice of Yes or No. This assessment provided information about the participant’s own experience with the activities. Finally, participants were asked 4) *Peer Group Gender*, “How many of your friends are X (same sex as participant)?”, with choices none, a few, some or most.

Plan for Analysis

Data were analyzed using Analysis of Variance (ANOVA) and repeated-measures ANOVAs to test hypotheses for between group differences, using age, gender, interpretative theory of mind competence, and stereotype awareness as the between group factors. The repeated-measures factors included stereotype norm (conforming or non-

conforming), condition (physical aggression, relational aggression or afterschool activity) or type of assessment, depending on the specific hypothesis. Follow up tests were conducted using the Bonferroni Correction to control for Type I errors. Dichotomous responses were coded 0 or 1. For ease, condition refers to the deviant group member's behavior. For instance, in the ballet condition, the deviant member advocates for doing ballet when the group wants to do football. Justifications were proportions of responses for each respective coding category, with the top three justifications analyzed for each question. Regression analyses were run on the *Interpretative Theory of Mind* score, to assess if differences in theory of mind competence are related to differences in evaluations of challenging gender stereotypic peer group norms.

Results

Gender Stereotypic Activities

Stereotype Measure

Do children and adolescents hold gender stereotypes about who usually plays football or does ballet? Descriptive statistics indicate that both children and adolescents, and boys and girls hold strong stereotypes about these behaviors, with over 75% of participants of each gender and age group affirming these stereotypes (See Table 2).

Peer Resistance to Group Norms about Gender Stereotypic Activities

Since results confirmed that children and adolescents do hold stereotypes suggesting that ballet is an activity usually done by girls, and football by boys, the next question was: do children think their peers will resist these stereotypes and challenge their groups? In order to address this question, two separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) ANOVAs were conducted,

one for the ballet and one for the football deviance conditions. For the ballet condition, when a group member wants to do ballet when the group wants to play football, an age interaction was found, $F(1,279) = 14.30, p < .001, \eta^2 = .04$, revealing that 9-year-olds ($M = .76, SD = .43$) were more likely to expect the group member to resist the group than were 13-year-olds ($M = .56, SD = .50$). Additionally, there was a version interaction effect, $F(1,279) = 48.69, p < .001, \eta^2 = .14$, which showed that participants were more likely to expect that a girl who wanted to do ballet when the rest of her girls' group wanted to play football would challenge the group ($M = .81, SD = .39$) than would a boy who wanted to do ballet when the rest of his boys' group wanted to play football ($M = .43, SD = .50$).

For the football condition, an age interaction was also found, $F(1,282) = 11.73, p = .001, \eta^2 = .04$, revealing that 9-year-olds were more likely ($M = .90, SD = .30$) to expect that the group member who wanted to play football when the group wanted to do ballet would tell the group than were 13-year-olds ($M = .72, SD = .45$). Thus, the results indicate that children are more likely to expect their peers to challenge the group about playing both football and ballet than are adolescents.

Next, in order to test the hypotheses that participants would be more likely to expect a group member would challenge the group to play football than ballet, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Condition: football, ballet) ANOVA with repeated measures on the last factor was conducted. A main effect for condition was found, $F(1,279) = 19.35, p < .001, \eta^2 = .06$, revealing that participants expected resistance would be more likely in the football ($M = .78, SD = .42$) than ballet condition ($M = .62, SD = .49$). This was driven by a condition

by version interaction, $F(1, 279) = 21.79, p < .001, \eta^2 = .07$, which revealed that participants perceived resistance to be equally likely when the group member was a boy who wanted to play football ($M = .80, SD = .39$) or a girl who wanted to do ballet ($M = .81, SD = .39$). However, as expected, when a girl wanted to play football ($M = .74, SD = .44$), participants were much more likely to expect that she would speak up and tell her group than when a boy wanted to do ballet ($M = .43, SD = .50$). Results revealed that participants may hold a shifting standard: they asserted that it is equally likely for children to speak up and challenge their group when the group member's desired activity aligns with stereotypes, but when that desire is stereotype non-conforming, participants were more likely to expect that a girl would challenge the group to play football than that a boy would challenge the group to do ballet.

Individual Resistance to Group Norms about Gender Stereotypic Activities

Children and adolescents do expect that their peers will challenge the group, but this is dependent on the nature of the challenge. It is harder for boys to challenge the group in gender non-conforming ways than it is for girls. The next question is do these gender stereotypes that regulate expectations about behavior influence children's and adolescents' own assertions about challenging the group? In order to answer this question, two separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) ANOVAs were conducted for individual likelihood of resistance in the ballet and football deviance conditions. Individual likelihood of resistance measured if participants thought they would be likely to tell the group that they disagreed with the group's chosen activity.

For the ballet condition, a version interaction effect was found, $F(1, 277) = 32.747, p < .001, \eta^2 = .10$, revealing that participants were more likely to challenge the group when challenging the group conformed to gender stereotypes ($M = .92, SD = .27$) than when it resisted such stereotypes ($M = .59, SD = .49$). Thus, participants, themselves, were influenced by gender stereotypes regarding ballet in similar ways to how they expected their peers to be influenced.

Next, in order to confirm that participants were also more likely to resist the peer group in order to play football than to do ballet, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Condition: football, ballet) ANOVA with repeated measures on the last factor was conducted on individual likelihood of resistance. An effect for condition was found, $F(1,277) = 25.51, p < .001, \eta^2 = .08$, confirming expectations that participants were more likely to challenge the group in order to play football ($M = .91, SD = .29$) than to do ballet ($M = .75, SD = .43$). Further, a condition by version interaction effect was found, $F(1,277) = 24.96, p < .001, \eta^2 = .08$, revealing that this difference was driven by differences in how participants evaluated challenging a group by resisting stereotypes. There were no differences between how they evaluated telling a girls group you want to do ballet ($M = .92, SD = .27$) or a boys group that you want to play football ($M = .92, SD = .27$). However, participants were more likely to challenge a girls group to play football ($M = .89, SD = .31$) than a boys group to do ballet ($M = .59, SD = .49$). Thus, stereotypes regarding boys doing ballet are especially strong and pervade participants' own expectations about if they would challenge their peer group.

Relationship between Peer and Individual Resistance to Group Norms about Gender Stereotypic Activities

Football. It was expected that participants would be more likely to resist the group by advocating for playing football than they would expect their peers to resist. In order to test this hypothesis, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Question: peer, individual resistance) ANOVA with repeated measures on the last factor was conducted for the football condition. Results revealed that participants were more likely to challenge the group to play football than they expected their peers to resist the group, $F(1,279) = 8.92, p < .01, \eta^2 = .03$. Further, a question by age group interaction was found, $F(1,279) = 11.17, p = .001, \eta^2 = .03$, revealing that while 9-year-olds did not differentiate between their own likelihood of resistance ($M = .90, SD = .30$) and their expectations for how likely their peer would be to resist the group ($M = .91, SD = .29$), 13-year-olds did differentiate (Peer Resistance: $M = .72, SD = .45$, Individual Resistance: $M = .91, SD = .43$). Thus, younger children may have more difficulty distinguishing between their own perspective and the group's perspective than do adolescents. Finally, a gender by question interaction was found, $F(1,279) = 4.932, p < .05, \eta^2 = .01$ which revealed that while female participants asserted that they would be more likely to challenge their group to play football ($M = .93, SD = .25$) than they expected their peers would ($M = .76, SD = .43$), $p < .001$, there were no differences for male participants (Peer Resistance: $M = .80, SD = .40$, Individual Resistance: $M = .87, SD = .33$). This suggests that female participants may believe that others adhere more strongly to stereotypes about football than they, themselves do.

Ballet. In order to assess if similar differences between individual and peer resistance to group norms were present for the ballet condition, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Question: peer, individual resistance) ANOVA with repeated measures on the last factor was conducted for the ballet condition. As in the football condition, results indicated that participants expected that their peers would be less likely to resist the group norm in order to advocate for doing ballet ($M = .62, SD = .49$) than would they, individually ($M = .75, SD = .43$), $F(1,275) = 8.76, p < .01, \eta^2 = .03$. Further, a question by age group interaction was found, $F(1,275) = 4.55, p < .05, \eta^2 = .01$, which showed that 9-year-olds did not differentiate between peer ($M = .76, SD = .43$) and individual ($M = .81, SD = .39$) likelihood of resistance, but that 13-year-olds did differentiate (Peer Resistance: $M = .56, SD = .50$, Individual Resistance: $M = .73, SD = .45$). Adolescents may be more likely to expect that their peers will be influenced by stereotypes than do children. A question by gender interaction was also found, similar to in the football condition, $F(1,275) = 8.85, p < .01, \eta^2 = .03$. Just as was found in the football condition, females differentiated more between individual ($M = .81, SD = .39$) and peer ($M = .60, SD = .43$) responses to challenging the group by doing ballet than did males (Peer Resistance: $M = .65, SD = .48$, Individual Resistance: $M = .69, SD = .46$). Perhaps females have more personal experience with stereotypes limiting their opportunities and thus are both more attuned to the likelihood that others will use stereotypes as well as more likely to resist these stereotypes themselves.

Group Favorability Towards the Deviant Group Member

After participants assessed the likelihood of resistance, they were told that the group member who disagreed with the group did actually deviate from the group, telling the group that he or she wanted to do a different activity. The next set of hypotheses involved how the group would respond to this deviant member. Do children and adolescents think that groups will dislike group members who want to engage in different gender stereotypic activities and does the gender of the group and the type of activity matter? It was expected that groups would not like any members who deviated from the group, but that they would be particularly negative towards members who deviated from the group and also challenged gender stereotypes (for instance a girl who wanted to play football and a boy who wanted to do ballet). Further, it was expected that groups would be least favorable towards a boy who wanted to do ballet, based on prior research indicating that gender non-conformity by boys is viewed as especially wrong (Horn, 2007). In order to test these hypotheses 2 separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) ANOVAs were conducted for group favorability in the ballet and football deviance conditions. Results for the football condition revealed a version by gender interaction, $F(1,281) = 7.62, p < .01, \eta^2 = .02$, which showed that while female participants did not differ in their expectations for group favorability of a boy who wanted to play football and a girl who wanted to play football (Boy: $M = 3.47, SD = 1.42$, Girl: $M = 3.67, SD = 1.42$), male participants did differ. Male participants expected that groups would be more favorable to a boy who wanted to play football when his group wanted to do ballet than a girl who wanted to play football when her group wanted to do ballet (Boy: $M = 4.15, SD = 1.46$, Girl: $M = 3.29, SD = 1.29$). Pairwise comparisons revealed that male and female participants differed

significantly in their responses to the boy who wanted to play football, $p < .05$ and that male participants differed significantly in their responses to the boy and girl who wanted to play football, $p < .01$.

For the ballet condition, a version effect was found, $F(1,280) = 63.20, p < .001, \eta^2 = .18$, revealing that participants expected groups to evaluate the girl who wanted to ballet much more positively than the boy who wanted to do ballet (Boy: $M = 2.46, SD = 1.51$, Girl: $M = 3.73, SD = 1.23$). Thus, while participants generally expected that groups would not like deviant members, they were attuned to the stereotypes associated with the deviant behaviors. All participants differed in their evaluations of a deviant member who wanted to do ballet, expecting that groups would be more favorable to a girl who wanted to do ballet when the group wanted to do football than a boy who wanted to do the same thing. Further, male participants allowed stereotypes to influence their evaluation for the football condition, expecting that groups would be more favorable to a boy who wanted to play football than to a girl who wanted to play football.

In order to assess if children and adolescents expected groups to differ in their evaluations of a deviant member who wanted to play football or do ballet, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Condition: ballet, football) ANOVA with repeated measures on the last factor was conducted. This revealed that participants expected groups would be more favorable to the deviant who wanted to play football than do ballet, $F(1,279) = 22.70, p < .001, \eta^2 = .07$ (Football: $M = 3.63, SD = 1.43$, Ballet: $M = 3.08, SD = 1.52$). Additionally, a condition by version effect was found, $F(1,279) = 23.08, p < .001, \eta^2 = .07$, indicating that participants expected groups would be more favorable to the girl who wanted to play

football than to the boy who wanted to do ballet, $p < .001$, but that they did not differ in their evaluations of the boy who wanted to play football and the girl who wanted to do ballet. In summary, children and adolescents generally did not think that groups would like these deviant members, especially if they were deviating by also suggesting the group engage in a counter-stereotypic activity.

Justifications for Group Favorability Towards the Deviant Member

The next question involved reasoning: do children and adolescents expect that groups will use different reasons to justify how favorable they are towards these deviant members? Analyses were conducted in order to test hypotheses concerning differences in children's and adolescents' reasoning about the group's favorability towards the deviant member by participants who thought the group would like versus would not like the deviant member. Specifically, participant responses to group favorability towards the deviant member were divided into a dichotomous variable (okay, not okay) using a mid-point split of 3.5. The top four forms of reasoning used by participants to reason about group favorability towards the deviant member were group functioning ("he isn't going along with what they want to do"), autonomy ("it's okay to be different and want to do something unique"), inclusion of diverse perspectives ("she can help them understand that football could be fun"), and gender identity or stereotypes ("they'll think he is gay for suggesting they do ballet"). Two separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Group Favorability: okay, not okay) X 2 (Justification: group functioning, autonomy, diverse perspectives, gender) ANOVA with repeated measures on the last factor were conducted, one for each condition (football and ballet).

Football. When the deviant member wanted to play football, participants used different forms of reasoning if they thought that the group would like the deviant member than if they thought the group would not like the deviant member. Specifically, an interaction was found for participant reasoning by group favorability evaluation, $F(3, 846) = 13.44, p < .001, \eta^2 = .04$. This interaction revealed that participants who thought that the group would be favorable to the deviant member who wanted to play football used all four forms of reasoning, with no significant differences in amount of use. They cited concerns about group functioning ($M = .30, SD = .45$), described the deviant member's autonomy ($M = .19, SD = .39$), noted the benefits of including diverse perspectives ($M = .18, SD = .37$), and discussed the traditional gender stereotypes associated with playing football ($M = .24, SD = .41$). Those participants who thought that the group would not like the deviant member who wanted to play football referenced primarily group functioning ($M = .45, SD = .48$) and gender identity and stereotypes ($M = .38, SD = .46$). They referenced group functioning and gender identity and stereotypes significantly more than autonomy or the inclusion of diverse perspectives with $ps < .001$, and referenced group functioning more than gender identity and stereotypes with $p < .05$. There was no difference between the use of autonomy ($M = .05, SD = .20$) and the inclusion of diverse perspectives ($M = .02, SD = .16$). Thus, while those who thought the group would like the deviant member relied upon a range of different reasons, those who thought the group would not like the deviant member focused more on group issues, including group functioning and gender based expectations for group activities.

There was also an interaction for age group by group favorability evaluation by reasoning: $F(3, 846) = 3.76, p < .05, \eta^2 = .01$. This revealed that 9 year olds and 13 year

olds who thought the group would like the deviant member who wanted to play football reason differently about group functioning, $p < .05$, with 9 year olds using less group functioning reasoning ($M = .20$, $SD = .40$) than did 13 year olds ($M = .34$, $SD = .46$).

Ballet. The ANOVA conducted for the ballet condition also revealed differences between participants who thought that the group would versus would not like the deviant member who wanted to play ballet, $F(3, 843) = 17.63$, $p < .001$, $\eta^2 = .05$. Similar to the football condition, participants who said that the group would like the deviant who wanted to play ballet cited a range of difference reasons, with no statistical difference in the proportion using each form of reasoning. They referenced group functioning ($M = .36$, $SD = .47$), the positive effects of including diverse perspectives ($M = .20$, $SD = .38$), the importance of the deviant member's autonomy ($M = .21$, $SD = .39$), and gender identity and gender stereotypes ($M = .16$, $SD = .34$). Participants who thought that the group would not like the deviant member relied most about gender stereotypes ($M = .47$, $SD = .47$), with almost half of participants citing gender stereotypes, followed by concerns with group functioning ($M = .32$, $SD = .46$). Very few participants referenced the inclusion of diverse perspectives ($M = .06$, $SD = .22$), or autonomy ($M = .01$, $SD = .08$). Participants made more references to group functioning and gender stereotypes than diverse perspectives or autonomy, $ps < .001$. There were no differences in the use of gender stereotypes and group functioning or in the use of diverse perspectives or autonomy.

Thus, overall, the reasoning results for group favorability in both the ballet and football conditions reveal that participants believe that groups who do not like deviant

members are concerned with how the deviant will impact group functioning and if the deviant behavior aligns with gender expectations or not.

Individual Favorability Towards the Deviant Member

The results measuring group favorability towards the deviant member indicate that gender stereotypes play a strong role in how children and adolescents think groups will respond to deviant members. Our next question was whether children attend to the same issues and concerns when individually evaluating how much they would favor deviant members. We expected that they would, individually, be less influenced by stereotypes and would show more support for the deviant members. In order to test hypotheses concerning individual favorability of the deviant member, 2 separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) ANOVAs were conducted for individual favorability in the ballet and football deviance conditions. For the football condition, there were no significant effects for gender, age group or version. For the ballet condition, there was a significant gender interaction effect, $F(1, 279) = 6.16, p < .05, \eta^2 = .02$, which revealed that female participants were more favorable ($M = 5.16, SD = 1.13$) to the deviant who wanted to do ballet than were male participants ($M = 4.91, SD = 1.23$). Additionally, there was a version interaction effect, $F(1, 279) = 32.14, p < .001, \eta^2 = .10$, which revealed that participants were more favorable to the girl who wanted to do ballet ($M = 4.94, SD = 1.19$) than to the boy who wanted to do ballet ($M = 3.92, SD = 1.71$). Thus, these findings indicate that, similar to the group favorability evaluations, gender stereotypes do influence how children and adolescents evaluate peers who challenge their group's norms regarding gender

stereotypic activities and, further, reveal that challenging stereotypes is viewed as less acceptable for boys than for girls.

The next hypotheses concerned differences between individual favorability towards the deviant who wanted to do ballet and the deviant who wanted to play football. It was expected that participants would prefer the deviant who wanted to play football. In order to assess if children and adolescents differed in their evaluations of a deviant member who wanted to play football or do ballet, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Condition: ballet, football) ANOVA with repeated measures on the last factor was conducted. An effect for condition was found, $F(1,278) = 46.10, p < .001, \eta^2 = .14$, which confirmed expectations that participants preferred the deviant who wanted to play football ($M = 5.05, SD = 1.18$) to the deviant who wanted to do ballet ($M = 4.42, SD = 1.60$). There was, additionally, a version by condition interaction effect, $F(1,278) = 24.42, p < .001, \eta^2 = .08$, which showed that participants were more favorable to the girl who wanted to play football ($M = 5.00, SD = 1.30$) than to the boy who wanted to do ballet ($M = 3.92, SD = 1.71$). There was no difference between the boy who wanted to do football ($M = 5.10, SD = 1.10$) and the girl who wanted to do ballet ($M = 4.94, SD = 1.19$). Thus, the differences between the conditions were accounted for by a rigid adherence to stereotypic expectations suggesting that ballet is an activity only appropriate for girls.

Justifications for Individual Favorability Towards the Deviant Member

Analyses were conducted in order to test hypotheses involving differences in children's and adolescents' reasoning about individual favorability towards the deviant member by participants who liked versus did not like the deviant member. As was done

for group favorability, participant responses to individual favorability were divided into a dichotomous variable (okay, not okay) using a mid-point split of 3.5. The top four forms of reasoning used by participants to reason about individual favorability towards the deviant member were group functioning (“if she wants to play something different it will mess up the group”), autonomy (“I like that he decided to do his own thing”), inclusion of diverse perspectives (“she will show them that everyone can play football”), and personal identification with the target (“I really like to play football, just like he does”). Note that for individual favorability, gender stereotypes were not one of the top forms of reasoning used. Two separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Individual Favorability: okay, not okay) X 4 (Justification: group functioning, autonomy, diverse perspectives, personal identification) ANOVAs with repeated measures on the last factor were conducted, one for each condition (football and ballet).

Football. For the football condition, an effect was found for individual favorability evaluation by reasoning, $F(3, 846) = 3.56, p < .05, \eta^2 = .01$. This revealed that participants who like the deviant member who wanted to play football primarily referenced autonomy ($M = .41, SD = .47$) and their personal identification with the target ($M = .20, SD = .38$). They also referenced the importance of including diverse perspectives ($M = .16, SD = .34$) and the role of group functioning ($M = .10, SD = .29$). Pairwise comparisons revealed that they used more autonomy reasoning than any other form of reasoning, $ps < .001$, that they used more reasoning referencing the inclusion of diverse perspectives than they did group functioning, $p < .05$, and that they used more references to including diverse perspectives than they did to group functioning, $p < .01$.

Participants who did not like the deviant member who wanted to play football made the most references to their personal identification with the target (i.e., “Well I wouldn’t want to play football, so I don’t like that she was telling my group to play,” $M = .26$, $SD = .44$). They also referenced group functioning ($M = .15$, $SD = .36$), autonomy ($M = .11$, $SD = .30$), inclusion of diverse perspectives (i.e., “even though he would bring a new activity to the group, I still don’t like him”, $M = .08$, $SD = .25$). Pairwise comparison revealed that the only significant difference was found between inclusion of diverse perspectives and personal identification with the target, $p < .05$. Thus, participants focused on a range of difference concerns when indicating that they did not like the deviant member who wanted to play football, while they were more focused on autonomy and their personal identification with the target if they did like the deviant member.

Ballet. As expected, differences were found between participants who liked the deviant member who wanted to do ballet and participants who did not like this deviant member, $F(3, 840) = 12.43$, $p < .001$, $\eta^2 = .04$. Specifically, almost half of participants who liked the deviant member who wanted to do ballet focused on autonomy ($M = .48$, $SD = .49$), with a smaller proportion focusing on group functioning ($M = .12$, $SD = .32$), the inclusion of diverse perspectives ($M = .13$, $SD = .32$), or personal identification with the target ($M = .11$, $SD = .29$). The use of autonomy differed from the use of each of the other forms of reasoning at $p < .001$. For participants who did not like the deviant member reasoning was more centered on their personal identification with the target ($M = .22$, $SD = .41$), with almost no participants referencing autonomy ($M = .05$, $SD = .19$) and a small number referencing group functioning ($M = .11$, $SD = .31$) and the inclusion of diverse perspectives ($M = .10$, $SD = .30$). Participants made significantly more references

to their personal identification with the target than to autonomy at $p < .05$. Note that participants who did not like the deviant member who wanted to do ballet were divided in their reasoning, referencing a number of different reasons.

Relation between Group and Individual Favorability Towards Deviant Members in the Context of Gender Stereotypic Activities

Were there differences, then, between participants' evaluations of how groups would respond to these deviant members and how they would personally respond? It was expected that participants would rate group favorability towards the deviant members who wanted to play ballet and football less than individual favorability, as they would expect that groups would rely more upon stereotypes about gender appropriate activities than would participants, themselves. In order to test for these differences 2 separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Question: peer, individual resistance) ANOVAs with repeated measures on the last factor was conducted for the football condition and the ballet condition. The ANOVA conducted for the football condition revealed that participants were more favorable ($M = 5.05$, $SD = 1.18$) to the deviant member who wanted to play football than they expected groups to be ($M = 3.63$, $SD = 1.43$), $F(1,280) = 280.59$, $p < .001$, $\eta^2 = .42$. The ANOVA conducted for the ballet condition revealed, similarly, that participants were more favorable ($M = 4.42$, $SD = 1.6$) to the deviant member who wanted to do ballet than they expected groups to be ($M = 3.08$, $SD = 1.52$), $F(1,277) = 168.57$, $p < .001$, $\eta^2 = .37$.

Likelihood of Exclusion of the Deviant Member

The next set of hypotheses surrounded the likelihood that the deviant member would be excluded from the group because of the member's decision to resist the group

norm. Do children and adolescents believe that peers who deviate from group norms involving gender stereotypic activities will be kicked out of the group for challenging the group? It was expected that participants would perceive exclusion as most likely for the boy who wanted to do ballet. In order to test these hypotheses, 2 separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) ANOVAs were conducted for likelihood of exclusion in the ballet and football deviance conditions. For the football condition, a gender by version interaction was found, $F(1, 279) = 6.48, p < .05, \eta^2 = .02$, indicating that male participants did not differ in their expectations about the likelihood of exclusion of a girl ($M = 3.08, SD = 1.57$) or a boy ($M = 2.73, SD = 1.58$) who wanted to play football. Female participants, however, did differ ($p < .05$), and were more likely to expect that a boy ($M = 3.49, SD = 1.51$) would be excluded for playing football than would a girl ($M = 2.93, SD = 1.52$). Thus, female participants may believe that boys groups are more likely to exclude than are girls groups, showing a form of ingroup preference. For the ballet condition, an effect was found for version, $F(1, 275) = 19.58, p < .001, \eta^2 = .06$, revealing that, as expected, participants thought that it was more likely that a boy ($M = 3.98, SD = 1.63$) would be excluded for wanting to do ballet than a girl ($M = 3.10, SD = 1.59$).

It was expected, additionally, that participants would believe that the member who deviated by wanting to play football would be less likely to be excluded than the member who wanted to do ballet. In order to test this hypothesis, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Condition: ballet, football) ANOVA with repeated measures on the last factor was conducted. This confirmed the expectation, $F(1, 272) = 22.72, p < .001, \eta^2 = .07$, revealing that, while

participants were generally ambivalent about the likelihood of exclusion of either member, they thought that the deviant member who wanted to do ballet ($M = 3.55$, $SD = 1.67$) was more likely to be excluded than the deviant member who wanted to play football ($M = 3.07$, $SD = 1.56$). Further, a condition by gender interaction, $F(1,272) = 6.30$, $p < .05$, $\eta^2 = .02$, revealed that while female participants did not differ in their evaluations of these two deviant members, male participants were more likely to expect that the deviant member who wanted to do ballet ($M = 3.65$, $SD = 1.72$) would be excluded than the member who wanted to play football ($M = 2.92$, $SD = 1.57$), $p < .001$. Additionally, a condition by version effect, $F(1,272) = 17.22$, $p < .001$, $\eta^2 = .06$, revealed that participants did not differ in their evaluations of the girl and the boy who wanted to play football (girl: $M = 3.01$, $SD = 1.54$, boy: $M = 3.14$, $SD = 1.58$), but that they were much more likely to expect that the boy who wanted to do ballet ($M = 3.98$, $SD = 1.63$) would be excluded than the girl who wanted to do ballet ($M = 3.10$, $SD = 1.59$), $p < .001$, see Figure 2.

Justifications for Likelihood of Exclusion of the Deviant Member

Analyses were conducted in order to test hypotheses involving differences in children's and adolescents' reasoning about the likelihood of excluding the deviant member, using a dichotomous variable (yes, no) for participants who thought the deviant should be versus should not be excluded. This variable was computed using a mid-point split of 3.5. The top three forms of reasoning used by participants to reason about the likelihood of excluding the deviant member who wanted to play football were group functioning, gender stereotypes and inclusion of diverse perspectives. For evaluations of

the deviant member who wanted to ballet, the top three forms of reasoning were group functioning, gender stereotypes and autonomy.

Football. In order to assess differences in reasoning about the likelihood of exclusion of a deviant member who wants to play football, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Likelihood of Exclusion: yes, no) X 3 (Justification: group functioning, gender stereotypes, diverse perspectives) ANOVA with repeated measures on the last factor was conducted. This revealed that participants did differ in their proportional use of reasoning, $F(2,560) = 74.61, p < .001, \eta^2 = .21$. They used primarily group functioning ($M = .56, SD = .48$), with a smaller proportion referencing gender stereotypes ($M = .16, SD = .44$), and the inclusion of diverse perspectives ($M = .07, SD = .25$). All groups differed significantly at $ps < .001$. There was no difference between participants who thought the deviant member would versus would not be excluded.

Ballet. In order to assess differences in reasoning about the likelihood of exclusion of a deviant member who wants to do ballet, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Likelihood of Exclusion: yes, no) X 3 (Justification: group functioning, gender stereotypes, autonomy) ANOVA with repeated measures on the last factor was conducted. This revealed that participants who thought that the deviant member would be excluded used different forms of reasoning than those who thought the deviant member would not be excluded, $F(2,552) = 11.58, p < .001, \eta^2 = .04$. Specifically, there were no differences in the use of group functioning, all participants frequently referenced group functioning (Likely: $M = .35, SD = .47$, Not Likely: $M = .52, SD = .47$). However, participants who thought that

exclusion was likely ($M = .42$, $SD = .48$) used much more reasoning involving gender stereotypes than those who did not ($M = .15$, $SD = .34$), $p < .01$, and used much less reasoning about autonomy ($M = .03$, $SD = .17$) than did those who thought the deviant would not be excluded ($M = .12$, $SD = .32$), $p < .05$, see Figure 3.

Inclusion Decisions

Our last question regarding gender stereotypic activities involved who participants would choose to include in a group. If asked to choose between someone who matches the gender identity of the group or who wants to engage in the same activity of the group, what will children and adolescents decide? It was expected that participants would be more likely to include an outgroup member (by gender) who wants to do the same activity as the group into a group when the group plays football than when they do ballet. In order to test this hypothesis, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Condition: ballet, football) ANOVA with repeated measures on the last factor was conducted. This analysis confirmed expectations, revealing that participants were more likely to include an outgroup member who wanted to play football into a group that was playing football ($M = .80$, $SD = .40$), than an outgroup member who wanted to do ballet into a group that was doing ballet ($M = .60$, $SD = .49$), $F(1,265) = 31.52$, $p < .001$, $\eta^2 = .10$. It is important to note that participants were, in fact, quite willing to include a child of the opposite gender into a group that wanted to play football, indicating that gender divisions may not be in place for all typically gender stereotypic activities.

Gendered Forms of Aggression

Stereotype Measure

The results presented so far reveal that children and adolescents do attend to and use gender stereotypes about activities when evaluating group members and group decisions. The next set of questions involved not gender stereotypic activities, but rather gendered forms of aggression. The first question was: do children and adolescents actually hold stereotypes indicating that physical aggression is connected to boys and relational aggression to girls? Descriptive statistics indicate that children and adolescents do hold stereotypes linking boys to physical aggression with almost 75% of participants endorsing this stereotype, but that fewer participants hold a stereotype linking girls to relational aggression (gossip, in this context) (See Table 2). These descriptive statistics suggest that children and adolescents hold stronger stereotypes about gendered forms of aggression for boys than for girls.

Relation between Peer and Individual Resistance to Group Norms about Aggression

Our first question involving group norms about aggression was: do children think that their peers will challenge group norms about aggression and would they, individually, challenge these norms? In order to address this question, 4 separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Question: peer, individual resistance) ANOVAs with repeated measures on the last factor were conducted, one for each condition (nice, rough, impartial, and gossip). When the group wanted to play rough and a group member wanted to play nice, participants asserted that they thought that the group member would speak up to challenge the group ($M = .69$, $SD = .46$), but that they would be even more likely to do so ($M = .86$, $SD = .35$), $F(1,278) = 23.06$, $p < .001$, $\eta^2 = .07$, see Figure 4. There were no differences, however, when the group wanted to play nicely and a group member wanted to play

rough, *n.s.*: both peer ($M = .74$, $SD = .44$) and individual resistance ($M = .80$, $SD = .40$) were evaluated as likely, see Figure 4.

Similarly, when the group wanted to be impartial and a group member wanted to gossip, both peer ($M = .71$, $SD = .46$) and individual ($M = .72$, $SD = .45$) resistance were evaluated as likely, there were no differences, *n.s.* When the group wanted to gossip and a group member wanted to be impartial, participants asserted that they thought that the group member would speak up to challenge the group ($M = .69$, $SD = .46$), but that they would be even more likely to do so ($M = .88$, $SD = .33$), $F(1,278) = 34.08$, $p < .001$, $\eta^2 = .10$, see Figure 4. Overall, then, participants did believe that their peers would stand up to groups and challenge their norms. They also asserted that they would do the same. However, when the group was engaging in aggressive behavior, participants, individually, asserted that they would be even more likely to challenge the group than they expected their peers to be.

Group Favorability Towards the Deviant Group Member

How do children and adolescents think that groups will respond to peer group members who do challenge the group about group norms involving aggression? It was expected that participants would assert that groups would not like deviant members, but that there would be variation based on the deviant member's action, with lower ratings for deviants who advocate for the group to engage in aggression. In order to test these hypotheses, repeated measures ANOVAs were conducted comparing responses to group favorability in the different conditions.

First, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Condition: rough, nice) ANOVA with repeated measures on the last

factor was conducted for group favorability. This revealed that participants thought groups would be more favorable toward a deviant member who wanted to play nicely ($M = 3.44$, $SD = 1.57$) than a deviant member who wanted to play rough ($M = 2.95$, $SD = 1.62$), even though both of these actions went against the group norm, $F(1,278) = 18.23$, $p < .001$, $\eta^2 = .06$.

Similarly, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Condition: impartial, gossip) ANOVA with repeated measures on the last factor was conducted for group favorability. This revealed that participants thought groups would be more favorable toward a deviant member who wanted to be impartial ($M = 3.27$, $SD = 1.58$) than a deviant member who wanted to gossip ($M = 2.55$, $SD = 1.53$), $F(1,281) = 35.88$, $p < .001$, $\eta^2 = .11$. It is important to note, however, that generally, participants thought that groups would not like any of these deviant members: none of the means cross above the mid-point of 3.5.

Two separate ANOVAs were also conducted to assess differences in group favorability for the different types of aggressive and non-aggressive behaviors. Did participants view one form of aggression as more acceptable to the group than another? First, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Condition: rough, gossip) ANOVA with repeated measures on the last factor was conducted for group favorability. This revealed that participants expected that groups would be more favorable to a deviant who wanted to play rough than a deviant who wanted to gossip, $F(1, 278) = 9.589$, $p < .01$, $\eta^2 = .03$. The 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Condition: nice, impartial) ANOVA with repeated measures on the last factor which was

conducted revealed no differences between the nice and impartial conditions. Thus, participants may have recognized the benefit of playing rough to the group when considering playing rough versus gossiping, but not seen similar distinctions between playing nice and being impartial.

Justifications for Group Favorability Towards the Deviant Member

Children do evaluate group favorability towards different types of deviants distinctly, but do children and adolescents expect that groups will use different reasons to justify how favorable they are towards these deviant members? Analyses were conducted in order to test hypotheses concerning differences in children's and adolescents' reasoning about the group's favorability towards the deviant member by participants who thought the group would like versus would not like the deviant member. As was done for the social activities, participant responses to group favorability towards the deviant member were divided into a dichotomous variable (okay, not okay) using a mid-point split of 3.5. The top four forms of reasoning used by participants to reason about group favorability towards the deviant member were group functioning, autonomy, welfare ("someone could get hurt if they play rough") and gender identity or stereotypes ("girls don't play rough, so they won't like her"). Four separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Group Favorability: okay, not okay) X 2 (Justification: group functioning, autonomy, welfare, gender) ANOVA with repeated measures on the last factor were conducted, one for each condition (rough, nice, gossip and impartial).

The ANOVA for the rough condition, when the group wanted to play nicely and the deviant wanted to play rough revealed a main effect for reasoning, $F(3, 840) = 24.71$,

$p < .001$, $\eta^2 = .08$, indicating that participants focused on both group functioning and welfare in their reasoning, relying less on references to autonomy and gender stereotypes, see Table 3. Pairwise comparisons showed that use of both group functioning and welfare differed significantly from use of both autonomy and gender stereotypes at $ps < .001$. Participants were centered on the fact that the deviant was both going against the group and doing something that could hurt others.

Did reasoning also differ for the deviant who wanted to play nicely when the group wanted to play rough? The ANOVA conducted for the nice condition showed a main effect for reasoning, $F(3, 417) = 38.65$, $p < .001$, $\eta^2 = .21$, and an interaction effect for reasoning split by group favorability judgment, $F(3, 417) = 6.51$, $p < .001$, $\eta^2 = .04$. The main effect showed that participants relied primarily upon group functioning reasoning, $p < .001$. The interaction effect revealed that when they thought that the group would like the deviant member, they used less group functioning and more autonomy than when they thought the group would not like the deviant member who wanted to play nicely, $ps < .001$, see Table 3. Thus, participants believe that, at times, groups may value individuality and autonomous thinking in their members.

When the deviant wanted to gossip when the group wanted to be impartial, there was also a main effect for reasoning, $F(3, 846) = 11.60$, $p < .001$, $\eta^2 = .03$, and an interaction effect for reasoning by group favorability evaluation, $F(3, 846) = 11.98$, $p < .001$, $\eta^2 = .04$. The main effect revealed that they relied primarily upon group functioning and welfare reasoning, with less use of autonomy and gender. Pairwise comparisons showed that use of both group functioning and welfare differed significantly from use of both autonomy and gender stereotypes at $ps < .001$. The interaction effect revealed that

participants used more references to group functioning and welfare when they thought the group would not like the deviant and more references to autonomy and gender stereotypes when they thought the group would like the deviant, see Table 3. Use of autonomy and welfare differed significantly at $ps < .001$ between those participants who thought the group would like versus would not like the deviant who wanted to gossip. Thus, participants recognize the psychological harm that gossiping can cause, while also understanding that the group would want people who also share their group norms.

Finally, in the condition where the group wanted to gossip and the deviant wanted to be impartial, a main effect for reasoning, $F(3, 849) = 38.50, p < .001, \eta^2 = .12$, and an interaction effect for reasoning by group favorability evaluation, $F(3, 849) = 17.29, p < .001, \eta^2 = .05$, were found. When the deviant wanted to be impartial, most participants relied upon group functioning reasoning, $ps < .001$, but made frequent references to welfare. When participants thought that the group would like the deviant member, they use all four forms of reasoning, but when they thought the group would not like the deviant member, they used mostly group functioning reasoning, see Table 3. They used more group functioning reasoning when they thought the group would not like the deviant, $p < .001$, and more welfare and autonomy when they thought the group would like the deviant, $ps < .01$. Thus, participants' reasoning reveals that they balance information about the group's goals, the welfare of others and gender stereotypes when making decisions about group favorability.

Individual Favorability Towards the Deviant Member

How do participants balance information about group loyalty, group norms, gender stereotypes and their sense of fairness when individually evaluating group

members who deviate from group norms about aggressive behavior? It was expected that participants themselves, would be most influenced by other's welfare, showing strong support for deviants who advocate for avoiding aggression when their groups are aggressive. In order to test these hypotheses, 4 separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) ANOVAs were conducted for individual favorability in each of the conditions.

When the deviant wanted to play rough when the group wanted to play nicely, an age effect was found, $F(1, 278) = 15.48, p < .001, \eta^2 = .05$, revealing that adolescents showed more support for the rough deviant ($M = 3.61, SD = 1.83$) than did younger children ($M = 2.67, SD = 1.77$). Similarly, when the deviant wanted to play nicely when the group wanted to play rough, an age effect was found, $F(1, 280) = 5.27, p < .05, \eta^2 = .01$. Here, a complimentary pattern was found: younger children showed more support for the nice deviant ($M = 4.92, SD = 1.60$) than did adolescents ($M = 4.43, SD = 1.64$). Additionally, a gender effect was found for the nice deviant, $F(1, 280) = 6.46, p < .05, \eta^2 = .02$, revealing that female participants ($M = 4.91, SD = 1.48$) were more favorable to the nice deviant than were male participants ($M = 4.22, SD = 1.74$). This may suggest an implicit adherence to gender stereotypes, with male participants showing less support for a deviant who advocates for an action that is counter-stereotypic for boys.

In the impartial condition, there was also an age effect, $F(1, 282) = 8.59, p < .05, \eta^2 = .03$, revealing that younger children are more supportive of the deviant who wants to be impartial ($M = 5.33, SD = 1.14$) than are adolescents ($M = 4.83, SD = 1.52$). There were no effects for the gossip condition. All participants agree that the deviant who gossips will not be liked ($M = 2.62, SD = 1.72$). Younger children, however, were more

supportive, than were adolescents, of both deviant members who challenged the aggressive behavior of the group. This may indicate the increasing social pressure to adhere to group norms in adolescence or a greater recognition of the benefits to the group of adhering to these norms.

Next, in order to assess differences across conditions, a series of 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Condition) ANOVAs with repeated measures on the last factor was conducted for each pair of conditions (for instance rough, nice). The ANOVA conducted for the rough and nice conditions revealed an effect for condition, $F(1,276) = 81.46, p < .001, \eta^2 = .22$, indicating that, as expected, participants were more positive towards the nice than the rough deviant. Further, there was an age by condition interaction, $F(1,276) = 16.97, p < .001, \eta^2 = .05$, that revealed that younger participants were more supportive of the nice deviant and less supportive of the rough deviant and the adolescents showed the reverse pattern. The ANOVA conducted for the rough versus the gossip condition revealed an effect for condition, $F(1,275) = 20.79, p < .001, \eta^2 = .07$. This showed that participants were more positive towards the deviant who wanted to play rough than the deviant who wanted to gossip. The ANOVA conducted on the gossip and the impartial condition revealed, as expected, that participants were more positive towards the impartial deviant than the gossip deviant, $F(1,279) = 302.77, p < .001, \eta^2 = .52$. Finally, the impartial versus nice ANOVA revealed an effect for condition, $F(1,280) = 15.91, p < .001, \eta^2 = .05$, showing that participants preferred the impartial deviant over the nice deviant. Thus, this set of analyses confirmed that participants are weighing these different conditions carefully and making reasoned evaluations of each condition. The justification data will

provide further insight into exactly what factors they weigh when making these evaluations, but it is interesting to note that all of the version effects were non-significant, which indicates that gender stereotypes about aggressive behavior were not at play in children's and adolescents' evaluations.

Justifications for Individual Favorability Towards the Deviant Member

What reasons do children and adolescents use to justify how much they like these deviant members? Analyses were conducted in order to test hypotheses concerning differences in children's and adolescents' reasoning about individual favorability towards the deviant member by participants who liked versus did not like the deviant member. Participant responses to individual favorability towards the deviant member were divided into a dichotomous variable (not okay, okay) using a mid-point split of 3.5. The top four forms of reasoning used by participants to reason about individual favorability towards the deviant member were group functioning, autonomy, welfare, and personal identification with the target (e.g., "well, I don't like to gossip either"). Note that, for group favorability, gender stereotypes were in the top four forms of reasoning, but for individual favorability they were much less frequently used. This indicates that, while participants believe groups may be motivated by stereotypes in their decision-making, they are not as influenced, personally, by such stereotypes. Four separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Individual Favorability: okay, not okay) X 2 (Justification: group functioning, autonomy, welfare, personal identification) ANOVA with repeated measures on the last factor were conducted, one for each condition (rough, nice, gossip and impartial).

The ANOVA conducted for the rough condition revealed a main effect for reasoning, $F(3, 837) = 26.01, p < .001, \eta^2 = .08$, indicating that participants referenced welfare most frequently, with less frequent references to group functioning, autonomy and personal identification with the target, $ps < .001$, see Table 4.. Additionally, there was a reasoning by age interaction, $F(3, 837) = 8.01, p < .001, \eta^2 = .02$. This revealed that younger participants made many more references to welfare ($M = .63, SD = .47$) than did adolescents ($M = .28, SD = .44$), $p < .001$. Further, there was an interaction for type of reasoning by individual favorability evaluation, $F(3, 837) = 37.73, p < .001, \eta^2 = .11$. Participants used more group functioning reasons when they were favorable to the rough deviant than when they were not favorable, $p < .05$. They also used more autonomy reasoning when they liked the rough deviant than when they did not like the rough deviant, $p < .001$. Finally, they used more welfare reasoning when they did not like the rough deviant than when they liked the rough deviant, $p < .001$, see Table 4.

For the nice deviant, a main effect for reasoning was found, $F(3, 843) = 10.32, p < .001, \eta^2 = .03$. Participants used more welfare reasoning than group functioning, $p < .05$, autonomy or welfare, $ps < .001$. Additionally, an age by reasoning interaction was found, $F(3, 843) = 4.71, p < .01, \eta^2 = .01$, which revealed that children used more welfare, ($M = .55, SD = .50$), $p < .001$, and less personal identification, ($M = .08, SD = .27$) $p < .05$, reasoning than did adolescents ($M_{welfare} = .28, SD = .44, M_{personal} = .17, SD = .37$). Finally, there was an interaction for reasoning by individual favorability evaluation, $F(3, 843) = 5.86, p = .001, \eta^2 = .02$. This revealed that participants used more group functioning and autonomy reasoning when they did not like the nice deviant, $ps < .01$, and more welfare reasoning when they do like the nice deviant, $p < .01$ see Table 4.

A main effect for the gossip condition was found, $F(3, 840) = 26.78, p < .001, \eta^2 = .08$. This revealed that participants used more welfare reasoning than any other form of reasoning, $p < .01$. Additionally, they used more autonomy than group functioning reasoning, $p < .05$, see Table 4. There was also an age group by reasoning interaction, $F(3, 840) = 4.83, p < .01, \eta^2 = .01$. Children used more references to welfare ($M = .66, SD = .46$) than did adolescents ($M = .43, SD = .48$), $p < .05$, and adolescents used more references to personal identification ($M = .21, SD = .39$) than did children ($M = .06, SD = .19$), $p < .05$. Finally, there was a reasoning by individual favorability interaction, $F(3, 840) = 30.86, p < .001, \eta^2 = .09$. This revealed that participants used more references to autonomy when they liked the deviant who gossips and more references to welfare when they did not like this deviant, $ps < .001$ see Table 4. Finally, for the impartial condition, there was an effect for reasoning, $F(3, 849) = 10.42, p < .001, \eta^2 = .03$. This revealed that participants used more references to welfare than group functioning, or autonomy, $ps < .001$, and personal identification with the target, $p < .05$ see Table 4. Overall, participants' individual evaluations of the deviant members were heavily influenced by their sense of welfare. Interestingly, there was also a consistent age-related pattern with children being more focused on welfare, while adolescents showed greater concern with their own personal identification with the target's behavior.

Group Favorability Versus Individual Favorability

Did participants evaluate these deviant members differently than how they expected groups would evaluate the deviant members? It was expected that participants would expect that groups would be very loyal to their norms, while participants themselves would attend more to the moral valence of the deviant member's behavior.

Participants would show more support for the deviants who wanted to challenge aggression than they would expect groups to show. In order to test these hypotheses, 4 separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Question: group, individual favorability) ANOVAs with repeated measures on the last factor were conducted, one for each condition (nice, rough, impartial, and gossip).

The ANOVA conducted for the rough condition revealed a significant difference between the individual and the group favorability, $F(1,275) = 5.74, p < .05, \eta^2 = .02$. Participants were slightly more favorable to the deviant member than they expected groups would be, see Figure 5. The ANOVA conducted for the nice deviant revealed, as expected that participants were much more favorable toward the nice deviant member than they expected groups to be, $F(1,280) = 97.36, p < .001, \eta^2 = .25$, see Figure 5.

The ANOVA conducted for the gossip condition was non-significant: individuals did not like a deviant member who wanted to gossip and thought that groups would also not like this deviant. For the impartial condition, there was an effect for question, $F(1,282) = 255.72, p < .001, \eta^2 = .47$, showing that participants were much more positive towards an impartial deviant than they expected the group to be, see Figure 5. Thus, as expected, participants were more favorable toward deviant members who resist aggression than they expected the group to be.

Likelihood of Exclusion

Do children and adolescents expect that someone who challenges the group's norms about aggressive behavior will be excluded from the group? It was expected that

exclusion would be viewed as a consequence for challenging the group, given children's frequent experiences with social exclusion.

As a first test of this hypothesis, 4 one sample t-tests were conducted, one for each condition, against a neutral test value of 3.5. Results indicated that participants did generally expect that exclusion would be likely for the deviants who wanted to act in an aggressive manner (gossip: $t(287) = 4.98, p < .001, d = .59, M = 3.99, SD = 1.71$, and rough: $t(284) = 3.09, p < .01, d = .58, M = 3.81, SD = 1.72$). For the impartial ($M = 3.40, SD = 1.69$) and nice ($M = 3.68, SD = 1.69$) conditions, the t-tests were non-significant, indicating that participant responses did not differ from the neutral mid-point of 3.5. Thus, participants thought that someone who challenged the group by advocating for gossiping or playing rough would be excluded from the group, but were unsure if someone who challenged the group to be impartial or to be nice would be excluded, see Figure 6.

In order to test for differences by age group, gender and version, 4 separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) ANOVAs were conducted for likelihood of exclusion in each of the conditions. The ANOVAs revealed age effects for the rough condition ($F(1, 275) = 6.85, p < .01, \eta^2 = .02$). Here children ($M = 4.2, SD = 1.79$) were more likely to expect exclusion of the rough deviant than were adolescents ($M = 3.63, SD = 1.67$). The nice condition showed a similar age effect, $F(1, 275) = 8.68, p < .01, \eta^2 = .03$, children: $M = 4.13, SD = 1.62$, adolescents: $M = 3.47, SD = 1.68$. Finally, in the gossip condition ($F(1, 278) = 14.96, p < .001, \eta^2 = .05$), an age effect also revealed that children ($M = 4.55, SD = 1.68$) were more likely to expect exclusion than were adolescents ($M = 3.74, SD = 1.67$). This

revealed that children were more likely to expect exclusion than were adolescents, in all conditions except the impartial condition, see Figure 7.

Overall, these results indicate that participants, especially children, do fear exclusion as a consequence for challenging the group, but that they are also attuned to the moral valence of the behavior a deviant advocates. They expect that one is less likely to be excluded if one is advocating that the group avoid aggression than if they encourage the group to engage in aggression.

Justifications for Likelihood of Exclusion

How do children reason about their evaluations of the likelihood of exclusion? Do they always expect that groups will focus on group functioning? Are there differences in reasoning between those participants who expect exclusion to occur and those who do not? In order to address these questions, analyses were conducted using a dichotomous variable created with a 3.5 median split on participants' likelihood of exclusion. For each condition, analyses were conducted on the top 4 forms of reasoning: group functioning, welfare, gender and the inclusion of diverse perspectives. For each condition, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Likelihood of Exclusion: likely, not likely) X 2 (Justification: group functioning, welfare, gender, diverse perspectives) ANOVA with repeated measures on the last factor was conducted

The ANOVA conducted for the rough deviant revealed an effect for justification, $F(3, 828) = 68.92, p < .001, \eta^2 = .20$. This revealed that participants used more group functioning reasoning than any other form, $ps < .001$, and used more references to protecting others' welfare than references to diverse perspectives or gender stereotypes,

$ps < .001$, see Table 5. Additionally, there was an interaction for reasoning by likelihood of exclusion, $F(3,828) = 9.06, p < .001, \eta^2 = .03$. This showed that participants referenced others' welfare when they expected that the deviant member who wanted to play rough would be excluded more than when they thought the deviant would not be excluded, $p < .001$. For the condition where the deviant member wanted to play nicely, an effect was found for reasoning, $F(3, 825) = 114.45, p < .001, \eta^2 = .29$, which revealed that participants primarily made references to group functioning, $ps < .001$. Thus, in both the rough and the nice condition, participants focused on group functioning, but also attended to other's welfare, especially when the deviant advocated playing rough.

For the gossip condition, there was an effect for reasoning, $F(3,837) = 88.13, p < .001, \eta^2 = .24$, which revealed, similarly to the rough condition, that participants used more group functioning reasoning than any other form of reasoning, $ps < .001$ and used more welfare reasoning than reasoning about gender stereotypes or diverse perspectives, $ps < .001$, see Table 5. Further, there was an interaction between type of reasoning and likelihood of exclusion evaluation, $F(3,837) = 14.56, p < .001, \eta^2 = .05$. This revealed that participants used more references to group functioning when they thought the deviant would not be excluded than when they thought the deviant would be excluded, $p < .05$. Participants made more references to others' welfare when they thought that the deviant who wants to gossip would be excluded than when they thought this deviant would not be excluded, $p < .001$. Finally, they made more references to the benefits of including diverse perspectives when they thought the deviant would not be excluded than when they thought that the deviant who wants to gossip would be excluded, $p < .05$.

For the impartial condition, there was also an effect for reasoning, $F(3,840) = 74.86, p < .001, \eta^2 = .21$. This showed that participants used more references to group functioning than any other category, $ps < .001$. There was also an interaction between reasoning and likelihood of exclusion, $F(3,840) = 19.21, p < .001, \eta^2 = .06$. This revealed that group functioning reasoning was used more by participants who thought that the impartial deviant would be excluded than those who thought this deviant would not be excluded, $p < .001$. References to the inclusion of diverse perspectives were more common among participants who thought the impartial deviant would not be excluded than among those who thought that the impartial deviant would be excluded, $p < .001$. Finally, participants who thought the impartial deviant would not be excluded referenced others' welfare more than those who thought the impartial deviant would be excluded, $p < .01$. Thus, children and adolescents weigh different concerns when thinking about if someone will be excluded for challenging their group. Group functioning was always a primary concern. Further, those who expected that deviants who advocated for aggression would be excluded focused more on others' welfare.

Likelihood of Exclusion by Group Favorability Evaluation

If you expect that the group will not like the deviant, are you more likely to expect that the deviant will be excluded? In order to address this question, a dichotomous variable was created using a mid-point split of 3.5 on group favorability to establish a variable which captured participants who thought that a group would not like versus would like a deviant member. Then, 4 separate a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Group Favorability: like, not like) ANOVAs were conducted on likelihood of exclusion, one for each condition. The

ANOVA conducted for the rough ($F(1,266) = 35.46, p < .001, \eta^2 = .11$), condition revealed an effect for group favorability, showing that participants who thought that groups would not like a deviant were more likely to expect that the deviant would be excluded ($M = 4.26, SD = 1.62$) than those who thought that groups would like a deviant ($M = 2.95, SD = 1.59$), see Figure 8. Similarly, for the nice condition ($F(1, 267) = 38.13, p < .001, \eta^2 = .12$), those who thought the group would like the deviant were less likely to expect exclusion ($M = 3.01, SD = 1.60$) than those who thought the group would not like the deviant ($M = 4.27, SD = 1.54$). Finally, the same pattern was shown in the impartial condition ($F(1,272) = 41.14, p < .001, \eta^2 = .13$): like: $M = 3.98, SD = 1.64$, not like: $M = 2.62, SD = 1.44$. While the main effect for group favorability for the gossip condition was not significant (like: $M = 4.22, SD = 1.64$, not like: $M = 3.22, SD = 1.72$), there was a significant interaction between group favorability evaluation and age group, $F(1,270) = 4.04, p < .05, \eta^2 = .01$. This revealed that 9 year olds thought that a deviant member who wanted to gossip would be excluded regardless of the whether the group did ($M = 4.55, SD = 1.67$) or did not ($M = 4.54, SD = 1.81$) like the deviant member, while 13 year olds recognized that the group's favorability toward the deviant member would have a greater impact on their evaluation ($M_{like} = 2.91, SD = 1.56, M_{not\ like} = 4.05, SD = 1.61$). Thus, participants used their understanding of groups and group desires when evaluating the likelihood of exclusion, and this skill may become more pronounced with age in some contexts.

Inclusion Choice

Children and adolescents are concerned about the group excluding someone who challenges the group's norm. However, what do children decide if asked to make a

choice between including someone who aligns with the group's norm but who is from a gender outgroup or someone who challenges the group norm, but is from the ingroup? For instance, would they prefer to include a boy in a girls group who would play nicely like the girls group, or would they rather include a girl who would play rough? Is maintaining the group norm central enough that children and adolescents will support including an outgroup member in the group? Further, how does this change depending on the group norm?

In order to address these questions, one sample t-tests were conducted against a test value of .50 for each condition on the proportion of participants who chose to include an outgroup member (by gender) who shared the group norm. This was done in order to assess whether participants were willing to include a gender outgroup member in order to preserve the group norm. Results indicated that for the rough, nice and impartial conditions, participants were likely to chose an outgroup member significantly more than chance (rough: $t(277) = 3.24, p = .001, d = .19, M = .59, SD = .49$, nice: $t(277) = 11.65, p < .001, d = .70, M = .79, SD = .41$, and impartial: $t(282) = 19.53, p < .001, d = 1.16, M = .88, SD = .32$, see Figure 9). For the gossip condition, participants did not vary from chance in their responses: $M = .51, SD = .50$. Thus, participants generally showed a willingness to include a gender outgroup member into the group in order to preserve the group norm. Interestingly, this is even the case when the group plays rough, perhaps reflecting a recognition of the benefits to the group of playing rough in a sports context.

In order to assess difference by age, age group and version, 4 separate 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) ANOVAs were conducted for inclusion choice in each of the conditions. There were no differences

for the rough condition or the gossip condition. For the impartial condition, a gender effect was found, $F(1,272) = 6.69, p < .01, \eta^2 = .02$. This revealed that female participants ($M = .93, SD = .25$) were more likely to include the outgroup member who wanted to be impartial into the impartial group than were male participants ($M = .82, SD = .38$). This is in line with previous research which suggests that female participants are often more inclusive than are male participants. For the nice condition, there was a version effect, $F(1, 268) = 7.07, p < .01, \eta^2 = .02$, which revealed that participants were more likely to include a girl into a boys group that wanted to play nicely ($M = .85, SD = .36$) than a boy into a girls group that wanted to play nicely ($M = .73, SD = .45$). This may reflect status differences, suggesting that participants see it as less acceptable for a boy to enter a girls group than a girl to enter a boys group, or an attention to the stereotype that girls play nicely.

The next ANOVA conducted for inclusion choice was a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Condition: rough, nice) ANOVA with repeated measures on the last factor. This revealed a significant effect for condition, $F(1, 261) = 23.37, p < .001, \eta^2 = .08$. This revealed, as expected, that participants were more willing to include an outgroup member who was nice into a group than an outgroup member who was rough. Finally, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Condition: gossip, impartial) ANOVA with repeated measures on the last factor was conducted for inclusion choice. A significant condition effect was found, $F(1,269) = 95.69, p < .001, \eta^2 = .26$, which revealed that participants were more willing to include an impartial outgroup member into a group than an outgroup member who wanted to gossip. Overall, these

results reveal, that, as expected, participants were more likely to include non-aggressive outgroup members into non-aggressive groups than aggressive outgroup members into aggressive groups, but that they are balancing information about group membership and group norm in making these decisions.

Justifications for Inclusion Choice

In order to better understand the reasoning behind participants' inclusion choices, participants' justifications for inclusion choice were examined. In order to address these questions, analyses were conducted using dichotomous variable created with a 3.5 median split on participants' likelihood of exclusion. For the each condition, analyses were conducted on the top 4 forms of reasoning: group functioning, welfare, gender and the inclusion of diverse perspectives. For each condition, a 2 (Gender: male, female) X 2 (Age Group: 9, 13 year olds) X 2 (Version: conform, resist) X 2 (Inclusion Choice: ingroup, outgroup) X 2 (Justification: group functioning, welfare, gender, diverse perspectives) ANOVA with repeated measures on the last factor was conducted.

The ANOVA for the rough condition revealed a main effect for reasoning, $F(3, 804) = 68.12, p < .001, \eta^2 = .20$, see Table 6. This revealed that participants were more likely to use group functioning than any other form of reasoning, $ps < .001$, and that participants also referenced welfare more often than the inclusion of diverse perspectives or gender stereotypes, $ps < .001$. There was also an inclusion choice by reasoning interaction, $F(3, 804) = 14.55, p < .001, \eta^2 = .05$. This revealed that participants who chose the ingroup member made more references to welfare and to the inclusion of diverse perspectives than those who did not, $ps < .01$. Additionally, those who chose the

outgroup member who wanted to play rough focused more on group functioning than those who chose the ingroup member, $p < .001$.

The ANOVA for the nice condition revealed a main effect for reasoning, $F(3, 804) = 36.00, p < .001, \eta^2 = .11$. This indicated that participants used more group functioning reasoning than any other category, $ps < .001$. The ANOVA for the gossip condition revealed a main effect for reasoning, $F(3, 828) = 70.39, p < .001, \eta^2 = .20$, which revealed that participants used more references to group functioning than any other form of reasoning, $ps < .001$, and more references to welfare than to stereotypes or the inclusion of diverse perspectives, $ps < .001$ see Table 6. Additionally, there was a inclusion choice by reasoning interaction, $F(3, 828) = 19.74, p < .001, \eta^2 = .06$. This revealed that participants who chose the ingroup member used more references to gender stereotypes ($p < .05$) and welfare ($p < .001$) and fewer references to group functioning ($p < .001$) than those who chose the outgroup member.

Finally, the ANOVA for the impartial condition revealed a main effect for reasoning, $F(3, 819) = 7.89, p < .001, \eta^2 = .02$. This showed that participants used more group functioning reasoning than any other form of reasoning, $ps < .01$. Overall, the reasoning results showed that participants frequently relied upon group functioning when deciding who to include in the group. However, for the rough and the gossip condition, when participants chose to include the ingroup member who advocated for non-aggressive behavior, they relied more upon welfare. Thus, participants were attuned to both the group norm and the moral valence of that norm when making their decision. Gender of the potential group members did not play as consequential a role in their evaluations.

Interpretative Theory of Mind Descriptive Statistics

For the interpretative theory of mind measure, scores were calculated on a 6 point scale from 0 = no theory of mind to 6 = full theory of mind. All participants performed very well on the interpretive theory of mind measure ($M = 4.87$, $SD = 1.37$). Further, performance was equally strong in both age groups (children: $M = 5.00$, $SD = 1.35$, adolescents: $M = 4.81$, $SD = 1.38$).

Interpretative Theory of Mind and Individual and Group Favorability

The results, thus far, indicate that children and adolescents reason very carefully about their evaluations of group members who want to challenge the group. In particular, they recognize that group perspectives may not be the same as their own individual perspectives. What is not known, however, is if group or individual perspectives are related to social-cognitive skills? Do children with more advanced abilities to recognize that two people may come to different conclusions about the same information (interpretative theory of mind) show an increased ability to take the perspective of either the group or the deviant member? In order to address these questions, multiple regression analyses were conducted for each condition using the interpretative theory of mind scale (from 0 = no theory of mind to 6 = full theory of mind), age and the interaction between age and interpretative theory of mind as predictor. Both age and theory of mind were centered prior to analyses and the interaction term was computed by multiplying age by theory of mind. First, age and theory of mind were entered into the model, and then age, theory of mind and the interaction term were entered into the model. For all analyses, the age by theory of mind interaction term was non-significant, so results presented will be for Model 1, with age and theory of mind included. For group

favorability, the regressions for each of the conditions except the impartial condition were non-significant. For the impartial condition, the multiple regression with age and theory of mind as predictors was statistically significant ($R^2 = 0.03$, $F(2,276) = 4.994$, $p < 0.01$), with theory of mind as the only significant predictor. Interpretative theory of mind accounted for 3.5% of the variance in group favorability toward the deviant member who wanted the group to be impartial. The unstandardized regression coefficient (β) for interpretative theory of mind was 0.20, meaning that for each additional unit increase in interpretative theory of mind skill, group favorability towards the impartial deviant increased 0.20 units. Thus, interpretive theory of mind played a small role in explaining group favorability evaluations.

For individual favorability, the regressions conducted for the deviants who advocated for aggressive behavior, in the rough and the gossip conditions, were non-significant. However, the regressions conducted for the impartial and the nice deviant members were significant. For the nice deviant, the multiple regression with age and theory of mind was statistically significant ($R^2 = 0.04$, $F(2,274) = 5.75$, $p < 0.01$), with both variables accounting for variation. The model accounted for 4% of the variance in individual favorability toward the deviant member who wanted the group to be nice. The unstandardized regression coefficient (β) for interpretative theory of mind was 0.15, meaning that for each additional unit increase in interpretative theory of mind skill, individual favorability towards the nice deviant increased 0.15 units. The unstandardized regression coefficient (β) for age was -0.08, meaning that for each additional unit increase in age, individual favorability towards the nice deviant decreased 0.08 units. For the impartial deviant, the multiple regression was statistically significant ($R^2 = 0.060$, F

(2,276) = 8.883, $p < 0.001$), with both age and theory of mind as significant predictors. The model accounted for 6% of the variance in individual favorability toward the deviant member who wanted the group to be impartial. The unstandardized regression coefficient (β) for interpretative theory of mind was 0.20, meaning that for each additional unit increase in interpretative theory of mind skill, individual favorability towards the impartial deviant increased 0.20 units. The unstandardized regression coefficient (β) for age was -0.07, meaning that for each additional unit increase in interpretative theory of mind skill, individual favorability towards the impartial deviant decreased 0.07 units. Thus, interpretative theory of mind competence accounted for a small amount of the variance for individual evaluations of the non-aggressive deviants.

Discussion

Gender Stereotypic Activities

The novel findings from this study indicate that children and adolescents rely upon stereotypes when evaluating peers who challenge their group norms regarding gender stereotypic social activities. Specifically, results revealed, across many measures, children's and adolescents' gender stereotypic expectations about girl-typed activities such as ballet, influenced their judgments about when and whether children should challenge peer group norms. Participants expected that a boy who wanted to challenge his group to try ballet was least likely to resist the group and they asserted that they would be less likely to resist the group if they were the boy who wanted to do ballet. These findings are explained, in part, by their assertion that groups would really dislike a gender non-conforming boy and that they, too, individually, would like a gender non-conforming boy the least. These results confirm the presence of a shifting standard

(Biernat & Manis, 1994), whereby challenging gender stereotypes by moving down the status hierarchy (boys acting in stereotypically female ways) is less acceptable than moving up the status hierarchy (girls acting in stereotypically male ways).

While girls encounter more societal barriers regarding exclusion (such as from participating in sports and math), boys encounter more psychological obstacles in terms of choice of activities. There are fewer barriers for boys, who are not literally excluded from girl-typed activities such as ballet, but the impacts on social favorability from peers is more pronounced than when girls desire to engage in boy-typed activities. For instance prior research with children has found that boys who engage in gender non-conforming behavior are judged more harshly than are girls who engage in gender non-conforming behavior (Smetana, 1986; Zucker, Wilson-Smith, Kurita, & Stern, 1995). Additionally, research with adolescents which examined perceptions of straight and gay peers who were gender conforming or gender non-conforming also confirms that boys who are gender non-conforming are judged the most harshly by their peers, regardless of their sexual orientation (Horn, 2007; Horn, 2008). Interestingly, research also indicates that, especially for boys, gender non-conformity in terms of activities is often judged less harshly than gender non-conformity in terms of appearance (Blakemore, 2003; Smetana, 1986). The current study assessed activities with strong visual associations (ballet and football). It is possible that participants were considering both the activity itself and the appearance of someone engaging in these gender non-conforming activities. In the current study, participants who did not condone challenging the group's gender stereotypes used stereotypes when reasoning about decisions to act in gender non-conforming ways. In particular, participants not only mentioned stereotypes citing who

should engage in each activity, but, also made assumptions that engaging in gender non-conforming activities also suggested sexual identity and orientation. Thus, participants were attuned to differences in societal expectations and referenced assumptions underlying stereotypes.

It is important to note, however, that while participants themselves were also least favorable toward the boy who wanted to do ballet, significant differences were found between individual and group favorability in both the football and ballet conditions. In other words, participants believed that their peers would be more negatively influenced by counter-stereotypic behavior and by challenging the group than would they, individually. Further, while participants believed that groups would use gender stereotypes as a reason to dislike someone who challenged their group's gender norms, participants themselves did not rely upon stereotypes. For both individual and group favorability, participants who rated favorability negatively made frequent reference to group functioning, discussing for instance, how engaging in a different activity would disrupt the group. However, for individual favorability, they cited their own personal preferences (i.e., "I wouldn't like to do ballet, either.") while for group favorability, they cited stereotypes (i.e., "Because most girls are girly and want to do girly things like ballet").

Those participants who liked the deviant member of the group or expected the group to like the deviant member frequently cited autonomy (i.e., "it's okay to be different"). Interestingly, however, they also frequently mentioned the benefit of including diverse perspectives in a group (i.e., "it's good to encourage them to try something new. It will open their minds"). This form of reasoning indicates that those

children and adolescents who support challenging gender stereotypes recognize the power that peers can have and see their peers as potential positive influences who can change group perceptions. Future research should further explore this type of reasoning to unpack precisely under what conditions children do perceive the benefits of including diverse perspectives and if inclusion of such perspectives can actual shape or change group norms.

Age-related differences were documented which indicate that, with age, group norms may become more embedded and, thus, more difficult to change, however. Specifically, younger children were more likely to expect that their peers would challenge the group and were more likely to assert that they, too, would challenge the group. Even though much research highlights a heavy focus on autonomy in adolescence (Smetana & Metzger, 2008) and popular culture assumes that teens are apt to rebel, peer pressure also plays a formidable role during adolescence (Brechwald & Prinstein, 2011) and peer cliques are firmly established during adolescence (Brown, 1990). This may explain, then, why younger children were more likely to expect peers to challenge their groups. Interestingly, however, children also did not differ in their assumptions of how likely their peers would be to challenge the group and how likely they said they would be to challenge the group. Adolescents, on the other hand, indicated that they were more likely to challenge the group than they expected their peers would be. This may suggest a greater sophistication by adolescents, who are able to understand that while they would really like to challenge gender stereotypes, their peers may not be so willing to do so. This finding extends previous research which indicates that adolescents may be more attuned to differences between individual and group perspectives than are children when

considering resource allocation decisions (Mulvey et al., under review). While in this study, all participants distinguished between group and individual favorability of the deviant members, younger children did not distinguish their own expectations for resisting the group from their understanding of how likely peers would be to challenge the group. Thus, together these findings indicate that children may show deficits in distinguishing their own view from that of another individual or group that manifest differently depending on the context.

This study also documented gender differences between how likely boys and girls were to resist group norms and to expect peers to resist group norms. In both the football and ballet conditions, girls asserted that they would individually be more likely to challenge the group than they expected a peer would be. For boys, there were no differences between their expectations for themselves and a peer. Thus, it may be that female participants have more personal experience with stereotypes limiting their opportunities and thus are both more attuned to the likelihood that others will use stereotypes as well as more likely to resist these stereotypes themselves.

Children and adolescents also showed a nuanced understanding of the repercussions for challenging the group and for exhibiting gender non-conforming behavior. Participants generally did not think that challenging the group's norms regarding gender stereotypic social activities would be grounds for exclusion from the group, except in the condition where a boy challenged his boys' group to do ballet. No previous research has examined the consequences in terms of social exclusion for challenging one's group, however research has indicate that boys who counter gender stereotypes are judged more harshly by their peers than are girls who counter gender

stereotypes (Smetana, 1986; Zucker et al., 1995). Participants focused primarily on group functioning when making these decisions, though gender stereotypes did play a role for those participants who believed that the deviant would be excluded from the group. These results indicate that children should be willing to challenge their group's norms regarding gender stereotypic activities, however. Though this may be more difficult for boys, children and adolescents do not believe that challenging the group's stereotypic norms will absolutely result in exclusion. Thus, children and adolescents may very well be able to influence their peers and help to eradicate gender stereotypes.

This suggestion that peers may be able to influence their group is supported by findings on inclusion in the present study. Gender stereotypes did not prove to be a barrier when considering who to invite to join your group. Instead, participants asserted that a group would be willing to allow an opposite gendered child to join their group playing football 80% of the time and would allow an opposite gendered child to join their group doing ballet 60% of the time. While there were differences between the ballet and football conditions, participants were, generally willing to include an outgroup member into their group. Thus, children may perceive that it is more difficult to encourage the entire group to change its norm to one which is gender non-conforming, than it is to accept a gender non-conforming child into a group with an established norm. This is supported by the lack of version effects, even in the ballet condition. While participants asserted that boys who challenged the group to engage in ballet would be treated the most harshly by their group, there were no differences between boys and girls who wanted to join a group of opposite gender children whose norm was to play ballet. Additionally, these findings suggest that future research should examine if seeing individual children

who challenge gender stereotypes by engaging in non-stereotypic activities can, in fact, change attitudes about gender stereotypes or serve as an example which normalizes non-stereotypic activities.

Gendered Forms of Aggression

While gender stereotypes drove many of the evaluations given by children and adolescents of a peer who challenges group norms involving gender stereotypic social activities (football and ballet), this was not the case for gendered forms of aggression. Most notably, while participants did still exhibit knowledge of the stereotype associating boys with physical aggression and (less so) girls with relational aggression, and while gender stereotypes did still play a significant role in participants' reasoning about their evaluations in some conditions, participants did not systematically evaluate gender non-conforming challenges to aggressive behavior as less acceptable than gender conforming challenges. On the other hand, evaluations of challenges to gendered forms of aggression were driven largely by the valence of the challenge: participants showed support for deviance which also countered aggression (physical and relational) and less support for deviance which encouraged aggressive behavior.

Children and adolescents asserted that they would be more likely to challenge their groups to be impartial when the group gossips and to be nice when the group plays rough than they were to challenge the group to engage in gossip or to play rough. Further, they individually asserted that they would be more likely to challenge the group than they expected a peer to be. These findings indicate that children and adolescents are driven to challenge aggressive behaviors. Further, it is important to note that the proportion of participants asserting that peers would challenge group norms was

generally quite high with rates of close to 70% or higher; this was paralleled with the findings for participants' own perspective (how they would react). Thus participants did believe that their peers would stand up to groups and challenge norms involving gendered forms of aggression. They also asserted that they would do the same. Interestingly, unlike in the gender stereotypic activities context, there were no systemic age-related differences. This may suggest that while adolescents expected peers to be more reserved in the context of stereotypic activities than did children, the strong moral dimension to group norms involving aggressive behavior may have impacted the judgments of both children and adolescents, with all participants indicating high levels of resistance. Given the consistent findings indicating the powerful influence of peers and the pervasive role of peer pressure (Brechwald & Prinstein, 2011), these results provide an optimistic picture.

Children and adolescents want to challenge aggressive group norms and expect that their peers will also often challenge such behaviors. This is centrally important as research has shown that bystanders can make a difference in reducing incidences of bullying and aggression (Barchia & Bussey, 2011; Salmivalli et al., 2011; Trach, Hymel, Waterhouse, & Neale, 2010). What these findings indicate is that we should look not only to third - party bystanders, but also to those children and adolescents who are actually part of peer groups which engage in aggression. Just because the group supports an aggressive norm does not mean that the members of this group agree with that norm, and, as these findings suggest, they may, in fact, be willing to actually challenge that norm.

Favorability Judgments. The results also indicated that children and adolescents believe that groups will adhere to their norms and dislike deviant members. Specifically, ratings for group favorability of a deviant were always negative—in all conditions they never crossed the mid-point. This confirms previous research on developmental subjective group dynamics (Abrams & Rutland, 2008), which has extensively demonstrated that groups do not like deviant members, especially in the context of social-conventional group norms. However, it also extends recent research which showed that groups also do not like deviants when the group norms are moral (Killen, Rutland, et al., 2012). The Killen, Rutland, et al. (2012) study demonstrated that groups will not like deviants who reject group norms involving equal or unequal distribution of resources.

The current study extended this finding to a different morally relevant context: aggression. The current study also extended these findings by demonstrating that groups dislike deviants regardless of if they deviate in gender stereotype conforming or non-conforming ways. However, the findings also reveal that participants do think that groups will attend to the nature of the norm; they were significantly more negative about deviants who advocated *for* aggressive behavior than about deviants who urged their group to reject aggressive behavior. Finally, participants also exhibited a careful awareness of group dynamics. They perceived that groups would be more favorable to a deviant who suggests that the group play rough than to a deviant who suggests that the group be impartial and avoid gossip, perhaps recognizing the ostensible benefit in terms of winning a sporting game that playing rough could garner. Participants' reasoning about group favorability also reflected their keen awareness of group dynamics: across all

conditions they focused on how the deviant's behavior would impact the functioning and conventions of the group in their justifications.

When rating their own individual favorability toward the nice and rough deviant members, children and adolescents exhibited age-related differences. Younger children were more supportive of the nice and impartial deviants and less supportive of the rough deviant than were adolescents. This finding extends findings indicating that younger children show greater support for equal distribution of resources than do adolescents (Almås, Cappelen, Sørensen, & Tungodden, 2010; Killen, Rutland, et al., 2012) by indicating that children show greater support for a wide range of behaviors that align with moral principles. The reasoning results indicate, that these differences are driven by differing foci in adolescents and children: children focused more on the harm to others that playing rough could cause while adolescents focused more on how playing rough or nicely would impact the group's ability to function. An important extension to this research would be to compare findings from normative populations and from highly aggressive populations. For instance, children who are part of groups for which aggression is the defining characteristic, such as gangs, will likely make much different judgments and may even be more likely to condone stereotypes associated with gender.

Gender differences also emerged for individual favorability ratings for the nice deviant: female participants were more supportive of this deviant than were male participants. This pattern aligns with gender stereotypes suggesting that playing nice is more acceptable for females while playing rough is more acceptable for males. Thus, even though research has shown that associations between boys and physical aggression may be founded largely on stereotypes (Card et al., 2008; Lansford et al., 2012),

implicitly these stereotypes appear to still influence individual evaluations. Findings such as these, which reveal that boys are less supportive of playing nice than are girls may help to explain why mean differences in rates of physical aggression (with boys demonstrating higher rates) were still found in the meta-analysis conducted by Card et al. (2008). This finding could also be explained by recent research that indicates that girls who are aggressive (physical or relational) are more likely to be rejected and excluded than are boys who are aggressive (Kochel et al., 2012). If being physically aggressive is more societally condoned for boys, this may be reflected in lower favorability ratings by boys' of gender non-conforming avoidance of physical aggression.

The results of analyses conducted to assess if participants were individually more favorable to the deviants than they expected groups to be were significant for the nice, rough, and impartial conditions. This confirms previous research indicating that children and adolescents can distinguish between individual and group perspectives regarding deviant members (Killen, Rutland, et al., 2012). Further, it is of note that this distinction was not only found in the non-aggressive conditions (nice and impartial), but also in the rough condition. In this case, participants recognize that while the group will focus on the fact that this individual is deviating from the group's norm, they individually understand that this deviant may, in fact, be trying to help the group win the soccer game.

Exclusion. While these results reveal that groups will not like deviant members, the study also provided insight into the consequences for deviance. Much previous literature has assessed the acceptability of exclusion. Pervasive findings, even cross-culturally (Hitti et al., 2011), show in most situations that children reject exclusion as morally unacceptable, citing harm to the target of exclusion (Killen, 2007; Killen &

Rutland, 2011). In the current study, participants assessed how likely it was that the member who challenged the group's norm would be excluded for encouraging the group to act differently.

Results regarding exclusion revealed critical age-related differences. Children believed that dissenting members would be excluded from the group in almost all conditions (all except impartial), while adolescents were more neutral in their judgments. This suggests, importantly, that younger children may be more concerned about exclusion as a potential consequence for challenging the group than are adolescents. This is counter intuitive, as much research has shown that adolescents are more focused on group functioning. It may be, however, that children's social groups are more fluid and less fixed than are adolescents', which could explain why they perceive exclusion to be more likely. However, previous research indicates that children perceive exclusion to be harmful to the target of exclusion, often citing other's welfare in their reasoning (Killen & Rutland, 2011) and that exclusion can have harmful negative consequences in terms of the excluded child's mental health and academic motivation (Buhs & Ladd, 2001; Buhs et al., 2006). Thus, even if children's social groups are more fluid than are adolescents' groups, exclusion from a group will still cause harm. These results, then, indicate that a significant barrier to challenging one's group, especially for children, is concern over social exclusion.

In addition to the age-related differences, there were also findings suggesting that all participants believed that group members who challenge the group to engage in aggressive behavior (rough and gossip conditions) will also be subject to exclusion. This is important as it suggests that children and adolescents think that groups with positive

moral norms will actively work to preserve these norms by excluding children who incite aggression. This is supported by findings indicating that children with externalizing problems and who exhibit aggressive behavior are more likely to be rejected by their peer groups (Rubin et al., 2006).

An interesting future extension of this research would be to examine how likelihood of exclusion would vary based on the social status of the dissenting member. For instance, many social groups include a leader or a group of leaders. Research with children indicates that children and adolescents who are perceived to be more popular also often exhibit aggressive behaviors, especially relationally aggressive behaviors (Rose, Swenson, & Waller, 2004). Additionally, research with adults indicates that one's social role in a group can impact how the group responds to deviance from group norms (Abrams, De Moura, Marques, & Hutchison, 2008; Pinto, Marques, Levine, & Abrams, 2010). The findings from the current study which indicate a relation between group favorability and exclusion also indicate the importance of further examining one's social role in a group. Participants asserted that if the group did not like a deviant, they were more likely to exclude this deviant. In this study, the group favorability question focused specifically on favorability in the context of challenging the group. Future research should also examine if groups differ in their responses to well-liked or popular group members who challenge the group and to more marginalized group members who challenge the group.

Another way to examine how groups would respond to challenges to their group norm is to examine if groups would rather include someone into their group who shares their gender identity (maintaining a homogenous gender group identity) but challenges

their group norm, or someone who shares their group norm, but is of the opposite gender. Results for inclusion choice indicate that as expected, participants were more likely to include non-aggressive outgroup members into non-aggressive groups than aggressive outgroup members into aggressive groups. In fact, they included outgroup members more than at chance for all groups except these gossiping group which suggests that participants balanced information about group membership and group norm in making these decisions. Interestingly, their reasoning often focused on group functioning, but those participants who chose to maintain the gender composition of the group often did explicitly reference the importance of maintaining a gender identity in the group.

Children employ complex social reasoning skills in all of their evaluations of group members who challenge the norms of their groups involving aggression. Though they did not overwhelmingly use stereotypes in their evaluations, participants balanced information about gender, group composition and functioning and moral principles in their evaluations.

Theory of Mind. An additional question in this study was whether any of the variance in participants' responses could be accounted for by their social-cognitive capacities, in particular interpretative theory of mind. Interpretative theory of mind (Carpendale & Chandler, 1996) was chosen as it is a more complex form of theory of mind, which would be more appropriate for the age groups tested in this study. However, participants at both age-groups exhibited a high level of competence on the measure. At the same time, associations were found between interpretative theory of mind and individual favorability ratings for the impartial and nice deviant members. This suggests that increased perspective-taking abilities may enable one to better understand why

someone would challenge the group, and thus to respond appropriately to such challenges. Yet, relations were only found for the individual favorability ratings, indicating that interpretative theory of mind aided participants in understanding the dissenting member's perspective, but not, for instance, the group's perspective. This indicates that more work needs to be done to better understand what social-cognitive skills are employed when making decisions in group contexts and evaluating group dynamics. This is especially true given that other theory of mind measures, specifically theory of social mind (Abrams et al., 2009), have been demonstrated to relate to children's understanding of group dynamics. In the future, researchers should consider other measures of social cognition and examine their relation to understanding of group dynamics. In particular, executive functioning has been shown to develop through adolescence, and thus may be a good candidate (Gogtay et al., 2004). Further, it may be useful to examine differences between populations of children who have been trained using Shure's (1992) I Can Problem Solve Methods versus those who have not. Additionally, examining differences in children's mindfulness (Greenberg & Harris, 2012) or social information processing skills (Dodge & Coie, 1987) may be useful. Across many of the measures in the current study, participants did, however, exhibit skill in distinguishing group and individual perspectives and understanding the nuances of social decision-making.

Conclusions

Contributions to the Literature

The findings from this study revealed the complexity of children's social lives and the challenges they face in responding to peer group norms with which they may not

agree, whether because they are founded upon stereotypes or because they cause harm to others. Gender segregation frequently occurs in children's lives. Research on the role of teachers in perpetuating gender segregation highlights the powerful role that functional labeling of gender can have (Patterson & Bigler, 2006). Further, children do spend much of their time with peers of the same gender and their affiliations with same gendered peers are stable over time (Martin & Fabes, 2001). Even more concerning, the more children affiliate with peers of the same gender, the more gender differentiated their behavior is (Martin & Fabes, 2001). However, in the current study, children and adolescents were quite willing to include someone of the opposite gender into their own group. This is an important new finding as it indicates that children do not always desire or prefer single-gender groups. This is especially important given research that indicates that when in mixed-gender contexts, children may be more willing to engage in non-stereotypic activities (Goble, Martin, Hanish, & Fabes, 2012). The results for the gender stereotypic activities indicate, however, that while children are optimistic about challenging gender stereotypes, these stereotypes still persist and do impact behaviors and social-decisions. Counter-stereotypic behavior was judged negatively especially for boys. This suggests that status differences between boys and girls still play a pervasive and pernicious role in children's lives and that we should be particularly attuned to how stereotypes can limit opportunities not only for girls, but also, importantly, for boys.

The results for the gendered forms of aggression confirm that associations between each gender and particular forms of aggression are likely just stereotypes as children and adolescents did not show the same types of distinctions between girls and boys challenging each type of aggression as they did for girls and boys challenging each

type of social activity. This provides important confirmation that new research findings indicating that the previously identified associations between gender and aggression may be unfounded (Card et al., 2008). Recent research has been conducted which documents that behaviorally boys engage in more physical aggression than do girls, but that there are no differences in rates of relational aggression. The current study included measures of judgments and reasoning, with findings suggesting that children evaluate aggressive behavior and deviance from group norms about aggression by focusing on the moral valence of the behavior and not on the gender of the individual engaging a particular behavior. Differences between evaluations of boys and girls groups were not documented in this study. Further, the study indicates that children and adolescents are willing to challenge their groups and that they expect their peers to do so as well, especially in support of non-aggressive behavior. However, results also revealed that children and adolescents do perceive social exclusion to be a very real consequence for challenging the peer group. This reveals the social competency of children: they are aware of the potential consequences in terms of exclusion, but are still hopeful that they and their peers will challenge the group.

Limitations and Future Directions

This study provides greater understanding of children's willingness to challenge their group, their expectations about their peers, and their understanding of the consequences of challenging the group. However, it does include limitations. For instance, likelihood of challenging the group was measured using a dichotomous variable, which cannot capture subtle nuances in difference in likelihood of challenging the group in different contexts. Future research should use a likert scale to measure

likelihood of resistance. Additionally, the gender stereotypic social activities tested, football and ballet, involve very strong gender stereotypes. Future research should examine a greater range of gender stereotypic activities. An additional limitation is that the only consequence for challenging the group which was tested was exclusion. It is possible that children did not think that they would be excluded for challenging the group, but rather that they may be teased or shunned for a short time. It would be interesting in the future to allow children to spontaneously generate possible consequences as well as to assess if they think the group will listen to the deviant member.

Future research should work to continue to identify what factors contribute to children actually challenging the group (as opposed to just asserting a desire to do so). Additionally, this study examined middle childhood and adolescence. An important future direction would be to examine these patterns in younger children. This extension would allow for an examination of the origins of a willingness to challenge the group. Further, using a younger sample may capture more relations with social-cognitive skills, as there is more variation in social cognitive abilities, for instance in theory of mind abilities, in younger children. In addition to examining different age groups, it would be interesting to test for differences by ethnic or cultural group. For instance, some cultural groups may hold much stronger gender stereotypes and thus, in those cultures, challenging these stereotypes may be viewed more negatively. Finally, future research should examine individual differences in children themselves, to help identify what types of children are likely to challenge their groups and to be effective in actually changing the norms of the group.

Implications

This study has implications for educators, and parents, in particular. The new findings indicate that parents and teachers should support children's autonomy within their social groups and allow them to negotiate social contexts without immediate intervention. Children and adolescents do want to challenge stereotypic and aggressive peer group norms and they are optimistic that their peers will do the same. Parents and teachers should allow peers to resolve these conflicts on their own, recognizing the sophistication of children's understanding of social group dynamics. At the same time, parents and teachers should work to create spaces whereby children can feel comfortable challenging gender stereotypes and gender segregation. The results also indicate that children and adolescents may be concerned about social exclusion as a consequence for challenging the peer group. Thus, parents and teachers should create opportunities for children to practice challenging peers, for instance through classrooms which are open to discussion and debate, to give children the skills needed to follow through on their desire to challenge unacceptable peer group norms and overcome concerns about exclusion.

In conclusion, this study provides evidence that children are willing to challenge gender stereotypic group expectations and norms, but that they are also often concerned about consequences such as exclusion. Further, this study reveals the complexity of children's and adolescents' understanding of group dynamics and the sophistication of their social reasoning in complex contexts.

Figure 1: Task Design

	Version 1 CRC: Conform, Resist, Conform	Version 2 RCR: Resist, Conform, Resist
	Conventional Story C (Conform)	Conventional Story R (Resist)
Story1: Girls Group	Group Norm: conform to stereotype “Let’s do ballet, that’s for girls.” Deviant Behavior: advocate rebellion “People think football is only for boys, let’s do football.”	Group Norm: advocate rebellion “People think football is only for boys, let’s do football.” Deviant Behavior: conform to stereotype “Let’s do ballet, that’s for girls.”
Story2: Boys Group	Group Norm: conform to stereotype “Let’s do football, that’s for boys.” Deviant Behavior: advocate rebellion “People think ballet is only for girls, let’s do ballet.”	Group Norm: advocate rebellion “People think ballet is only for girls, let’s do football.” Deviant Behavior: conform to stereotype “Let’s do football, that’s for boys.”
	Moral Story R (Resist)	Moral Story C (Conform)
Story 3: Girls Group	Group Norm: Impartiality “It’s important not to gossip all the time, even though we won’t find out why those kids act that way.” Deviant Behavior: Gossip “It’s okay to gossip all the time, because we don’t know why those kids act that way.”	Group Norm: Gossip “It’s okay to gossip all the time, because we don’t know why those kids act that way.” Deviant Behavior: Impartiality “It’s important not to gossip all the time, even though we won’t find out why those kids act that way.”
Story 4: Boys Group	Group Norm: Gossip “It’s okay to gossip all the time, because we don’t know why those kids act that way.” Deviant Behavior: Impartiality “It’s important not to gossip all the time, even though we won’t find out why those kids act that way.”	Group Norm: Impartiality “It’s important not to gossip all the time, even though we won’t find out why those kids act that way.” Deviant Behavior: Gossip “It’s okay to gossip all the time, because we don’t know why those kids act that way.”
	Moral Story C (Conform)	Moral Story R (Resist)
Story 5:	Group Norm: Caring	Group Norm: Rough

Girls Group	<p>“It’s important to play nicely, even if you don’t score as many points.”</p> <p>Deviant Behavior: Aggressive</p> <p>“It’s okay to push and shove just to score points.”</p>	<p>“It’s okay to push and shove just to score points.”</p> <p>Deviant Behavior: Caring</p> <p>“It’s important to play nicely, even if you don’t score as many points.”</p>
Story 6: Boys Group	<p>Group Norm: Rough</p> <p>“It’s okay to push and shove just to score points.”</p> <p>Deviant Behavior: Caring</p> <p>“It’s important to play nicely, even if you don’t score as many points.”</p>	<p>Group Norm: Caring</p> <p>“It’s important to play nicely, even if you don’t score as many points.”</p> <p>Deviant Behavior: Aggressive</p> <p>“It’s okay to push and shove just to score points.”</p>

Table 1: Summary of Hypotheses and Analyses

Measure	Analysis and Hypothesis
Likelihood of Resistance	<p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) ANOVA.</p> <p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) X 2 (condition) ANOVA with repeated measures on the last factor, for each pair of conditions.</p> <p>It is expected that participants will be more likely to expect targets will resist the group when the group is stereotype non-conforming than when the group conforms to stereotypes. Further, resistance will be less likely for the gender stereotypic activities when the group adheres to stereotypes than when they do not. It is expected that participants will least expect resistance from a boy who wants to do ballet when his group wants to do football. It is expected that children will see deviance as more likely than will adolescents. It is expected that participants, themselves, will be more attuned to challenging aggressive behavior than will they expect their peers to be. Thus, it is expected that participants will rate their own likelihood of resistance to aggressive norms to be higher than the ratings they provide for their peers.</p>
Likelihood of Resistance Versus Individual Likelihood of Resistance	<p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) X 2 (question: likelihood of resistance, individual likelihood of resistance) ANOVA with repeated measures on the last factor, for each condition.</p> <p>It is expected that participants, themselves, will be more attuned to challenging aggressive behavior than will they expect their peers to be. Thus, it is expected that participants will rate their own likelihood of resistance to aggressive norms to be higher than the ratings they provide for their peers.</p>

Group favorability toward dissenting member	<p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) ANOVA.</p> <p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) X 2 (condition) ANOVA with repeated measures on the last factor, for each pair of conditions.</p> <p>Participants will likely expect that groups will not like dissenting members, but this will vary if the dissenter is resisting or condoning aggression. Further, groups will be more favorable to dissenting members who resist aggression, as aggression will likely be viewed as a moral transgression. There may be a shifting standard, with participants asserting that groups will find deviance towards non-conforming behaviors as less acceptable, especially for boys who want to do ballet.</p>
Group favorability toward dissenting member: reasoning	<p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) X 2 (group favorability: ok, not okay) X 3 (reasoning: top three forms of reasoning) ANOVA with repeated measures on the last 2 factors, for each of the 6 conditions.</p> <p>Participants will cite group functioning, stereotypes, and other societal justifications when asserting that the group will not like dissenting members who resist stereotypes.</p> <p>Participants will cite moral reasons when asserting that the group will like dissenting members who do not gossip or push and shove.</p>
Individual favorability toward dissenting member	<p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) ANOVA.</p> <p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) X 2 (condition) ANOVA with repeated measures on the last factor, for each pair of conditions.</p> <p>Participants will support dissenting members who adhere to generic moral principles and resist aggression, regardless of their gender. Non-conformity may be seen as more positive for girls, than for boys, because participants may believe that boys should not move down the status ladder to act in stereotypically female ways.</p>
Individual favorability toward dissenting member: reasoning	<p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) X 2 (group favorability: ok, not okay) X 3 (reasoning: top three forms of reasoning) ANOVA with repeated measures on the last 2 factors, for each of the 6 conditions.</p> <p>Participants will reference moral reasons when supporting the dissenting</p>

	<p>members who avoid aggression, and will reference harm, in particular, when evaluating the dissenting members who engage in aggression more negatively.</p> <p>With age, participants will reference autonomy, saying that it is up to the dissenting member.</p>
Likelihood of exclusion	<p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) ANOVA.</p> <p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) X 2 (condition) ANOVA with repeated measures on the last factor, for each pair of conditions.</p> <p>Participants will identify exclusion as a likely repercussion for resisting the group by encouraging the group to avoid aggression.</p> <p>Children may be more likely to expect exclusion than are adolescents.</p>
Likelihood of exclusion: reasoning	<p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) X 2 (group favorability: ok, not okay) X 3 (reasoning: top three forms of reasoning) ANOVA with repeated measures on the last 2 factors, for each of the 6 conditions.</p> <p>Participants who believe exclusion is likely will cite group functioning and stereotypes more often than those who do not.</p>
Inclusion choice	<p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) ANOVA.</p> <p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) X 2 (condition) ANOVA with repeated measures on the last factor, for each pair of conditions.</p> <p>Dissenting member who go against moral principles (by gossiping and playing rough) are less likely to be included.</p>
Inclusion choice: reasoning	<p>2 (age group: 4th, 8th) X 2 (gender: male, female) X 2 (stereotype norm: conforming, resisting) X 2 (group favorability: ok, not okay) X 3 (reasoning: top three forms of reasoning) ANOVA with repeated measures on the last 2 factors, for each of the 6 conditions.</p> <p>Reasoning will likely focus on moral reasons when participants choose to include the non-aggressive target, regardless of if the target is conforming to or resisting stereotypes.</p>

	Age Group		Gender	
	9 year olds	13 year olds	Females	Males
Football	75.6	82.6	75.8	86.1
Ballet	85.6	83.6	79.1	89.8
Rough	73.9	73.5	69.7	77.8

Table 2: Percentage of Participants Affirming Stereotypes Regarding Who Usually Engages in Each Type of Activity or Behavior

Nice	61.1	64.3	59.6	67.2
Gossip	47.8	57.0	51.0	58.1
Impartial	41.6	71.9	60.4	64.9

Figure 2: Likelihood of Exclusion of the Deviant Member

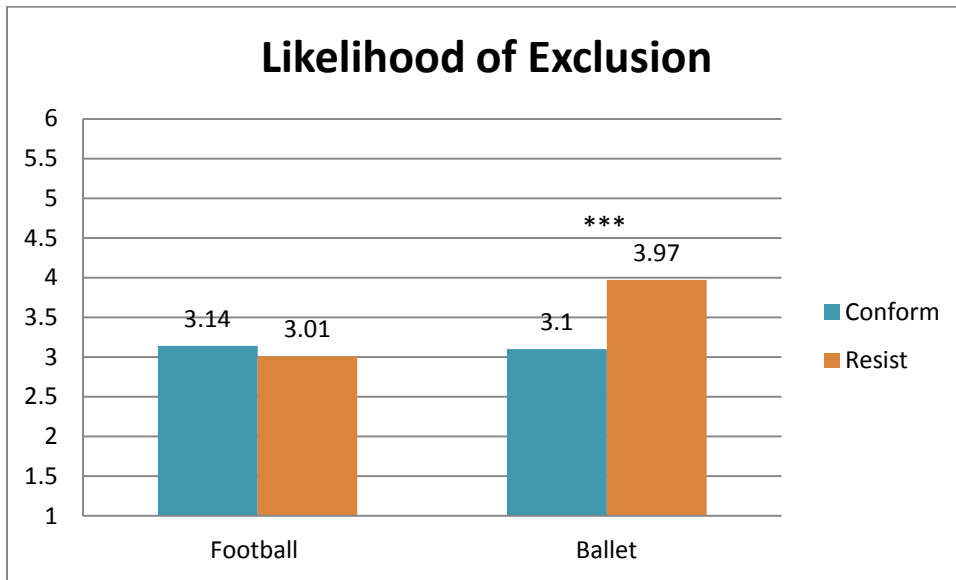


Figure 3: Deviant who Wants to do Ballet: Reasoning by Exclusion Likelihood

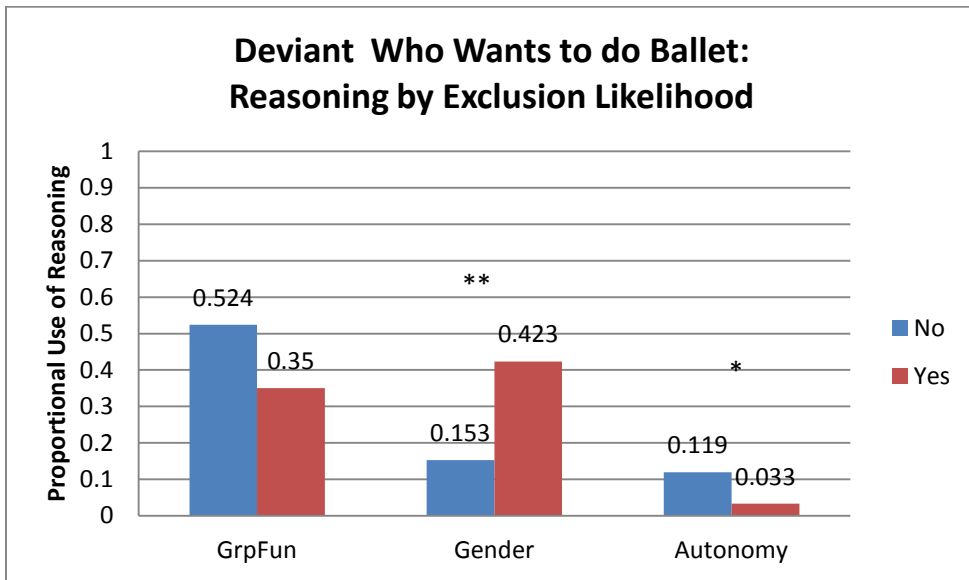


Figure 4: Likelihood of Resistance Versus Individual Likelihood of Resistance to Group Norms about Aggression

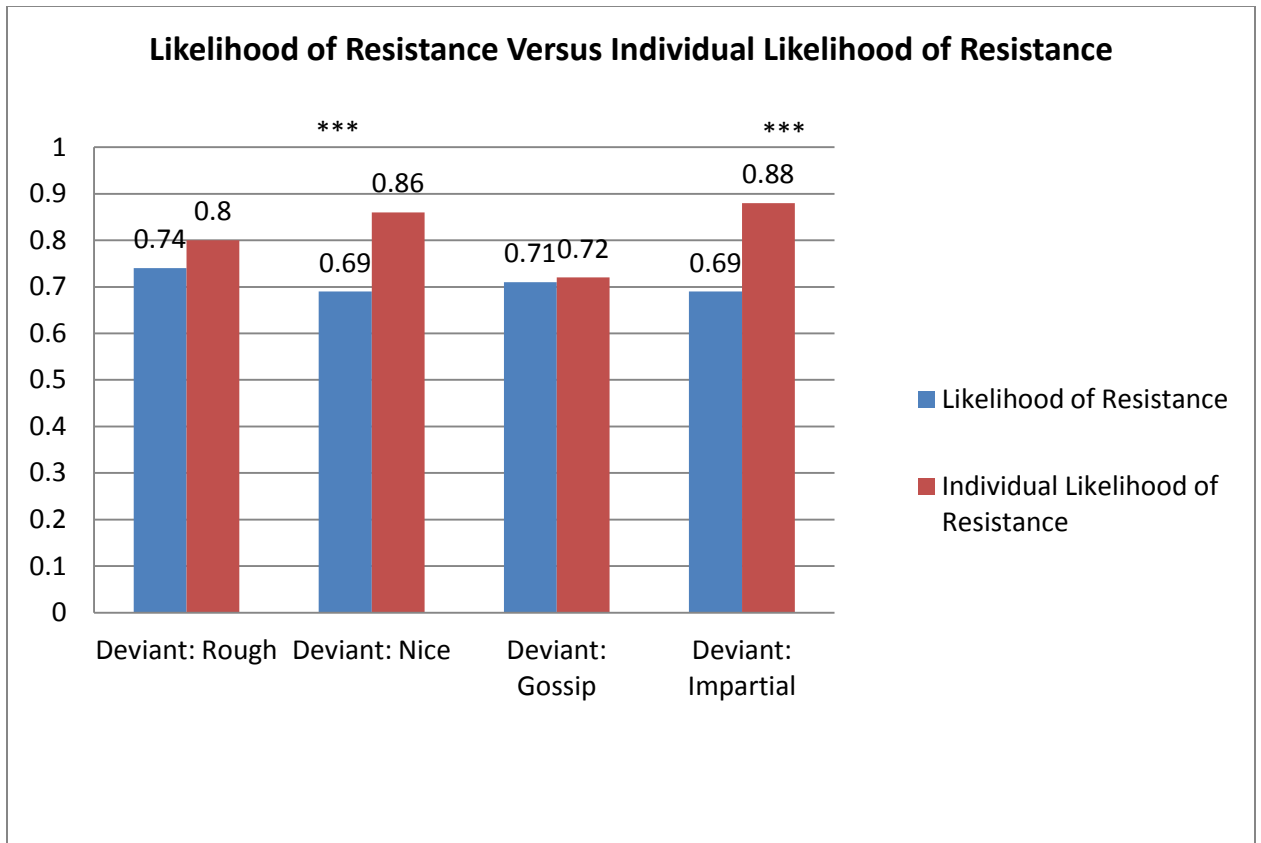


Table 3: Means and Standard Deviations for Proportion of Justifications Used for Evaluations of Group Favorability Toward the Deviant Member

	Group Functioning	Autonomy	Welfare	Gender
Rough				
Not Okay	.40 (.47)	.03 (.17)	.41 (.47)	.05 (.21)
Okay	.37 (.48)	.21 (.41)	.14 (.34)	.08 (.26)
Total	.39 (.47)	.09 (.28)	.32 (.45)	.06 (.22)
Nice				
Not Okay	.50 (.49)	.02 (.15)	.07 (.22)	.04 (.19)
Okay	.32 (.46)	.14 (.35)	.12 (.32)	.00 (.00)
Total	.42 (.49)	.08 (.26)	.09 (.27)	.02 (.14)
Gossip				
Not Okay	.37 (.47)	.01 (.10)	.44 (.48)	.05 (.20)
Okay	.25 (.43)	.21 (.41)	.10 (.30)	.14 (.34)
Total	.34 (.46)	.06 (.23)	.36 (.47)	.07 (.25)
Impartial				
Not Okay	.59 (.47)	.04 (.18)	.11 (.29)	.11 (.29)
Okay	.28 (.44)	.17 (.37)	.24 (.41)	.12 (.32)
Total	.46 (.48)	.09 (.28)	.17 (.35)	.12 (.31)

Note. Okay and Not Okay evaluations of favorability are based on a midpoint split of 3.5 for responses to a Likert scale ranging from 1 = Really Not Okay to 6 = Really Okay

Table 4: Means and Standard Deviations for Proportion of Justifications Used for Evaluations of Individual Favorability Toward the Deviant Member

	Group Functioning	Autonomy	Welfare	Personal Identification
Rough				
Not Okay	.13 (.32)	.01 (.09)	.63 (.46)	.12 (.30)
Okay	.25 (.43)	.62 (.44)	.11 (.31)	.17 (.38)
Total	.19 (.38)	.13 (.33)	.39 (.47)	.14 (.34)
Nice				
Not Okay	.24 (.42)	.04 (.21)	.18 (.38)	.21 (.39)
Okay	.14 (.33)	.21 (.40)	.41 (.48)	.13 (.32)
Total	.16 (.36)	.17 (.37)	.36 (.47)	.15 (.33)
Gossip				
Not Okay	.08 (.26)	.03 (.17)	.64 (.46)	.16 (.35)
Okay	.14 (.34)	.36 (.48)	.17 (.35)	.17 (.36)
Total	.09 (.29)	.13 (.32)	.50 (.48)	.16 (.35)
Impartial				
Not Okay	.20 (.40)	.02 (.15)	.31 (.46)	.15 (.34)
Okay	.06 (.23)	.18 (.37)	.45 (.47)	.17 (.36)
Total	.08 (.27)	.16 (.36)	.42 (.47)	.17 (.35)

Note. Okay and Not Okay evaluations of favorability are based on a midpoint split of 3.5 for responses to a Likert scale ranging from 1 = Really Not Okay to 6 = Really Okay

Figure 5 Group Favorability Versus Individual Favorability

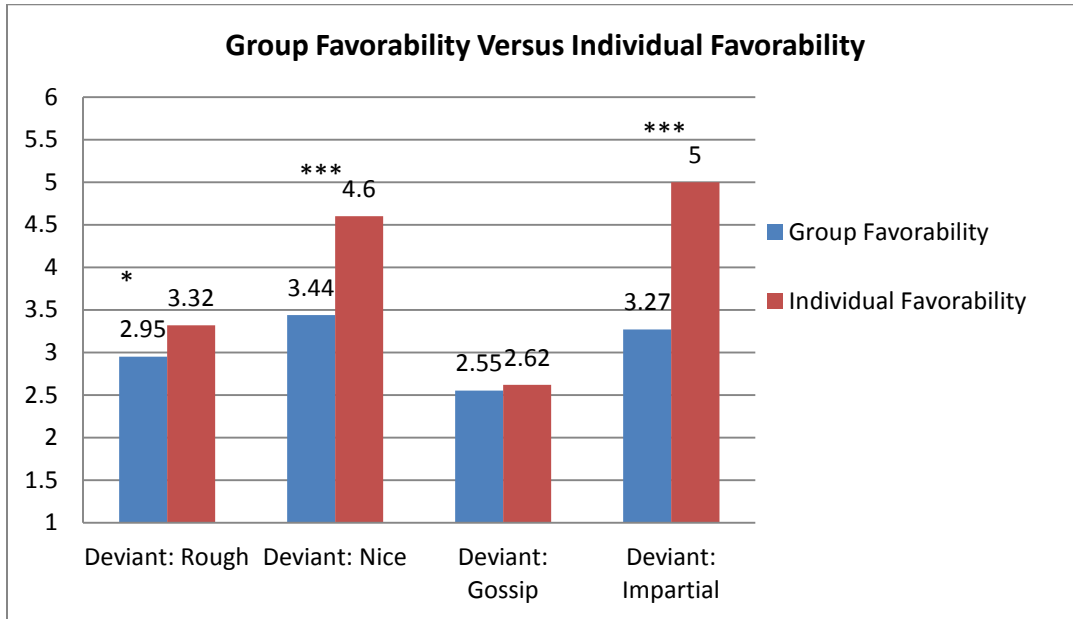


Figure 6 Likelihood of Exclusion

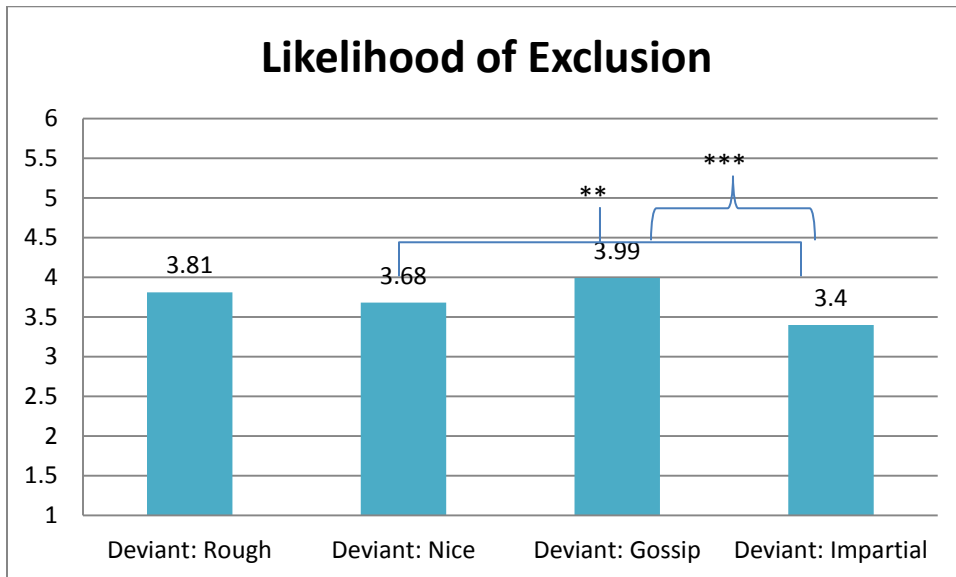


Figure 7 Likelihood of Exclusion by Age Group

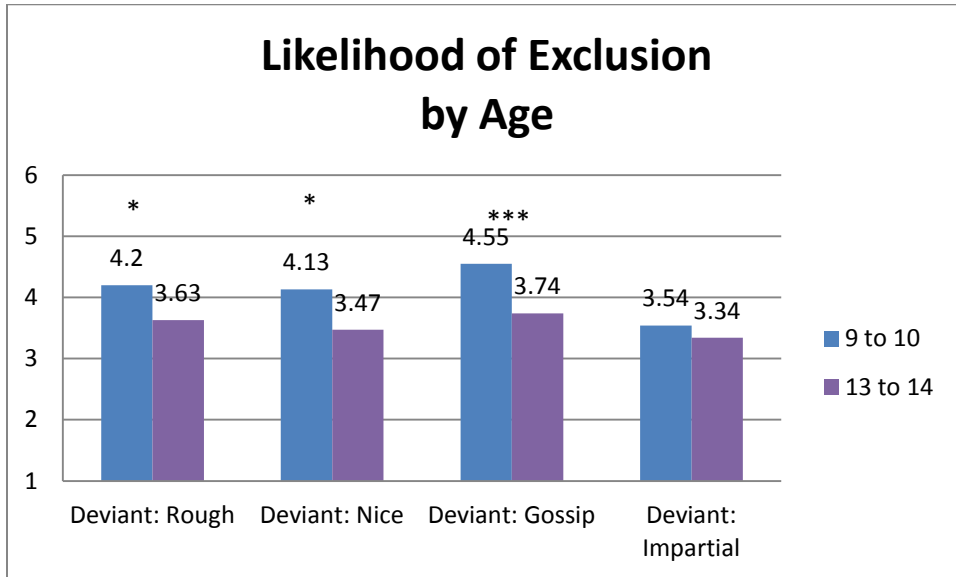


Table 5: Means and Standard Deviations for Proportion of Justifications Used for Evaluations of Likelihood of Exclusion of the Deviant Member

	Group Functioning	Diverse Perspectives	Welfare	Gender Stereotypes
Rough				
Not Likely	.58 (.49)	.04 (.18)	.09 (.26)	.06 (.23)
Likely	.44 (.48)	.01 (.11)	.36 (.47)	.07 (.23)
Total	.50 (.49)	.02 (.14)	.23 (.41)	.07 (.23)
Nice				
Not Likely	.56 (.49)	.09 (.27)	.11 (.30)	.03 (.17)
Likely	.59 (.48)	.02 (.14)	.10 (.28)	.08 (.27)
Total	.58 (.49)	.05 (.21)	.10 (.29)	.06 (.22)
Gossip				
Not Likely	.61 (.47)	.06 (.23)	.06 (.21)	.07 (.24)
Likely	.49 (.49)	.00 (.00)	.36 (.46)	.03 (.15)
Total	.54 (.48)	.02 (.15)	.23 (.40)	.04 (.19)
Impartial				
Not Likely	.40 (.47)	.15 (.36)	.16 (.34)	.09 (.27)
Likely	.67 (.45)	.00 (.00)	.07 (.25)	.10 (.28)
Total	.53 (.48)	.08 (.27)	.12 (.31)	.10 (.27)

Note. Likely and Not Likely evaluations of favorability are based on a midpoint split of 3.5 for responses to a Likert scale ranging from 1 = Really Not Likely to 6 = Really Likely

Figure 8: Likelihood of Exclusion by Dichotomous Group Favorability Evaluation

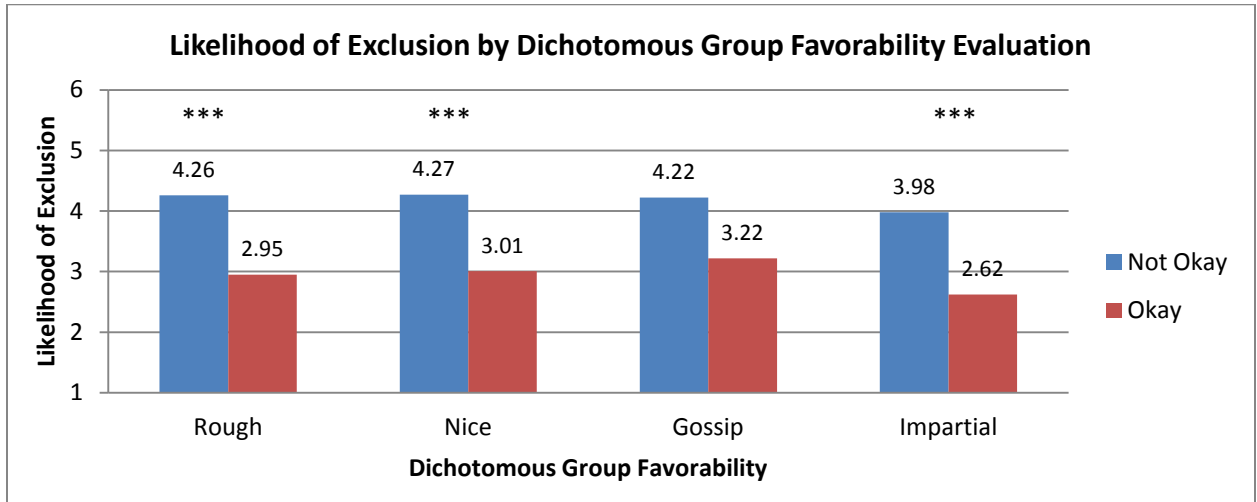


Figure 9: Inclusion of Outgroup Member Who Matches the Group Norm

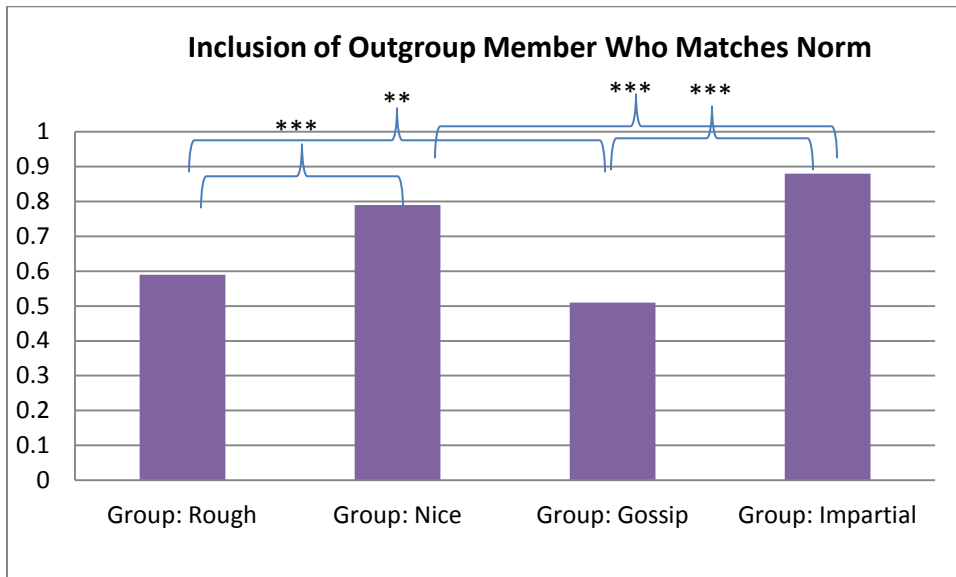


Table 6: Means and Standard Deviations for Proportion of Justifications Used for

Inclusion Choice

	Group Functioning	Diverse Perspectives	Welfare	Gender Stereotypes
Rough				
Ingroup	.33 (.46)	.09 (.29)	.03 (.45)	.13 (.26)
Outgroup	.66 (.46)	.01 (.09)	.13 (.33)	.07 (.28)
Total	.52 (.48)	.04 (.19)	.20 (.39)	.09 (.27)
Nice				
Ingroup	.40 (.48)	.09 (.29)	.07 (.24)	.15 (.33)
Outgroup	.56 (.49)	.15 (.34)	.13 (.33)	.06 (.21)
Total	.52 (.49)	.13 (.33)	.11 (.31)	.08 (.25)
Gossip				
Ingroup	.34 (.45)	.07 (.24)	.33 (.45)	.13 (.31)
Outgroup	.65 (.46)	.03 (.16)	.16 (.36)	.05 (.17)
Total	.50 (.48)	.05 (.20)	.24 (.41)	.09 (.25)
Impartial				
Ingroup	.34 (.47)	.06 (.21)	.03 (.17)	.34 (.45)
Outgroup	.48 (.49)	.18 (.37)	.16 (.35)	.08 (.25)
Total	.46 (.49)	.16 (.36)	.14 (.33)	.11 (.29)

Appendix A: Institutional Review Board Approval and Consent Forms



UNIVERSITY OF
MARYLAND
INSTITUTIONAL REVIEW BOARD

1204 Marie Mount Hall
College Park, MD 20742-5125
TEL 301.405.4212
FAX 301.314.1475
irb@um.edu
www.umresearch.um.edu/IRB

DATE: November 26, 2012

TO: Melanie Killen
FROM: University of Maryland College Park (UMCP) IRB

PROJECT TITLE: [331482-2] Gender Exclusion, Intentionality, and Theory of Social Mind
REFERENCE #: 11-0332
SUBMISSION TYPE: Amendment/Modification

ACTION: APPROVED
APPROVAL DATE: November 26, 2012
EXPIRATION DATE: June 2, 2013
REVIEW TYPE: Expedited Review

REVIEW CATEGORY: Expedited review category # 7. *Subpart D applies, 45CFR46.404*

Thank you for your submission of Amendment/Modification materials for this project. The University of Maryland College Park (UMCP) IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a project design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the project and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the project via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this committee prior to initiation. Please use the appropriate revision forms for this procedure which are found on the IRBNet Forms and Templates Page.

All UNANTICIPATED PROBLEMS involving risks to subjects or others (UPIRSOs) and SERIOUS and UNEXPECTED adverse events must be reported promptly to this office. Please use the appropriate reporting forms for this procedure. All FDA and sponsor reporting requirements should also be followed.

All NON-COMPLIANCE issues or COMPLAINTS regarding this project must be reported promptly to this office.

This project has been determined to be a Minimal Risk project. Based on the risks, this project requires continuing review by this committee on an annual basis. Please use the appropriate forms for this procedure. Your documentation for continuing review must be received with sufficient time for review and continued approval before the expiration date of June 2, 2013.

Please note that all research records must be retained for a minimum of three years after the completion of the project.

If you have any questions, please contact the IRB Office at 301-405-4212 or irb@umd.edu. Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within University of Maryland College Park (UMCP) IRB's records.

**Initial Application Approval**

DO NOT REPLY TO THIS EMAIL ADDRESS AS IT IS UNMONITORED

To: Principal Investigator, Dr. Melanie Killen, Human Development

Student, Kelly Lynn Mulvey, Human Development

James M. Hagberg

From: IRB Co-Chair

University of Maryland College Park

Re: IRB Protocol: 11-0332 - Gender Exclusion, Intentionality, and Theory of
Social Mind

Approval Date: June 02, 2011

Expiration
June 02, 2012
Date:

Application: Initial

Review Path: Expedited

The University of Maryland, College Park Institutional Review Board (IRB) Office approved your Initial IRB Application. This transaction was approved in accordance with the University's IRB policies and procedures and 45 CFR 46, the Federal Policy for the

Protection of Human Subjects. Please reference the above-cited IRB Protocol 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 will be sent via mail. The IRB approval expiration date has been stamped on the informed consent document. Please note that research participants must sign a stamped version of the informed consent form and receive a copy.

Continuing Review: If you intend to continue to collect data from human subjects or to analyze private, identifiable data collected from human subjects, beyond the expiration date of this protocol, you must [submit a Renewal Application](#) to the IRB Office 45 days prior to the expiration date. If IRB Approval of your protocol expires, all human subject research activities including enrollment of new subjects, data collection and analysis of identifiable, private information must cease until the Renewal Application is approved. If work on the human subject portion of your project is complete and you wish to close the protocol, please [submit a Closure Report](#) to irb@umd.edu.

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 an apparent immediate hazard to the subjects. If you would like to modify an approved protocol, please [submit an Addendum request](#) to the IRB Office.

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

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

The UMCP IRB is organized and operated according to guidelines of the United States Office for Human Research Protections and the United States Code of Federal Regulations and operates under Federal Wide Assurance No. FWA00005856.

1204 Marie Mount Hall

College Park, MD 20742-5125

TEL 301.405.4212

FAX 301.314.1475

irb@umd.edu

<http://www.umresearch.umd.edu/IRB>

UNIVERSITY OF MARYLAND, COLLEGE PARK
Institutional Review Board
Initial Application for Research Involving Human Subjects

Last Rev. 05/12/2010



UNIVERSITY OF
MARYLAND

Institute for Child Study/Department of Human Development Melanie Killen, Ph.D.
3304 Benjamin Building Office: 301.405.3176
College Park, MD 20742-1131 E-mail: mkillen@umd.edu

Dear Parents or Guardians,

We are conducting a project on how children and adolescents make decisions about inclusion in peer groups. We would like to ask your permission to interview your son or daughter for this project. We are interested in studying how children and adolescents judge peer groups who include group members based on different reasons. We will tell participants short stories about afterschool clubs that include group members as well based on who they think should participate in group activities. These issues are central to how children and adolescents evaluate peer relationships and group identity.

Interviews will be administered by trained research assistants from the University of Maryland to students who provide assent to participate. The interview will take about 30 minutes to complete. Students will be asked to evaluate scenarios in which individuals are asked to make choices regarding peer inclusion. All information is confidential. **Please look over the description 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 the form to your child's teacher.**

The information obtained from this study will help teachers, policy makers, counselors and school administrators design curriculum and interventions to promote tolerance and mutual respect among children and adolescents. 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 daughter/son to participate. We have found that students enjoy the opportunity to express their opinions about their peer relations.

Thank you,

A handwritten signature in cursive script that reads "Melanie Killen".

Melanie Killen, Ph.D.
Professor of Human Development
Associate Director, Center for Children, Relationships, and Culture
Office: 301.405.3176 or email: mkillen@umd.edu

UNIVERSITY OF MARYLAND, COLLEGE PARK
 Institutional Review Board
 Initial Application for Research Involving Human Subjects

List 10/10/08, 9/1/2008

PARENTAL CONSENT FORM

Project Title	Gender Exclusion, Intentionality, and Theory of Social Mind	
Why is this research being done?	This is a research project being conducted by Dr. Melanie Killen at the University of Maryland, College Park. We are inviting you to participate in this research project because your child is 3-6 years old. The purpose of this research project is to better understand how children think about when it is okay or not okay to include or exclude peers in familiar school contexts.	
What will your child be asked to do?	Your child will be interviewed for 30 minutes. The interview will be conducted in the school in a specially designated area. Trained research assistants from the University of Maryland, College Park, will conduct the interview and will be available to answer any questions. Your child will be told hypothetical stories about children making inclusion and exclusion decisions regarding peer groups, and asked what s/he thinks about the various decisions.	
What about confidentiality?	We will do our best to keep your personal information confidential. To help protect confidentiality, your child's name will not be attached to the interview. S/he will be given an ID number. We will not share his/her answers with anyone, including his/her teachers, principal, or parents. If we write a report or article about this research project, your identity will be protected to the maximum extent possible. Your information may be shared with representatives of the University of Maryland, College Park or governmental authorities if you or someone else is in danger or if we are required to do so by law.	
What are the risks of this research?	There are no known risks associated with participating in this research project.	
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 friendship and peer inclusion and exclusion. 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 school districts.	
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 withdrawal 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 you allow your child to participate in this research project.	
Signature and Date	NAME OF CHILD	
	SIGNATURE OF PARENT	
	DATE	



UNIVERSITY OF MARYLAND, COLLEGE PARK
Institutional Review Board
Initial Application for Research Involving Human Subjects

IRB Form 1676a (2012)

PARENTAL CONSENT FORM

Project Title	Gender Exclusion, Intentionality, and Theory of Social Mind	
Why is this research being done?	This is a research project being conducted by Dr. Melanie Killen at the University of Maryland, College Park. We are inviting you to participate in this research project because your child is in 4 th grade. The purpose of this research project is to better understand how children think about when it is okay or not okay to include or exclude peers in familiar school contexts.	
What will your child be asked to do?	Your child will be interviewed for 30 minutes. The interview will be conducted in the school in a specially designated area. Trained research assistants from the University of Maryland, College Park, will conduct the interview and will be available to answer any questions. Your child will be told hypothetical stories about children and adolescents making inclusion and exclusion decisions regarding school funds for afterschool groups, and asked what s/he thinks about the various decisions.	
What about confidentiality?	We will do our best to keep your personal information confidential. To help protect confidentiality, your child's name will not be attached to the interview. S/he will be given an ID number. We will not share his/her answers with anyone, including his/her teachers, principal, or parents. If we write a report or article about this research project, your identity will be protected to the maximum extent possible. Your information may be shared with representatives of the University of Maryland, College Park or governmental authorities if you or someone else is in danger or if we are required to do so by law.	
What are the risks of this research?	There are no known risks associated with participating in this research project.	
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 friendship and peer inclusion and exclusion. The results will help us learn more about what kids and teenagers 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 school districts.	
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 withdrawal 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 you allow your child to participate in this research project.	
Signature and Date	NAME OF CHILD	
	SIGNATURE OF PARENT	
	DATE	



UNIVERSITY OF MARYLAND, COLLEGE PARK
Institutional Review Board
Initial Application for Research Involving Human Subjects

8th grade parent letter: Consent



UNIVERSITY OF
MARYLAND

Institute for Child Study/Department of Human Development Melanie Killen, Ph.D.
3304 Benjamin Building Office: 301.405.3176
College Park, MD 20742-1131 Email: mkillen@umd.edu

Dear Parents or Guardians,

We are conducting a project on how children and adolescents make decisions about inclusion in peer groups. We would like to ask your permission to survey your son or daughter for this project. We are interested in studying how children and adolescents judge peer groups who include group members based on different reasons. We will tell participants short stories about afterschool clubs that include group members as well based on who they think should participate in group activities. These issues are central to how children and adolescents evaluate peer relationships and group identity.

Surveys will be administered by trained research assistants from the University of Maryland to students who provide assent to participate. The survey will take about 30 minutes to complete. Students will be asked to evaluate scenarios in which individuals are asked to make choices regarding peer inclusion. All information is confidential. This research is voluntary. You have the ability to withdraw your child at any time. This is not a school or class requirement and will not affect your child's grade. **Please look over the description 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 the form to your child's teacher.**

The information obtained from this study will help teachers, policy makers, counselors and school administrators design curriculum and interventions to promote tolerance and mutual respect among children and adolescents. 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 daughter/son to participate. We have found that students enjoy the opportunity to express their opinions about their peer relations.

Thank you,

A handwritten signature in cursive script that reads "Melanie Killen".

Melanie Killen, Ph.D.
Professor of Human Development
Associate Director, Center for Children, Relationships, and Culture
Office: 301.405.3176 or email: mkillen@umd.edu

UNIVERSITY OF MARYLAND, COLLEGE PARK
 Institutional Review Board
Initial Application for Research Involving Human Subjects

PARENTAL CONSENT FORM

Project Title	Gender Exclusion, Intentionality, and Theory of Social Mind	
Why is this research being done?	This is a research project being conducted by Dr. Melanie Killen at the University of Maryland, College Park. We are inviting you to participate in this research project because your child is in 8 th grade. The purpose of this research project is to better understand how children think about when it is okay or not okay to include or exclude peers in familiar school contexts.	
What will your child be asked to do?	Your child will be surveyed for 30 minutes. The survey will be conducted in the school in a specially designated area. Trained research assistants from the University of Maryland, College Park, will conduct the survey and will be available to answer any questions. Your child will be told hypothetical stories about children and adolescents making inclusion and exclusion decisions regarding school funds for afterschool groups, and asked what s/he thinks about the various decisions.	
What about confidentiality?	We will do our best to keep your personal information confidential. To help protect confidentiality, your child's name will not be attached to the survey. S/he will be given an ID number. We will not share his/her answers with anyone, including his/her teachers, principal, or parents. If we write a report or article about this research project, your identity will be protected to the maximum extent possible. Your information may be shared with representatives of the University of Maryland, College Park or governmental authorities if you or someone else is in danger or if we are required to do so by law.	
What are the risks of this research?	There are no known risks associated with participating in this research project.	
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 friendship and peer inclusion and exclusion. The results will help us learn more about what kids and teenagers 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 school districts.	
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 withdrawal 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 you allow your child to participate in this research project.	
Signature and Date	NAME OF CHILD	
	SIGNATURE OF PARENT	
	DATE	

IRB APPROVED
 EXPIRES ON

JUN 02 2012

 UNIVERSITY OF MARYLAND
 COLLEGE PARK

UNIVERSITY OF MARYLAND, COLLEGE PARK
Institutional Review Board
Initial Application for Research Involving Human Subjects

IRB Form 10/12/10

8th grade parent letter: Assent



UNIVERSITY OF
MARYLAND

Institute for Child Study/Department of Human Development Melanie Killen, Ph.D.
3304 Benjamin Building Office: 301.405.3176
College Park, MD 20742-1131 Email: mkillen@umd.edu

Dear Parents or Guardians,

We are conducting a project on how children and adolescents make decisions about inclusion in peer groups. We would like to ask your permission to survey your son or daughter for this project. We are interested in studying how children and adolescents judge peer groups who include group members based on different reasons. We will tell participants short stories about afterschool clubs that include group members as well based on who they think should participate in group activities. These issues are central to how children and adolescents evaluate peer relationships and group identity.

Surveys will be administered by trained research assistants from the University of Maryland to students who provide assent to participate. The survey will take about 30 minutes to complete. Students will be asked to evaluate scenarios in which individuals are asked to make choices regarding peer inclusion. All information is confidential. This research is voluntary. You have the ability to withdraw your child at any time. This is not a school or class requirement and will not affect your child's grade.

The information obtained from this study will help teachers, policy makers, counselors and school administrators design curriculum and interventions to promote tolerance and mutual respect among children and adolescents. 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 daughter/son to participate. We have found that students enjoy the opportunity to express their opinions about their peer relations.

Thank you,

A handwritten signature in cursive script that reads "Melanie Killen".

Melanie Killen, Ph.D.
Professor of Human Development
Associate Director, Center for Children, Relationships, and Culture
Office: 301.405.3176 or email: mkillen@umd.edu

UNIVERSITY OF MARYLAND, COLLEGE PARK
Institutional Review Board
Initial Application for Research Involving Human Subjects

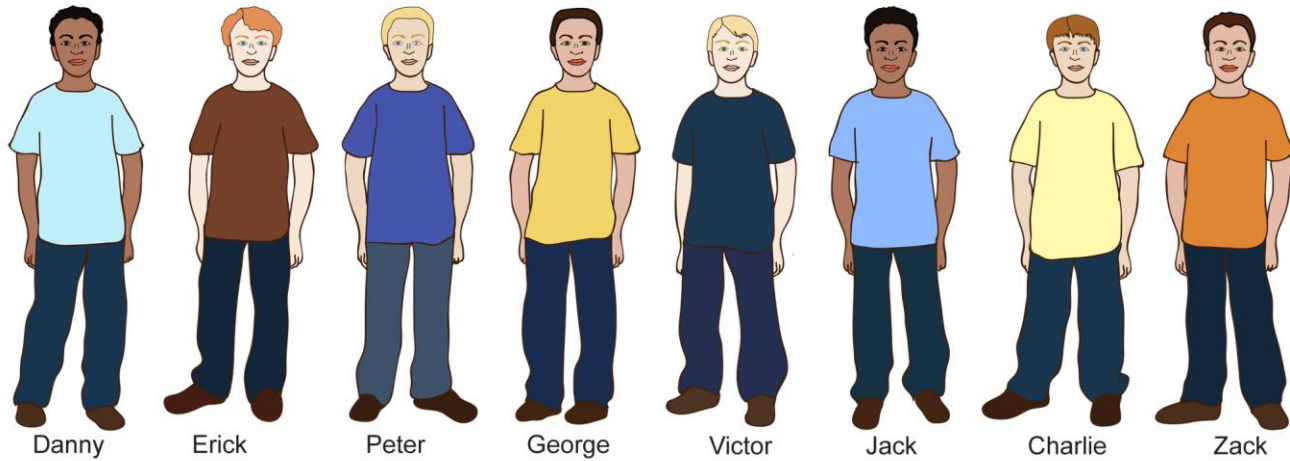
STUDENT ASSENT FORM

Project Title	Gender Exclusion, Intentionality, and Theory of Social Mind	
Why is this research being done?	This is a research project being conducted by Dr. Melanie Killen at the University of Maryland, College Park. We are inviting you to participate in this research project because you are in 8 th grade. We are completing this project to better understand how children and adolescents think about when it is okay or not okay to include or exclude kids your age in groups.	
What will your child be asked to do?	You will complete a survey for 30 minutes at school. Trained research assistants from the University of Maryland, College Park, will give the surveys and will be available to answer any questions. You will be told stories about children and adolescents making inclusion and exclusion decisions regarding afterschool groups, and asked what you think about what they decide.	
What about confidentiality?	We will do our best to keep your personal information private. Your name will not be attached to the interview. You will be given an ID number. 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 identity will be protected as much as possible. Your information may be shared with representatives of the University of Maryland, College Park or governmental authorities if you or someone else is in danger or if we are required to do so by law.	
What are the risks of this research?	There are no known risks to participating in this research project.	
What are the benefits of this research?	This research is not designed to help you personally, but to help us learn about friendship and peer inclusion and exclusion. The results will help us learn more about what kids and teenagers think about social relationships. We hope that in the future, other people will be able to learn from this research and use it to help kids and teenagers.	
Do I have to be in this research? May I stop participating at any time?	Your participation is strictly voluntary. You can ask any questions at any time. You 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 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 you allow You to participate in this research project.	
Signature and Date	NAME OF PARTICIPANT	
	SIGNATURE OF PARTICIPANT	
	DATE	

IRB APPROVED
EXPIRES ON

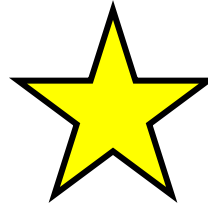
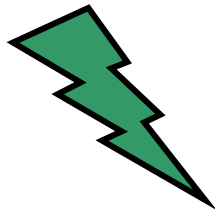
JUN 02 2012

UNIVERSITY OF MARYLAND
COLLEGE PARK

Appendix B: Sample Tasks**Gender Exclusion Task****Let's get started!****This group of boys is your group of friends:**

1 – What color do you want to be your group color? _____

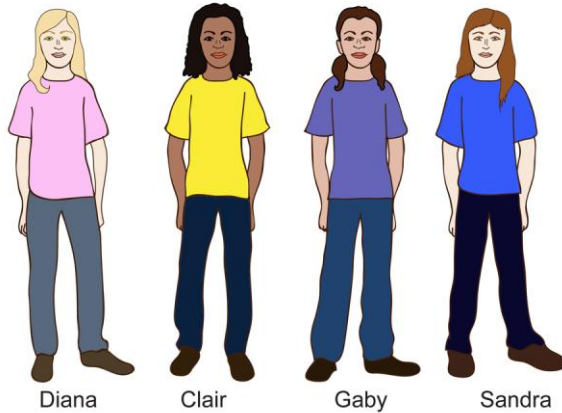
2 – Circle the symbol that you would like for your group:



INTRODUCTION: SC-C

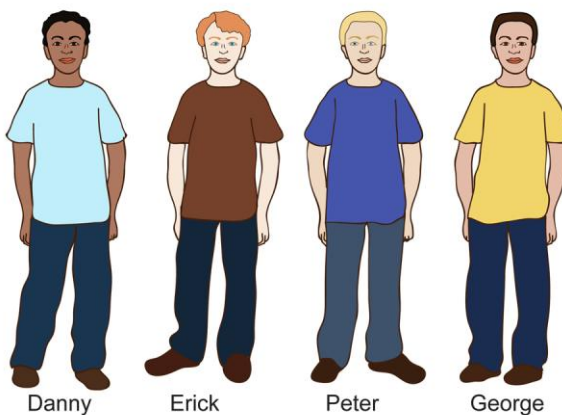
These groups of friends have to choose an afterschool activity and can pick between football or ballet.

Their group, the girls group,



Always likes to choose ballet because they say: “We like to do ballet, that’s for girls.”

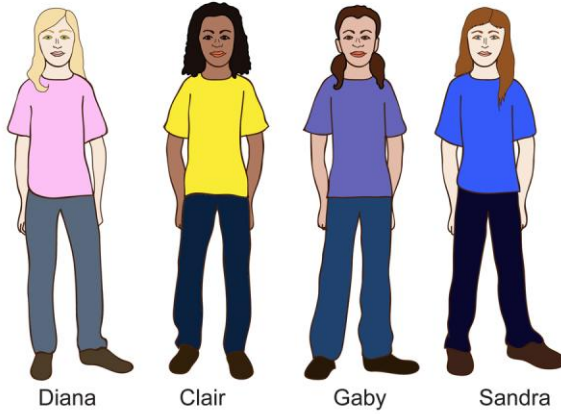
Your group, the boys group,



Always likes to choose football because they say: “We like to play football, that’s for boys.”

STORY 1

Now remember, their group, the girls group,



Always likes to choose ballet because they say: “We like to do ballet, that’s for girls.” Kay, who is also in this group, wants to be different from the other members of their group. She thinks “People think football is only for boys, let’s play football.”

Q0: What do **you** think **she** would do?

- Tell the group what she thinks
- Go along with the group and not tell them what she thinks

Q0B: What would **you** do?

- Tell the group what you think
- Go along with the group and not tell them what you think

Q1: Let’s say she tells **their group** what she thinks. How okay or not okay will **they** think what she says is?

1	2	3	4	5	6
Really Not Okay					Really Okay

Q1B: Why? (Please fill out the lines with their answer.)

Q2: Now, what about **you**? When **you** hear her, how okay or not okay do **you** think what she says is?

1	2	3	4	5	6
Really Not Okay					Really Okay

Q2B: Why? (Please fill out the lines with their answer.)

Remember, Kay says: "People think football is only for boys, let's play football." The group has to decide what to do.

Q3: Do you think **the girls group** will tell her she can't be in the group anymore?

	NO	YES			
How much?					
1	2	3	4	5	6
Can Really Stay				Really Cannot Stay	

Q3B: Why? (Please fill out the lines with their answer.)

Some new kids want to join the group. There is only room for one more member. They have to choose who to invite to join. Remember, their group says: "We like to do ballet, that's for girls."

Q1: Who should their group invite:

Karen, who wants to be in this group and would say "I like to play football."

OR

Donald, who wants to be in this group and would say "I like to do ballet."

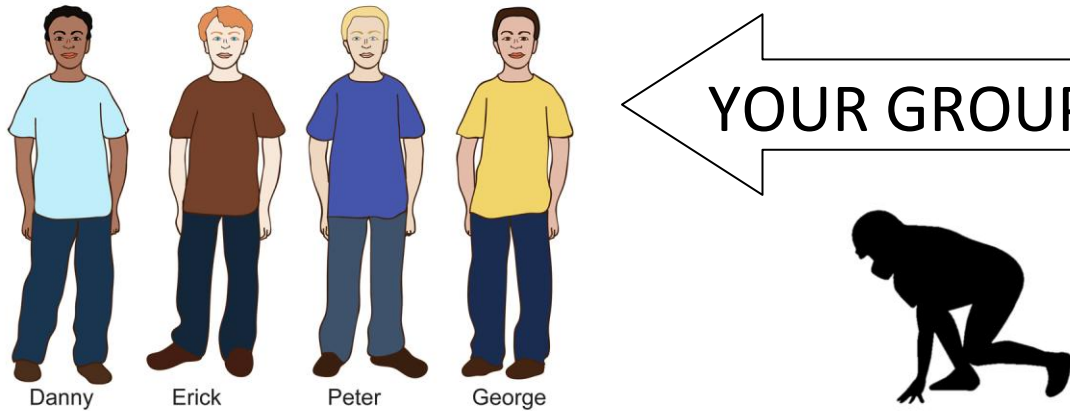
KAREN

DONALD

Q2: Why? (Please fill out the lines with their answer.)

STORY 2

Your group, the boys group,



Always likes to choose football because they say: "We like to play football, that's for boys."

Marcus, who is also in this group, wants to be different from the other members of his group. He thinks "People think ballet is only for girls, let's do ballet."

Q0: What do **you** think **he** would do?

- Tell the group what he thinks
- Go along with the group and not tell them what he thinks

Q0B: What would **you** do?

- Tell the group what you think
- Go along with the group and not tell them what you think

Q1: Let's say he tells **your group** what he thinks. How okay or not okay will **they** think what he says is?

1	2	3	4	5	6
Really Not Okay					Really Okay

Q1B: Why? (Please fill out the lines with their answer.)

Q2: Now, what about **you**? When **you** hear him, how okay or not okay do **you** think what he says is?

1	2	3	4	5	6
Really Not Okay					Really Okay

Q2B: Why? (Please fill out the lines with their answer.)

Remember, Marcus says: "People think ballet is only for girls, let's do ballet."
The group has to decide what to do.

Q3: Do you think **the boys group** will tell him he can't be in your group anymore?

NO YES

How much?

1	2	3	4	5	6
Can Really Stay					Really Cannot Stay

Q3B: Why? (Please fill out the lines with their answer.)

Some new kids want to join the group. There is only room for one more member. They have to choose who to invite to join. Remember, your group says: "We like to play football, that's for boys."

Q1: Who should your group invite:

Frank who wants to be in the group and would say "I like to do ballet."

OR

Sally who wants to be in the group and would say "I like to play football."

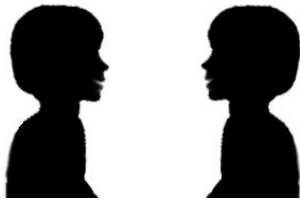
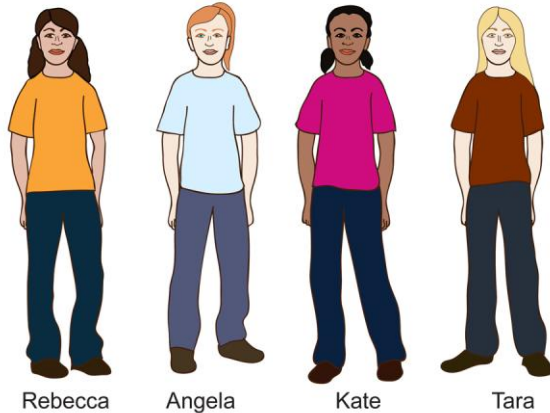
FRANK SALLY

Q2: Why? (Please fill out the lines with their answer.)

INTRODUCTION: M-R

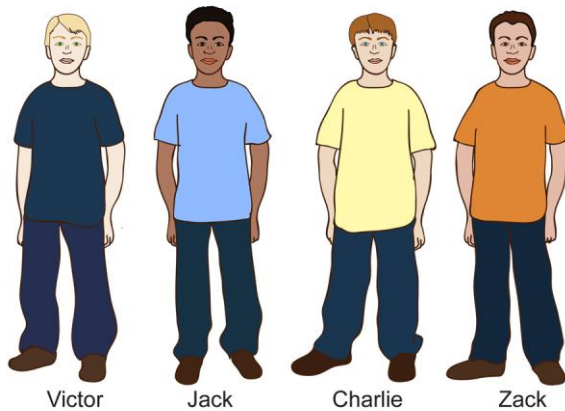
Let's say there are some kids at school who always sit alone, and act differently from the other kids.

Their group, the girls group,



Says: "Even though you don't know why those kids act that way, you shouldn't always gossip about those kids."

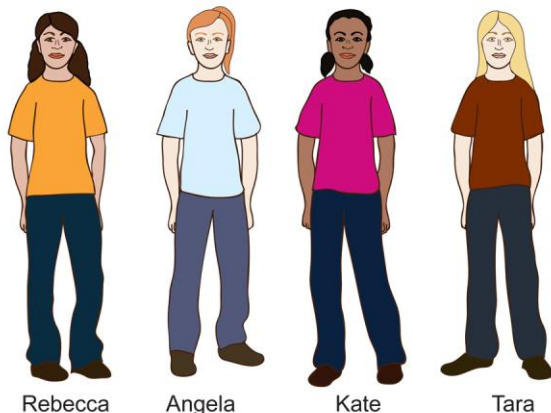
Your group, the boys group,



Says: "It's okay to gossip because you don't know why those kids act that way."

STORY 3

Now remember, their group, the girls group,



Says “Even though you don’t know why those kids act that way, you shouldn’t always gossip about those kids.”

Betsy, who is also in this group, wants to be different from the other members of their group. She thinks “It’s okay to gossip all the time, because we don’t know why those kids act that way.”

Q0: What do **you** think **she** would do?

- Tell the group what she thinks
- Go along with the group and not tell them what she thinks

Q0B: What would **you** do?

- Tell the group what you think
- Go along with the group and not tell them what you think

Q1: Let’s say she tells **their group** what she thinks. How okay or not okay will **they** think what she says is?

1	2	3	4	5	6
Really Not Okay					Really Okay

Q1B: Why? (Please fill out the lines with their answer.)

Q2: Now, what about **you**? When **you** hear her, how okay or not okay do **you** think what she says is?

1	2	3	4	5	6
					Really Okay
Really Not Okay					

Q2B: Why? (Please fill out the lines with their answer.)

Remember, Betsy says: "It's okay to gossip all the time, because we don't know why those kids act that way." The group has to decide what to do.

Q3: Do you think **the girls group** will tell her she can't be in their group anymore?

NO YES

How much?

1	2	3	4	5	6
					Really Cannot Stay
Can Really Stay					

Q3B: Why? (Please fill out the lines with their answer.)

Some new kids want to join the group. There is only room for one more member. They have to choose who to invite to join. Remember, their group: "Even though you don't know why those kids act that way, you shouldn't always gossip about those kids."

Q1: Who should their group invite:

Katelyn, who wants to be in this group and would say "It's okay to gossip all the time, because we don't know why those kids act that way."

OR

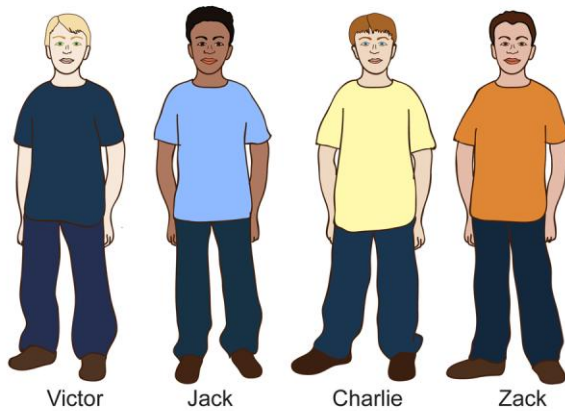
David, who wants to be in this group and would say "Even though you don't know why those kids act that way, you shouldn't always gossip about those kids."

KATELYN DAVID

Q2: Why? (Please fill out the lines with their answer.)

STORY 4

Remember, your group, the boys group,



Victor

Jack

Charlie

Zack



Says: "It's okay to gossip all the time, because we don't know why those kids act that way."

Jacob, who is also in this group, wants to be different from the other members of his group. He thinks "Even though you don't know why those kids act that way, you shouldn't always gossip about those kids."

Q0: What do **you** think **he** would do?

Tell the group what he thinks

Go along with the group and not tell them what he thinks

Q0B: What would **you** do?

Tell the group what you think

Go along with the group and not tell them what you think

Q1: Let's say he tells **your group** what he thinks. How okay or not okay will **they** think what he says is?

1

2

3

4

5

6

Really Not Okay

Really Okay

Q1B: Why? (Please fill out the lines with their answer.)

Q2: Now, what about **you**? When **you** hear him, how okay or not okay do **you** think what he says is?

1	2	3	4	5	6
Really Not Okay					Really Okay

Q2B: Why? (Please fill out the lines with their answer.)

Remember, Jacob says: “Even though you don’t know why those kids act that way, you shouldn’t always gossip about those kids.” The group has to decide what to do.

Q3: Do you think **the boys group** will tell him he can’t be in your group anymore?

	NO	YES				
How much?						
	1	2	3	4	5	6
	Can Really Stay					Really Cannot Stay

Q3B: Why? (Please fill out the lines with their answer.)

Some new kids want to join the group. There is only room for one more member. They have to choose who to invite to join. Remember, your group that says: “It’s okay to gossip all the time, because we don’t know why those kids act that way.”

Q1: Who should your group invite:

Arthur who wants to be in the group and would say “Even though you don’t know why those kids act that way, you shouldn’t always gossip about those kids.”

OR

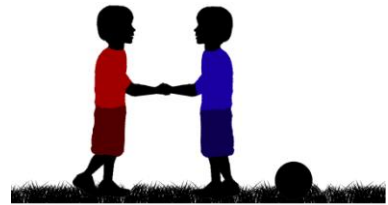
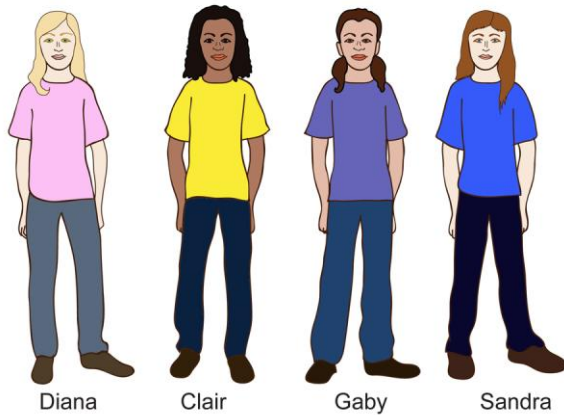
Emily who wants to be in the group and would say “It’s okay to gossip all the time, because we don’t know why those kids act that way.”

ARTHUR EMILY

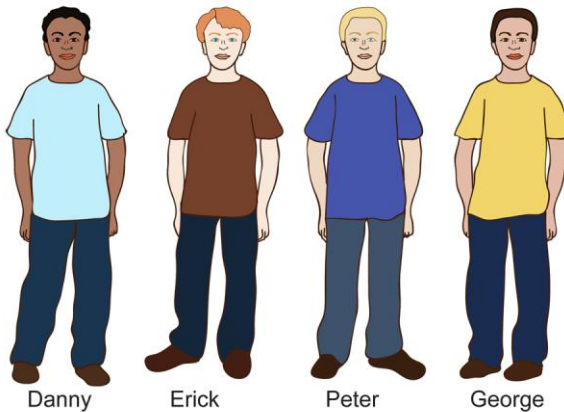
Q2: Why? (Please fill out the lines with their answer.)

INTRODUCTION: M-C

These are groups of friends at your school.

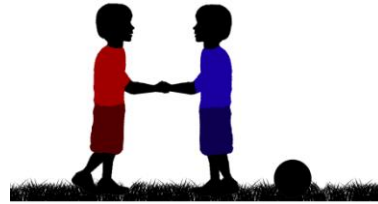
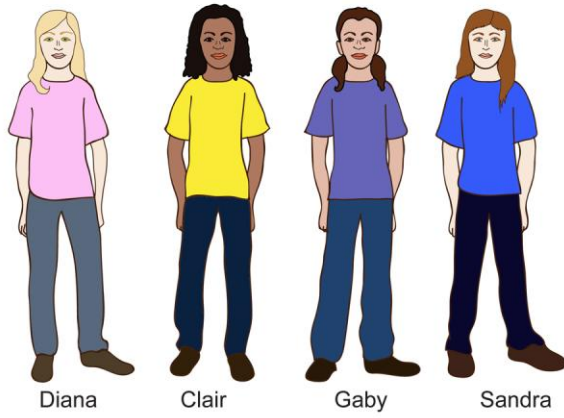


When playing soccer, their group, the girls group, says “Even though you want to score points, you shouldn’t always push and shove.”



When playing soccer, your group, the boys group, says “It’s okay to push and shove just to score points.”

STORY 5



Now remember, their group, the girls group, says “Even though you want to score points, you shouldn’t always push and shove.”

Stephanie, who is also in this group, wants to be different from the other members of their group. She thinks “It’s okay to push and shove just to score points.”

Q0: What do **you** think **she** would do?

- Tell the group what she thinks
- Go along with the group and not tell them what she thinks

Q0B: What would **you** do?

- Tell the group what you think
- Go along with the group and not tell them what you think

Q1: Let’s say she tells **their group** what she thinks. How okay or not okay will **they** think what she says is?

- | | | | | | |
|-----------------|---|---|---|---|-------------|
| 1 | 2 | 3 | 4 | 5 | 6 |
| Really Not Okay | | | | | Really Okay |

Q1B: Why? (Please fill out the lines with their answer.)

Q2: Now, what about **you**? When **you** hear her, how okay or not okay do **you** think what she says is?

1	2	3	4	5	6
Really Not Okay					Really Okay

Q2B: Why? (Please fill out the lines with their answer.)

Remember, Stephanie says: "It's okay to push and shove just to score points." The group has to decide what to do.

Q3: Do you think **the girls group** will tell her she can't be in their group anymore?

NO YES

How much?

1	2	3	4	5	6
Can Really Stay					Really Cannot Stay

Q3B: Why? (Please fill out the lines with their answer.)

Some new kids want to join the group. There is only room for one more member. They have to choose who to invite to join. Remember, their group says: "Even though you want to score points, you shouldn't always push and shove."

Q1: Who should their group invite:

Alice who wants to be in the group and would say "It's okay to push and shove just to score points."

OR

Gary who wants to be in the group and would say "Even though you want to score points, you shouldn't always push and shove."?

ALICE GARY

Q2: Why? (Please fill out the lines with their answer.)

Q2: Now, what about **you**? When **you** hear him, how okay or not okay do **you** think what he says is?

1	2	3	4	5	6
Really Not Okay				Really Okay	

Q2B: Why? (Please fill out the lines with their answer.)

Remember, Michael says: “Even though you want to score points, you shouldn’t always push and shove.”

The group has to decide what to do.

Q3: Do you think **the boys group** will tell him he can’t be in your group anymore?

	NO	YES	
How much?			
1	2	3	4
Can Really Stay			Really Cannot
	Stay		

Q3B: Why? (Please fill out the lines with their answer.)

Some new kids want to join the group. There is only room for one more member. They have to choose who to invite to join. Remember, your group says: “It’s okay to push and shove just to score points.”

Q1: Who should your group invite:

Molly, who wants to be in this group and would say “It’s okay to push and shove just to score points.”

OR

Dan, who wants to be in this group and would say “Even though you want to score points, you shouldn’t always push and shove.”

MOLLY DAN

Q2: Why? (Please fill out the lines with their answer.)

Interpretative Theory of Mind Task

Story 7:

John and Allan are playing a game. They are supposed to “wait for a ring” before they take the next turn.



John says they should wait for the telephone to ring.



Allan says they should wait for a ring that you wear.

Q1. Is it okay for John to say they should wait for the telephone to ring and Allan to say they should wait for a ring that you wear?

YES NO

Q2: Why? (Please fill out the lines with their answer.)

Q3: Now a group of boys comes over and hears about the game. What will they say they should wait for: a telephone to ring, a ring to wear, or would you not know what they will say?

TELEPHONE RING RING TO WEAR UNSURE

Q4: Why? (Please fill out the lines with their answer.)

Story 8:

John and Allan see this picture.

John says it is a duck. Allan says it is a rabbit.

Q1. Is it okay for John to say it is a duck and Allan to say it is a rabbit?

YES NO

Q2: Why? (Please fill out the lines with their answer.)

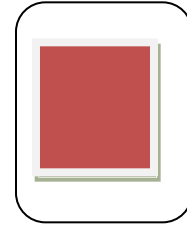
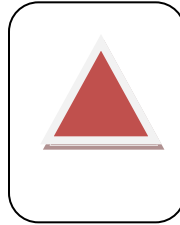
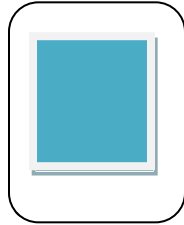
Q3: Now a group of boys comes over and sees the picture. What will they say it is, a duck, a rabbit or would you not know what they will say?

DUCK RABBIT UNSURE

Q4: Why? (Please fill out the lines with their answer.)

Story 9:

John and Allan need to find a penny, which is hidden under one of these three cards:



The penny is under the card with the block on it.

John says it is under card 1. Allan says it is under card 3.

Q1. Is it okay for John to say it is under Card 1 and Allan to say it is under Card 3?

YES NO

Q2: Why? (Please fill out the lines with their answer.)

Q3: Now a group of boys comes over and sees the picture. Where will they say it is, under Card 1, Card 3 or would you not know what they would say?

CARD 1 CARD 3 UNSURE

Q4: Why? (Please fill out the lines with their answer.)

Gender Stereotype Task

- Q1. Who usually plays football ?
- | | | | |
|--|------|-------|------|
| | BOYS | GIRLS | BOTH |
|--|------|-------|------|
- Q2. Do you play football ?
- | | | |
|--|-----|----|
| | YES | NO |
|--|-----|----|
- Q3. How many of your friends play football?
- | | | | |
|------|-------|------|------|
| NONE | A FEW | SOME | MOST |
|------|-------|------|------|
- Q4. Who usually does ballet?
- | | | | |
|--|------|-------|------|
| | BOYS | GIRLS | BOTH |
|--|------|-------|------|
- Q5. Do you do ballet?
- | | | |
|--|-----|----|
| | YES | NO |
|--|-----|----|
- Q6. How many of your friends do ballet?
- | | | | |
|------|-------|------|------|
| NONE | A FEW | SOME | MOST |
|------|-------|------|------|
- Q7. Who usually gossips?
- | | | | |
|--|------|-------|------|
| | BOYS | GIRLS | BOTH |
|--|------|-------|------|
- Q8. How many of your friends gossip?
- | | | | |
|------|-------|------|------|
| NONE | A FEW | SOME | MOST |
|------|-------|------|------|
- Q9. Who usually avoids gossip?
- | | | | |
|--|------|-------|------|
| | BOYS | GIRLS | BOTH |
|--|------|-------|------|
- Q10. How many of your friends avoid gossip?
- | | | | |
|------|-------|------|------|
| NONE | A FEW | SOME | MOST |
|------|-------|------|------|
- Q11. Who usually pushes and shoves?
- | | | | |
|--|------|-------|------|
| | BOYS | GIRLS | BOTH |
|--|------|-------|------|
- Q13. How many of your friends push and shove?
- | | | | |
|------|-------|------|------|
| NONE | A FEW | SOME | MOST |
|------|-------|------|------|
- Q14. Who usually plays nice?
- | | | | |
|--|------|-------|------|
| | BOYS | GIRLS | BOTH |
|--|------|-------|------|
- Q15. How many of your friends play nice?
- | | | | |
|------|-------|------|------|
| NONE | A FEW | SOME | MOST |
|------|-------|------|------|
- Q16. How many of your friends are boys?
- | | | | |
|------|-------|------|------|
| NONE | A FEW | SOME | MOST |
|------|-------|------|------|

Appendix C: Coding System

Justification Categories for Coding Surveys: Gender and Aggression Study

Below are the justification codes to use for the Transcribed Interview Protocol for each participant. These codes are recorded on the accompanying Justification Coding Sheet. There are four groups of categories: *Moral*, *Social-Conventional*, *Psychological*, and *Undifferentiated*. Each category has subcategories. Definitions and Ex:s are below.

Moral: Justification codes 1 – 2 are referred to as “Moral” because justice, fairness or rights of a victim are involved. Includes all positive and negative references to the moral domain.

Social-Conventional: Justification codes 3-6 are “Social-conventional” because group functioning, group identity, rules and authority form the basis for the response. Includes all positive and negative references to group functioning.

Psychological Justifications 7-9 are “Psychological” because they involve focus on individual concerns.

Undifferentiated: Category 10 is undifferentiated (“It’s bad”; “It’s good; He’s weird.”). Justification 11 is for “Other” responses that do not fit in any other category (keep notes on “Other”).

Missing or Uncodable: Justification 99 is for missing data or uncodable.

CODING DECISIONS: You may use two codes if the response warrants two codes. If more than two are indicated, choose the two most developed codes/reasoning.

CATEGORIES

I. Moral Domain

1. Concern for Other’s Welfare

A. Psychological Harm: References negative intentions towards others (teasing or being mean); acknowledgment of how it feels to be excluded.

EX: It will hurt his feelings.

EX: Gossip always makes someone feel bad.

EX: She’s jealous.

EX: Gossip is mean.

EX: If he kicked me out I would be upset.

EX: How would they feel if it happened to them.

B. Physical Harm: References physical harm to another.

EX: She’ll hurt someone else if she pushes.

EX: If you push and shove, someone might get hurt.

EX: It’s mean/bad to push or shove.

EX: It’s nicer to not hit.

C. Fairness/Equity/Rights: Appeals to principles regarding fairness, equity, and rights.

EX: That wouldn’t be fair to kick him out.

EX: It’s never right to kick someone out of a group.

EX: It’s not fair to talk about someone behind their back.

EX: It’s the right thing to do.

II. Societal Domain

2. Group Functioning and Conventions: Conventions of the group designed to promote the group/encourage group functioning. Recognizing that the act does not disrupt the group.
 - EX: It's just a game.
 - EX: Doesn't hurt the group.
 - EX: He's not doing what they want him to do. They outnumber him.
 - EX: They all want to play one way. He needs to go along.
 - EX: You need to follow the rules of the club.
 - EX: It's good for the group.
 - EX: He was only thinking about the group
 - EX: She fits in. (*She is the child matching the norm and NOT gender*)
 - EX: Then you know who's in the group
 - EX: He's going against the group.
 - EX: He didn't do what the group said to.
 - EX: You need to follow the rules. (*If group norm matches social norm*)

3. Inclusion of Diverse Perspectives
 - EX: She will change their minds. (*Unless references fairness*)
 - EX: He can teach the group.
 - EX: She'll be able to support the deviant member.

4. Gender Group Identity/Stereotypes: Appeals to group identity as boys or girls.
 - A. General/Non-stereotypic:
 - EX: She's a girl!
 - EX: He fits in. (*He is the child matching the gender of group and NOT the Norm*).
 - EX: She would be out of place (*She is the child matching the gender of group and NOT the norm*).
 - B. Adhering to Stereotypes: Appeals to gender stereotypes.
 - EX: He's acting gay.
 - EX: That's such a girly thing to do.
 - EX: Well, girls are supposed to play nicely.
 - EX: Football is only for boys.
 - EX: Girls always gossip, so they'd like her.

III. Psychological Domain

5. Autonomy: Individuality and personal choice.
 - EX: It's good to be different.
 - EX: She's just trying to be different.
 - EX: She's being honest/telling the truth (if about football/ballet).

6. Personal Identification with the target:
 - A. Moral Domain: involves moral issues such as psychological or physical harm.

EX: Well, I never gossip about others.

B. Societal Domain: involves conventions or activities.

EX: I prefer football, too.

IV. Other

7. Undifferentiated

EX: It's good.

EX: He's nice

EX: You can't kick her out. (*Use only if no further reasoning is*

given)

8. Other

99. Uncodable

References

- Abrams, D., De Moura, G. R., Marques, J. M., & Hutchison, P. (2008). Innovation credit: When can leaders oppose their group's norms? *Journal of Personality and Social Psychology, 95*, 662-678. doi: 10.1037/0022-3514.95.3.662
- Abrams, D., & Rutland, A. (2008). The development of subjective group dynamics. In S. R. Levy & M. Killen (Eds.), *Intergroup relations and attitudes in childhood through adulthood* (pp. 47-65). Oxford, UK: Oxford University Press.
- Abrams, D., Rutland, A., & Cameron, L. (2003). The development of subjective group dynamics: Children's judgments of normative and deviant in-group and out-group individuals. *Child Development, 74*, 1840-1856.
- Abrams, D., Rutland, A., Cameron, L., & Ferrell, J. (2007). Older but wiler: In-group accountability and the development of subjective group dynamics. *Developmental Psychology, 43*, 134-148.
- Abrams, D., Rutland, A., Cameron, L., & Marques, J. M. (2003). The development of subjective group dynamics: When in-group bias gets specific. *British Journal of Developmental Psychology, 21*, 155-177.
- Abrams, D., Rutland, A., Ferrell, J. M., & Pelletier, J. (2008). Children's judgments of disloyal and immoral peer behavior: Subjective group dynamics in minimal intergroup contexts. *Child Development, 79*, 444-461.
- Abrams, D., Rutland, A., Pelletier, J., & Ferrell, J. M. (2009). Children's group nous: Understanding and applying peer exclusion within and between groups. *Child Development, 80*, 224-243.

- Almås, I., Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science*, 328, 1176-1178. doi: 10.1126/science.1187300
- 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 perspective. In S. Levy & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 66-86). Oxford, U.K.: Oxford University Press.
- Astington, J. W. (2001). The future of theory-of-mind research: Understanding motivational states, the role of language, and real-world consequences. Commentary on 'Meta-analysis of theory-of-mind development: The truth about false belief.'. *Child Development*, 72, 685-687. doi: 10.1111/1467-8624.00305
- Atkin, L. M., & Gummerum, M. (2012). Moral and group-based criteria in children's evaluation of peers after transgressions. *Infant and Child Development*, 21, 189-197. doi: 10.1002/icd.744
- Badenes, L. V., Estevan, R. A. C., & Bacete, F. J. G. (2000). Theory of mind and peer rejection at school. *Social Development*, 9, 271-283. doi: 10.1111/1467-9507.00125
- Barchia, K., & Bussey, K. (2011). Predictors of student defenders of peer aggression victims: Empathy and social cognitive factors. *International Journal of Behavioral Development*, 35, 289-297. doi: 10.1177/0165025410396746
- Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, 66, 5-20.

- Blakemore, J. E. O. (2003). Children's beliefs about violating gender norms: Boys shouldn't look like girls, and girls shouldn't act like boys. *Sex Roles, 48*, 411-419. doi: 10.1023/a:1023574427720
- Brame, B., Nagin, D. S., & Tremblay, R. E. (2001). Developmental trajectories of physical aggression from school entry to late adolescence. *Journal of Child Psychology and Psychiatry, 42*, 503-512. doi: 10.1111/1469-7610.00744
- Brechwald, W. A., & Prinstein, M. J. (2011). Beyond homophily: A decade of advances in understanding peer influence processes. *Journal of Research on Adolescence, 21*, 166-179. doi: 10.1111/j.1532-7795.2010.00721.x
- Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, B., Dodge, K. A., Fergusson, D., Horwood, J. L., Loeber, R., Laird, R., Lynam, D. R., Moffitt, T. E., Pettit, G. S., & Vitaro, F. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. *Developmental Psychology, 39*, 222-245. doi: 10.1037/0012-1649.39.2.222
- Brown, B. B. (1990). Peer groups and peer cultures. . In S. S. Feldman & G. R. Elliot (Eds.), *At the threshold: The developing adolescent* (pp. 171-196). Cambridge, MA: Harvard University Press.
- Brown, C. S., Bigler, R. S., & Chu, H. (2010). An experimental study of the correlates and consequences of perceiving oneself to be the target of gender discrimination. *Journal of Experimental Child Psychology, 107*, 100-117. doi: 10.1016/j.jecp.2010.04.010

- Buhs, E. S., & Ladd, G. W. (2001). Peer rejection as antecedent of young children's school adjustment: An examination of mediating processes. *Developmental Psychology, 37*, 550-560.
- Buhs, E. S., Ladd, G. W., & Herald, S. L. (2006). Peer exclusion and victimization: Processes that mediate the relation between peer group rejection and children's classroom engagement and achievement? *Journal of Educational Psychology, 98*, 1-13. doi: 10.1037/0022-0663.98.1.1
- Buhs, E. S., McGinley, M., & Toland, M. D. (2010). Overt and relational victimization in Latinos and European Americans: Measurement equivalence across ethnicity, gender, and grade level in early adolescent groups. *The Journal of Early Adolescence, 30*, 171-197. doi: 10.1177/0272431609350923
- Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child Development, 79*, 1185-1229. doi: 10.1111/j.1467-8624.2008.01184.x
- Carpendale, J., & Lewis, C. (2006). *How children develop social understanding*. Oxford, U.K.: Blackwell Publishing.
- Carpendale, J. I., & Chandler, M. J. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development, 67*, 1686-1706. doi: 10.2307/1131725
- Chandler, M. J., & Helm, D. (1984). Developmental changes in the contribution of shared experience to social role-taking competence. *International Journal of Behavioral Development, 7*, 145-156. doi: 10.1016/s0163-6383(84)80207-6

- Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms on reactive and proactive aggression. *Child Development, 67*, 993-1002. doi: 10.2307/1131875
- Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development, 66*, 710-722. doi: 10.2307/1131945
- Crick, N. R., Ostrov, J. M., & Werner, N. E. (2006). A Longitudinal Study of Relational Aggression, Physical Aggression, and Children's Social-Psychological Adjustment. *Journal of Abnormal Child Psychology: An official publication of the International Society for Research in Child and Adolescent Psychopathology, 34*, 131-142. doi: 10.1007/s10802-005-9009-4
- DeWall, C. N., Twenge, J. M., Gitter, S. A., & Baumeister, R. F. (2009). It's the thought that counts: The role of hostile cognition in shaping aggressive responses to social exclusion. *Journal of Personality and Social Psychology, 96*, 45-59. doi: 10.1037/a0013196
- Dodge, K. A., & Coie, J. D. (1987). Social information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology, 53*, 1146-1158.
- Downey, G., Lebolt, A., Rincón, C., & Freitas, A. L. (1998). Rejection sensitivity and children's interpersonal difficulties. *Child Development, 69*, 1074-1091. doi: 10.2307/1132363
- Eccles, J. S., Roeser, R., Vida, M., Fredricks, J., & Wigfield, A. (2006). Motivational and Achievement Pathways Through Middle Childhood. In L. Balter & C. S. Tamis-

- LeMonda (Eds.), *Child psychology: A handbook of contemporary issues (2nd ed.)*. (pp. 325-355). New York, NY US: Psychology Press.
- Estell, D. B., Cairns, R. B., Farmer, T. W., & Cairns, B. D. (2002). Aggression in inner-city early elementary classrooms: Individual and peer-group configurations. *Merrill-Palmer Quarterly*, *48*, 52-76. doi: 10.1353/mpq.2002.0002
- Goble, P., Martin, C. L., Hanish, L. D., & Fabes, R. A. (2012). Children's gender-typed activity choices across preschool social contexts. *Sex Roles*, *67*, 435-451. doi: 10.1007/s11199-012-0176-9
- Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., Nugent, T. F., Herman, D. H., Clasen, L. S., Toga, A. W., Rapoport, J. L., & Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 8174-8179. doi: 10.1073/pnas.0402680101
- Goldstein, S. E., & Tisak, M. S. (2010). Adolescents' social reasoning about relational aggression. *Journal of Child and Family Studies*, *19*, 471-482. doi: 10.1007/s10826-009-9319-1
- Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA US: The MIT Press.
- Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture*. (pp. 257-293). New York, NY US: Cambridge University Press.

- Greenberg, M. T., & Harris, A. R. (2012). Nurturing mindfulness in children and youth: Current state of research. *Child Development Perspectives*, 6, 161-166. doi: 10.1111/j.1750-8606.2011.00215.x
- Hartup, W. W. (2005). The Development of Aggression: Where Do We Stand? In R. E. Tremblay, W. W. Hartup & J. Archer (Eds.), *Developmental origins of aggression*. (pp. 3-22). New York, NY US: Guilford Press.
- Hitti, A., Mulvey, K. L., & Killen, M. (2011). Social exclusion and culture: The role of group norms, group identity and fairness *Anales de Psicologia*.
- Horn, S. (2003). Adolescents' reasoning about exclusion from social groups. *Developmental Psychology*, 39, 71-84.
- Horn, S. S. (2007). Adolescents' acceptance of same-sex peers based on sexual orientation and gender expression. *Journal of Youth and Adolescence*, 36, 363-371. doi: 10.1007/s10964-006-9111-0
- Horn, S. S. (2008). The multifaceted nature of sexual prejudice: How adolescents reason about sexual orientation and sexual prejudice. In S. Levy & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 173-190). Oxford, U.K.: Oxford University Press.
- Jastrow, J. (1899). The mind's eye. *Popular Science Monthly*, 54, 299-312.
- Juvonen, J., & Galván, A. (2008). Peer influence in involuntary social groups: Lessons from research on bullying. In M. J. Prinstein & K. A. Dodge (Eds.), *Understanding peer influence in children and adolescents*. (pp. 225-244). New York, NY US: Guilford Press.

- Kelly, M. C., Mulvey, K. L., Hitti, A., & Killen, M. (May 2010). *Attribution of intentions: Theory of mind, morality and stereotypes*. Paper presented at the Association for Psychological Science, Boston, MA.
- Kelly, M. C., Mulvey, K. L., Hitti, A., & Moin, M. (2011). *Attribution of intentions: The role of stereotype knowledge and Theory of Mind in moral judgments*. Paper presented at the biennial meeting of the Society for Research in Child Development, Montreal, Canada.
- Kelly, M. C., Mulvey, K. L., Hitti, A., Moin, M., & Killen, M. (May 2011). *The impact of gender stereotypes and false belief on moral judgments*. Paper presented at the Association for Psychological Science Washington, D.C. .
- Killen, M. (2007). Children's social and moral reasoning about exclusion. *Current Directions in Psychological Science*, 16, 32-36.
- 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., Mulvey, K. L., & Hitti, A. (2012). Social exclusion in childhood: A developmental intergroup perspective. *Child Development*. doi: 10.1111/cdev.12012

- Killen, M., Mulvey, K. L., Richardson, C., Jampol, N., & Woodward, A. (2011). The accidental transgressor: Morally-relevant theory of mind. *Cognition, 119*, 197-215. doi: 10.1016/j.cognition.2011.01.006
- 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., & Rutland, A. (2011). *Children and social exclusion: Morality, prejudice, and group identity*. New York, NY: Wiley/Blackwell Publishers.
- Killen, M., Rutland, A., Abrams, D., Mulvey, K. L., & Hitti, A. (2012). Development of intra- and intergroup judgments in the context of moral and social-conventional norms. *Child Development*. doi: 10.1111/cdev.12011
- 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: Elsevier.
- 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.
- Kochel, K. P., Miller, C. F., Updegraff, K. A., Ladd, G. W., & Kochenderfer-Ladd, B. (2012). Associations between fifth graders' gender atypical problem behavior and peer relationships: A short-term longitudinal study. *Journal of Youth and Adolescence, 41*, 1022-1034. doi: 10.1007/s10964-011-9733-8

- Komolova, M., & Wainryb, C. (2011). 'What I want and what you want': Children's thinking about competing personal preferences. *Social Development, 20*, 334-352. doi: 10.1111/j.1467-9507.2010.00589.x
- Lansford, J. E., Skinner, A. T., Sorbring, E., Giunta, L. D., Deater-Deckard, K., Dodge, K. A., Malone, P. S., Oburu, P., Pastorelli, C., Tapanya, S., Uribe Tirado, L. M., Zelli, A., Al-Hassan, S. M., Peña Alampay, L., Bacchini, D., Bombi, A. S., Bornstein, M. H., & Chang, L. (2012). Boys' and Girls' Relational and Physical Aggression in Nine Countries. *Aggressive Behavior, 38*, 298-308. doi: 10.1002/ab.21433
- Latane, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York: Appleton Crofts Century Publishing.
- Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences, 8*, 529-533. doi: 10.1016/j.tics.2004.10.001
- 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, 67*.
- Maccoby, E. E. (2002). Gender and group process: A developmental perspective. *Current Directions in Psychological Science, 11*, 54-58. doi: 10.1111/1467-8721.00167
- Margie, N. G., Killen, M., Sinno, S., & McGlothlin, H. (2005). Minority children's intergroup attitudes about peer relationships. *British Journal of Developmental Psychology, 23*, 251-259.
- Martin, C. L., & Fabes, R. A. (2001). The stability and consequences of young children's same-sex peer interactions. *Developmental Psychology, 37*, 431-446.

- Martin, C. L., & Halverson, C. F. (1981). A schematic processing model of sex typing and sex stereotyping in children. *Child Development, 52*, 1119-1134.
- Moller, L. C., Hymel, S., & Rubin, K. H. (1992). Sex typing in play and popularity in middle childhood. *Sex Roles, 26*, 331-353. doi: 10.1007/bf00289916
- Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The rejection of moral rebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology, 95*, 76-93. doi: 10.1037/0022-3514.95.1.76
- Mull, M. S., & Evans, E. M. (2010). Did she mean to do it? Acquiring a folk theory of intentionality. *Journal of Experimental Child Psychology, 107*, 207-228. doi: 10.1016/j.jecp.2010.04.001
- Mulvey, K. L., Hitti, A., & Killen, M. (2010). The development of stereotyping and exclusion. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*, 597-606. Retrieved from <http://dx.doi.org/10.1002/wcs.66> doi:10.1002/wcs.66
- Mulvey, K. L., Hitti, A., & Killen, M. (2013). Intentionality, morality, and exclusion: How children navigate the social world. In M. Banaji & S. Gelman (Eds.), *Navigating the social world: What infants, children and other species can teach us*. (pp. 377-384). Oxford: Oxford University Press.
- Mulvey, K. L., Hitti, A., Rutland, A., Abrams, D., & Killen, M. (under review). Reasoning about resource allocation in an intergroup context from an individual and a group perspective.
- Murray-Close, D., & Ostrov, J. M. (2009). A longitudinal study of forms and functions of aggressive behavior in early childhood. *Child Development, 80*, 828-842. doi: 10.1111/j.1467-8624.2009.01300.x

- NCWGE. (2007). Title IX athletic policies: Issues and data for education decision makers: National Coalition for Women and Girls in Education
- Nesdale, D. (2008). Peer group rejection and children's intergroup prejudice. In S. R. Levy & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 32-46). Oxford, U.K.: Oxford University Press.
- Nesdale, D., Maass, A., Durkin, K., & Griffiths, J. (2005). Group norms, threat, and children's racial prejudice. *Child Development, 76*, 652-663.
- Nesdale, D., Maass, A., Griffiths, J., & Durkin, K. (2003). Effects of in-group and out-group ethnicity on children's attitudes towards members of the in-group and out-group. *British Journal of Developmental Psychology, 21*, 177-192.
- Nesdale, D., Maass, A., Kiesner, J., Durkin, K., Griffiths, J., & Ekberg, A. (2007). Effects of peer group rejection, group membership, and group norms, on children's outgroup prejudice. *International Journal of Behavioral Development, 31*, 526-535. doi: 10.1177/0165025407081479
- Nucci, L. P. (1981). The development of personal concepts: A domain distinct from moral or societal concepts. *Child Development, 52*, 114-121.
- Orobio de Castro, B., Veerman, J. W., Koops, W., Bosch, J. D., & Monshouwer, H. J. (2002). Hostile attribution of intent and aggressive behavior: A meta-analysis. *Child Development, 73*, 916-934. doi: 10.1111/1467-8624.00447
- Ostrov, J. M., & Godleski, S. A. (2010). Toward an integrated gender-linked model of aggression subtypes in early and middle childhood. *Psychological Review, 117*, 233-242.

- Park, Lee-Kim, Killen, Kim, & Park. (2011). Korean children's reasoning about parental expectations regarding gender-stereotypic activities.
- Park, Y., & Killen, M. (2010). When is peer rejection justifiable?: Children's understanding across two cultures. *Cognitive Development, 25*, 290-301.
- 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.
- Piaget, J. (1932). *The moral judgment of the child*. New York: Free Press.
- Piehler, T. F., & Dishion, T. J. (2007). Interpersonal dynamics within adolescent friendships: Dyadic mutuality, deviant talk, and patterns of antisocial behavior. *Child Development, 78*, 1611-1624. doi: 10.1111/j.1467-8624.2007.01086.x
- Pinto, I. R., Marques, J. M., Levine, J. M., & Abrams, D. (2010). Membership status and subjective group dynamics: Who triggers the black sheep effect? *Journal of Personality and Social Psychology, 99*, 107-119. doi: 10.1037/a0018187
- Ritchey, P. N., & Fishbein, H. D. (2001). The lack of an association between adolescent friends' prejudices and stereotypes. *Merrill-Palmer Quarterly, 47*, 188-206.
- Rose, A. J., Swenson, L. P., & Waller, E. M. (2004). Overt and relational aggression and perceived popularity: Developmental differences in concurrent and prospective relations. *Developmental Psychology, 40*, 378-387. doi: 10.1037/0012-1649.40.3.378
- Rubin, K., Bukowski, W., & Parker, J. (2006). Peers, relationships, and interactions. In W. Damon & R. Lerner (Eds.), *Handbook of Child Psychology* (pp. 571-645). NY: Wiley Publishers.

- Ruble, D. N., Martin, C. L., & Berenbaum, S. (2006). Gender development. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Vol. 3, Personality and Social Development (6th ed.)* (pp. 858-932). New York: Wiley Publishers.
- Rutland, A., Killen, M., & Abrams, D. (2010). A new social-cognitive developmental perspective on prejudice: The interplay between morality and group identity. *Perspectives on Psychological Science, 5*, 279-291. doi: 10.1177/1745691610369468
- Salmivalli, C., Voeten, M., & Poskiparta, E. (2011). Bystanders matter: Associations between reinforcing, defending, and the frequency of bullying behavior in classrooms. *Journal of Clinical Child and Adolescent Psychology, 40*, 668-676. doi: 10.1080/15374416.2011.597090
- Scholl, B. J., & Leslie, A. M. (2001). Minds, modules, and meta-analysis. Commentary on 'Meta-analysis of theory-of-mind development: The truth about false belief.'. *Child Development, 72*, 696-701. doi: 10.1111/1467-8624.00308
- Shakoor, S., Jaffee, S. R., Bowes, L., Ouellet-Morin, I., Andreou, P., Happé, F., Moffitt, T. E., & Arseneault, L. (2012). A prospective longitudinal study of children's theory of mind and adolescent involvement in bullying. *Journal of Child Psychology and Psychiatry, 53*, 254-261. doi: 10.1111/j.1469-7610.2011.02488.x
- Shaw, L. A., & Wainryb, C. (2006). When victims don't cry: Children's understandings of victimization, compliance, and subversion. *Child Development, 77*, 1050-1062.
- Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1961). *Intergroup cooperation and competition: The Robbers cave experiment*. Norman: OK: University Book Exchange.

- Shiverick, S. M., & Moore, C. F. (2007). Second-order beliefs about intention and children's attributions of sociomoral judgment. *Journal of Experimental Child Psychology, 97*, 44-60.
- Shure, M. B. (1992). *I Can Problem Solve: An interpersonal cognitive problem-solving program: Kindergarten & primary grades*. Champaign, IL US: Research Press.
- Signorella, M. L., Bigler, R. S., & Liben, L. S. (1993). Developmental differences in children's gender schemata about others: A meta-analytic review. *Developmental Review, 13*, 147-183.
- Sinno, S., & Killen, M. (2009). Moms at work and dads at home: Children's evaluations of parental roles. *Applied Developmental Science, 13*, 16 - 29.
- 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. (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., & Metzger, A. (2008). Don't ask, don't tell (your mother or father): Disclosure and nondisclosure in parent-adolescent relationships. In M. Kerr, H. Stattin & R. Engels (Eds.), *What can parents do: New insights into the role of parents in adolescent problem behavior* (pp. 65-87). New York: Wiley.

- Stone, V. E., & Gerrans, P. (2006). What's domain-specific about theory of mind? *Social Neuroscience, 1*, 309-319. doi: 10.1080/17470910601029221
- Sumter, S. R., Bokhorst, C. L., Steinberg, L., & Westenberg, P. M. (2009). The developmental pattern of resistance to peer influence in adolescence: Will the teenager ever be able to resist? *Journal of Adolescence, 32*, 1009-1021. doi: 10.1016/j.adolescence.2008.08.010
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Monterey, CA: Brooks-Cole.
- Terwogt, M. M., & Rieffe, C. (2003). Stereotyped beliefs about desirability: Implications for characterizing the child's theory of mind. *New Ideas in Psychology, 21*, 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.
- Trach, J., Hymel, S., Waterhouse, T., & Neale, K. (2010). Bystander responses to school bullying: A cross-sectional investigation of grade and sex differences. *Canadian Journal of School Psychology, 25*, 114-130. doi: 10.1177/0829573509357553
- 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? *Infant and Child Development, 14*, 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. (2006). The development of morality. In N. Eisenberg, W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 789-857). Hoboken, NJ: Wiley.
- Wellman, H. M., & Cross, D. (2001). Theory of mind and conceptual change: Reply. *Child Development, 72*, 702-707. doi: 10.1111/1467-8624.00309
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*, 655-684. doi: 10.1111/1467-8624.00304
- Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*, 502-517.
- Wellman, H. M., Lopez-Duran, S., & LaBounty, J. (2008). Infant attention to intentional action predicts preschool theory of mind. *Developmental Psychology, 44*, 618-623.
- Zucker, K., Wilson-Smith, D., Kurita, J., & Stern, A. (1995). Children's appraisals of sex-typed behavior in their peers. *Sex Roles, 33*, 703-725. doi: 10.1007/bf01544775