UNIVERSITY OF CALIFORNIA SAN DIEGO

The Role of Culture in the Development of Prejudice and Moral Reasoning

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Experimental Psychology

by

Haleh Yazdi

Committee in charge:

Professor David Barner, Chair
Professor Gail Heyman, Co-Chair
Professor Pascal Gagneux
Professor Dana Nelkin
Professor Adena Schachner

2022

The Dissertation of Haleh Yazdi is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

DEDICATION

For my mother,

her mother,

and the courageous women in Iran

fighting for a better tomorrow.

زن، زندگی، آزادی

EPIGRAPHS

It is time for parents to teach young people early on that
in diversity there is beauty and there is strength.
We all should know that diversity makes for a rich tapestry,
and we must understand that all the threads of the tapestry are
equal in value no matter their color;
equal in importance no matter their texture.

—Maya Angelou, *Wouldn't Take Nothing for My Journey Now*

People of different religions and cultures live side by side in almost every part of the world,
and most of us have overlapping identities which unite us with very different groups.
We can love what we are, without hating what – and who – we are not.

—Kofi Annan, *Nobel Prize Lecture*

The Sneetches got really quite smart on that day.
The day they decided that Sneetches are Sneetches.
And no kind of Sneetch is the best on the beaches.
That day, all the Sneetches forgot about stars
and whether they had one, or not, upon thars.

— Dr. Suess, *The Sneetches and Other Stories*

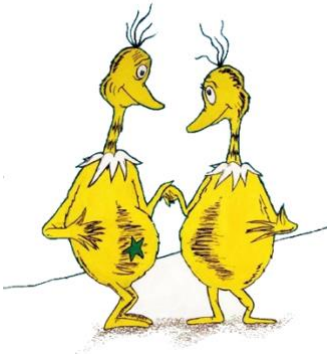

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This dissertation is the culmination of years of collaborative endeavors involving brilliant advisors, generous colleagues, and caring friends and family. Without each of these systems of support, this work would not have been possible. It takes a village to raise a doctoral student and I am so grateful for mine.

I am most grateful to my advisors David Barner and Gail Heyman for generously sharing their time, expertise, and support with me throughout these years. For taking me on as their student and allowing me to be securely attached to both of them: I was given the freedom to pursue the ideas that interested me the most with the reassurance that their expertise and guidance was always available to me. Dave's willingness to meet with me before I was his student and to share my enthusiasm for topics that I hoped to one day explore are the reason behind this work. His expert guidance, rigorous approach to research, open mind, and trust in me shaped me into the researcher and person that I am today. I am equally indebted to Gail for encouraging me to recognize the value of my experiences and for showing me how to marry my personal interests with my academic pursuits. Gail's steadfast belief in me and skillful approach to solving any problem helped see me through numerous moments of doubt and roadblocks. The compassion and concern that she has for each of her students is truly remarkable and I am so fortunate to have been on the receiving end of it. My Ph.D. journey encompasses some of my most enriching, creative, and formative years because of the superb mentorship that I received from Dave and Gail each step of the way.

I also give thanks to the members of the Language and Development Lab who served as my academic family and helped make this work possible. The experiments in this dissertation are due to the hard work of extraordinary research assistants, kind graduate students, and

diligent lab coordinators, all of whom are individually thanked in the acknowledgements of each chapter. I particularly want to acknowledge Rose Schneider whose brilliance as a colleague and encouragement as a friend were instrumental in the completion of Chapter 3 and in helping me get to the finish line.

I extend my appreciation to the Center for Academic Research & Training in Anthropogeny (CARTA), Annette Merle-Smith, and Pascal Gagneux for expanding my community and providing me with extraordinary experiences that fostered my personal and academic growth. I am so thankful for the financial support given to me by this network, which allowed me to freely pursue the topics that I am most passionate about. I am also grateful to Friends of the International Center for awarding me my first travel award, which spearheaded my cross-cultural work and the fellowships and grants that followed.

The years of rigorous work were made easier thanks to an incredible community of family, friends and colleagues who infused my days with laughter and kindness. These wonderful people include Abda Mahmood, Telli Davoodi, Iman Momken, Iris Oved, Liz Seckel, Tanushree Agrawal, Tiffany Wang, Jae Engle, Hessam Kashani, Adriana Pitea, Liesel Spangler, Caren Walker, Joanne Oh, Mehrnoosh Razavi, and many others who uplifted me and rooted for my success. I especially want to thank Kyle Bentz (KCB) for coming on this journey with me and treating my goals as if they were his own. His encouragement, attentive care, and pesto pizzas made everything better.

I owe my deepest gratitude to my parents, Majid Yazdi and Bahareh Momken, for making my education a priority, instilling in me a love for learning, and leading by example as they earned their own doctorate degrees in a foreign country while raising three small children. I thank them for their sacrifices, unwavering support, and unconditional love. Finally, I feel

much appreciation and adoration for my grandmother, Nahid Zohouri, who has been a strong advocate for my success and made it her personal mission to ensure that my research in Iran was possible. Her exceptional character and compassion for others have served as my greatest example.

Chapter 1, in full, is a reprint of the material as it appears in: Yazdi, H., Heyman, G.D., & Barner, D. (2020) Children are sensitive to reputation when giving to both ingroup and outgroup members. Journal of Experimental Child Psychology, 194, 104814. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in full, is a reprint of the material as it appears in: Yazdi, H., Barner, D., & Heyman, G.D. (2020) Children's intergroup attitudes: Insights from Iran. *Child Development, 91* (5), 1733-1744. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in part is currently being prepared for submission for publication of the material. Yazdi, H. Schneider, R.M., Yoon, E.J., Frank, M.C., Srinivasan, M., Dunham, Y. & Barner, D. The dissertation author was the primary investigator and author of this paper.

CURRICULUM VITAE

2008 Bachelor of Arts, Psychology and Sociology
University of Colorado Boulder

2013 Master of Science, Experimental Psychology
University of Oxford, United Kingdom

2022 Doctor of Philosophy, Experimental Psychology
University of California San Diego

PUBLICATIONS

Yazdi, H., Barner, D., & Heyman, G.D. (2020). Children's intergroup attitudes: Insights from Iran. *Child Development, 91*(5), 1733-1744.

Yazdi, H., Heyman, G.D., & Barner, D. (2020). Children are sensitive to reputation when giving to both ingroup and outgroup members. *Journal of Experimental Child Psychology,* 194, 104814.

Heyman, G.D. & Yazdi, H. (2019). The role of individuation in the development of intergroup relations. *Infant and Child Development*, *28*(2), e2121.

Yazdi, H., Barner, D., & Heyman, G.D. (2016). The influence of reputation concerns and social biases on children's sharing behavior. *Proceedings of the 38th Annual Meeting of the Cognitive Science Society.*

ABSTRACT OF THE DISSERTATION

The Role of Culture in the Development of Prejudice and Moral Reasoning

by

Haleh Yazdi

Doctor of Philosophy in Experimental Psychology

University of California San Diego, 2022

Professor David Barner, Chair

Professor Gail D. Heyman, Co-Chair

Prejudices develop early in childhood and can drive disparities in how children treat members of different social groups. These biases can progress into xenophobic and discriminatory acts in adulthood, making it crucial to address them early in life. To effectively do so, we must first identify to what extent prejudices are inevitable or driven by cultural factors. This dissertation investigates the role of culture in children's development of group biases and other moral processes.

In Chapter 1, I find that assigning children to artificially constructed minimal groups (e.g., an Orange or Green group) is sufficient to induce an ingroup bias in children's sharing behavior, and this bias overrides the desire to appear fair and generous to others. These findings suggest that children are predisposed to favor their ingroup over outgroups, but the implications for real-world groups are unclear. In Chapter 2, I address the limitations of Chapter 1 by examining how Iranian children perceive real-world outgroups that differ from their own in similarity, sociopolitical relations, and status. I find that children do not view all outgroups interchangeably, but rather base their group preferences on the relative status of the group in question. These findings highlight the need for more research in non-Western societies to further understand the complexities of children's intergroup attitudes. Finally, in Chapter 3, I highlight the issues that arise when researchers use a standard set of measures developed primarily for Western groups to conduct cross-cultural comparisons. I propose a new two-stage model that combines standardized methods with culturally tailored items to achieve greater validity of measures and reliability of findings across different cultural groups. I demonstrate that this two-stage model is effective in capturing the moral/conventional distinction in children from Canada, India, Iran, and Korea.

This dissertation provides key insights into the cognitive and cultural mechanisms that shape childhood prejudice and highlights new approaches for assessing the role of culture in children's moral reasoning.

INTRODUCTION

Nearly every generation and society has experienced the damaging effects of intergroup conflict, the source of which originates in childhood. Prejudices toward groups develop early in life—even 3-year-olds recognize similarities and differences between themselves and others across traits such as gender and skin tone and actively use this information to categorize individuals into preferred ingroups and disliked outgroups (Nesdale & Flesser, 2001). These biases drive disparities in how children treat group members with regard to social exclusion and harm (Abrams & Killen, 2014) and can progress into discriminatory acts in adulthood, making it crucial to address them early in development. Developmental psychologists have sought to understand the origins of prejudice, but this is a challenging undertaking since we have yet to determine to what extent innate factors, such as the need for group belonging, and cultural factors, such as historical events, shape its trajectory. Another obstacle to our understanding of how group biases form is that humans differ from one another on a multitude of factors, and it is unclear how much findings derived from a subset of cultural groups and environmental contexts can speak to universal human behavior.

Much of the research on childhood prejudice has been conducted in lab settings with children from relatively peaceful Western societies, thereby limiting our understanding of how intergroup attitudes emerge in the real world in different cultural settings. The research presented in this dissertation explores how cultural input shapes the development of prejudice and moral reasoning through experiments conducted in a controlled laboratory setting and in the real world with children from different cultural groups. These are among the first studies to examine these social-cognitive processes in children from Canada, India, Iran, Korea, and the United States—five countries that differ with regard to resource availability, cultural cohesion, and social stability. The aim of this work is to better understand how group biases develop, the role of culture in

shaping its trajectory, and the extent to which findings derived from one cultural group generalize to other groups of children.

Chapter 1 of this dissertation examines whether children are predisposed to treat ingroup and outgroup members differently by testing how children share resources with artificial groups that carry no cultural significance. Chapter 2 of this dissertation explores how cultural input shapes children's preferences for groups that exist in the real world and have critical social and political implications for their lives. Finally, Chapter 3 investigates whether methods that have been primarily validated and developed for Western groups can capture the moral reasoning abilities of children across the globe.

Background Information

A widely held belief among developmental psychologists is that prejudices form early in life, are inevitable, and emerge from children's preference for similar individuals and dislike of dissimilar individuals. In line with these theories, findings from infants and preschoolers indicate that children generally prefer people who share their language (Kinzler, Dupoux, & Spelke, 2007), gender (Hilliard & Liben, 2010; Quinn et al., 2008; Shutts, Banaji, & Spelke, 2010), and age (Aboud, 1988; French, 1984), and this preference extends to race and ethnicity as they get older (Banaji, Baron, Dunham, & Olson, 2008; Kinzler & Spelke, 2011). Children are also more likely to befriend, trust, and cooperate with members of their ingroup over outgroup members (Elashi & Mills, 2014; Fehr, Bernhard, & Rockenbach, 2008; Rotenberg & Cerda, 1994), demonstrating that children's cognitive biases give rise to differential treatment of group members.

Theories of social group formation

To understand how and why children form prejudices toward different groups of people, we need to identify the extent to which specific cultural factors and general cognitive factors shape

these processes. There are several existing theories that speak to the cognitive mechanisms that predispose children to forming group biases. The first is self-categorization theory (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987), which argues that children and adults are inclined to group themselves and others into social categories (e.g., old and young, male and female) to help establish order and make sense of the social world. Children use social categories to indicate why certain people belong to the same category (Hirschfeld, 1996; Heyman & Gelman, 2000) and to form predictions for how group members should look and behave (Diesendruck & haLevi, 2006; Rhodes, 2012). For example, preschoolers use gender categories to infer how girls and boys should behave (e.g., boys play with trucks and girls play with dolls; Killen, Piscane, Lee-Kim, & Ardila-Rey, 2001), age to infer who has more knowledge (Pillow & Weed, 1997), race to infer who has power (Dukler & Liberman, 2022), and national identity to infer whether group members pose a political threat (Bennett et al., 2004). According to self-categorization theory, children are more likely to use social categories as signifiers of differences between groups if the category distinction is made more salient and accessible, and if group members fit widely held stereotypes (Oakes, Turner, & Haslam, 1991).

As children become increasingly aware of social categories, they also begin to establish their social identity—aspects of their self-image that are derived from the social groups they belong to (Tajfel & Turner, 1979). According to social-identity theory, children's need for belonging and desire to have a positive self-concept motivates them to form their social identity in the context of their membership within social groups (Baumeister & Leary, 1995). One way that children maintain a positive social identity is by viewing their group members as more similar to them and more favorable than members of outgroups. As children associate more of their identity

with their ingroup, they come to strengthen their ingroup bias and view their own group as superior to others as a means for enhancing their own self-concept (Brown, 1995; Hogg & Abrams, 1988).

A third explanation for why intergroup biases develop is essentialist thinking, or believing that categories reflect something deep, stable, and informative about their members (Gelman, 2003; Rhodes, Leslie, Saunders, Dunham, & Cimpian, 2015). Developmental work finds that children's group biases are especially strong toward groups they believe to have essentialist properties, meaning that group membership is perceived as permanent, biologically determined, and indicative of underlying similarities among group members (Diesendruck & haLevi, 2006; Gelman, 2003; Hirschfeld, 2005). For example, preschool children who hold essentialist beliefs about gender perceive boys and girls as fundamentally different from each other and expect them to behave in ways that fit gender stereotypes (Gelman, Collman & Maccoby, 1986; Rhodes & Gelman, 2009). For some social categories, such as gender, children are taught essentialist views at a young age, but for others, such as race or status, children's views vary more as a function of their cultural experiences (Rhodes & Gelman, 2009; Shutts, Roben, & Spelke, 2013; Waxman, 2010).

Cultural input and experience play a significant role in determining what type of category children essentialize. For instance, Israeli children essentialize religious-ethnic categories at a younger age and more strongly than American children, who show a similar pattern for race (Diesendruck, Goldfein-Elbaz, Rhodes, Gelman, & Neumark, 2013). Chilean children of high socio-economic status essentialize poverty more than those of low socio-economic status (Del Río & Strasser, 2011), and findings from Indian adults show that higher caste individuals essentialize caste membership more than lower caste individuals (Mahalingam, 2007). Together, these findings suggest that children expect some social categories to reflect essential kinds, but for others,

essentialist views are a product of social and cultural learning (Heyman & Giles, 2006; Mahalingam, 2003). The work in this dissertation aims to uncover some of the ways that cultural input shapes children's biases and behaviors toward different social groups.

Measuring children's intergroup biases: Minimal vs. real-world groups

In measuring children's intergroup biases, there is tension between maximizing experimental control to isolate causal relations and understanding how these judgments take shape in ecologically valid, real-world contexts. In this dissertation I begin with one study that maximizes experimental control, using what is known as a minimal group paradigm, and two studies that take the latter approach.

Minimal groups

Minimal groups are arbitrarily constructed groups that are novel, equal in status, and have no real-world significance (Billig & Tajfel, 1973; Tajfel, Billig, Bundy, & Flament, 1971). They are often established by randomly assigning individuals to one of two color-based categories and using visible identifiers such as t-shirts or wristbands to signify group membership (Dunham, Baron, & Carey, 2011; Schug, Shusterman, Barth, & Patalano, 2013; Sparks, Schinkel, & Moore, 2017). The minimal group paradigm is an effective tool for highlighting group membership while controlling for other dimensions that are inherent in real-world groups (e.g., status, size, familiarity). Some developmental studies find that children show an immediate ingroup preference in their attitudes and behaviors upon being assigned to a minimal group, suggesting that the mere process of group categorization can induce an "us" versus "them" mentality (Billig & Tajfel, 1973; Bigler, 1995; Bigler, Jones, & Lobliner, 1997; Richter, Over, & Dunham, 2016; Tajfel et al., 1971). On the other hand, other work has found that assigning children to minimal groups without introducing other shared characteristics among groups members is insufficient to produce an

ingroup bias in their cooperative behavior (Patterson & Bigler, 2006; Rabbie & Horwitz, 1969; Rabbie & Wilkens, 1971; Spielman, 2000). Therefore, it is possible that minimal group membership probes children's cognitive biases, but the groups must carry greater meaning for children to act on these biases.

Few studies have contrasted minimal group membership against other social concerns to test the strength of minimal groups. In Chapter 1, I used minimal color groups (e.g., the Orange or Green group) to examine how children treat ingroup and outgroup members when the desire to favor one's own group is pitted against the motive to appear fair and generous. I show that randomly assigning children to minimal groups is sufficient to induce a robust ingroup preference in their sharing behavior, and this preference persists when competing social concerns, such as the desire to maintain a positive reputation, are present. The findings that I present in Chapter 1 reveal that both children and adults have a propensity to favor their ingroup on the sole basis of group membership.

The minimal group paradigm, while an effective tool for uncovering the basis of intergroup biases, has several critical limitations. First, it is unclear to what extent findings from artificial social groups generalize to groups that exist in the real world. Unlike minimal groups, existing groups have multiple dimensions, such as status, history of conflict, and cultural differences, which children take into consideration when forming beliefs about group members (Newheiser, Dunham, Merrill, Hoosain, & Olson, 2014). The minimal group paradigm does not speak to how children weigh different, and often competing, group dimensions when determining which groups they want to associate with. Another limitation is that the minimal group paradigm assumes that children rely on group labels alone to make judgements about group members, but it is possible that children make inferences about the group in question that extend beyond its minimal qualities.

For instance, in real-world scenarios, children make inferences about personality traits from gender groups (Berndt & Heller, 1986; Taylor, 1996), social status from racial groups (Dukler & Liberman, 2022; Mandalaywala, Tai, & Rhodes, 2020), and religious affiliation from ethnic groups (Birnbaum, Deeb, Segall, Ben-Eliyahu, & Diesendruck, 2010). Therefore, it is possible that children infer meaning from minimal groups and ascribe characteristics to group members beyond the scope of the minimal group paradigm. In Chapter 2, I address some of these limitations and show that by asking children about real-world groups that are socially, historically and culturally significant to them, we can achieve a more comprehensive understanding of children's development of intergroup attitudes.

Real-world groups

Studies with real-world groups typically involve comparing children's judgements of individuals from their own group versus a different group, where groups are defined with reference to pre-existing social distinctions that even young children are likely to be familiar with such as gender, race, nationality, religion, or language. The chief advantage of using real-world groups over artificially created groups is that the measures and findings have more transparent ecological validity. Children acquire information about existing social groups through mere exposure, cultural messages, nonverbal cues, and linguistic cues (Diesendruck & Deblinger-Tangi, 2014; Rhodes, Leslie, Bianchi, & Chalik, 2017; Rhodes, Leslie, & Tworek, 2012; Skinner, Meltzoff, & Olson, 2017; Waxman, 2010)— social input that is not captured within a typical minimal group paradigm. By gauging children's attitudes toward real-world groups, we can examine how children's biases are shaped by years of cultural input rather than a more abstract learning of a minimal group boundary.

In contrast to minimal groups, where only one group dimension is made salient, real-world groups have multiple dimensions (e.g., race, gender, status), which children become increasingly aware of with age (Baron & Banaji, 2006; Hirschfeld, 1988; Rhodes & Gelman, 2009). Most studies highlight only one group dimension at a time to test the causal relationship between activating a specific domain and children's biases, so it remains unclear how children perceive different group dimensions. A few studies show that children prioritize shared interests (e.g., sports activities) over shared physical features (e.g., skin color) when selecting friends (McGlothlin, Killen & Edmonds, 2005) and that they prefer members of their gender ingroup over members of their minimal group when the two dimensions are pitted against one another (Yang, Yang, Guo, & Dunham, 2022). However, the limited empirical work in this area leaves open the question of whether children prioritize certain group dimensions over others when forming group preferences.

It is possible that children do not distinguish between group dimensions and view all outgroups as interchangeable—as observed in minimal group studies. Alternatively, perhaps children do make distinctions between group dimensions and prioritize some over others. In Chapter 2, I explore this question by examining whether children in Iran differentiate between real-world outgroups that vary from their own group on the dimensions of perceived similarity, social status, and socio-political relations. I find evidence that children do not view all outgroups as interchangeable, but instead prioritize some dimensions over others when determining their group preferences. In particular, Iranian children favored American children, the outgroup which they perceived as having the highest social status, over children from groups that they perceived as having lower status who they are more similar to with regards to religion, culture, and proximity (e.g., Iranian children from a different school, Arab children). By examining how Iranian children perceive groups with which they have social and political ties, my findings illuminate a critical

aspect of intergroup relations that has major implications for children's lives. Chapter 2 highlights the significance of assessing children's beliefs about real-world groups and the need for more research with children from non-Western societies in developmental research more broadly.

WEIRD groups are overrepresented in psychology research

The vast majority of research studies in the behavioral sciences have been conducted in Western, educated, industrialized, rich and democratic (WEIRD) societies (Apfelbaum, Phillips, & Richeson, 2014; Henrich, Heine, & Norenzayan, 2010; Legare & Harris, 2016; Meadon & Spurret, 2010), which represent 12% of the world's population yet constitute 96% of the samples in psychological research (Arnett, 2009; Hardin, Robitschek, Flores, Navarro, & Ashton, 2014). This is a major problem, since WEIRD populations are not representative of human culture, and the findings from these studies are limited in what they can tell us about the universality of human behaviors.

A growing body of evidence from non-Western groups show considerable cultural variation in fundamental cognitive, affective, and behavioral processes. For instance, cultural experience shapes how people perceive and process basic sensory-perceptual information such as taste (Moskowitz, Kumaraiah, Sharma, Jacobs, & Sharma, 1975), odor (Majid & Burenhult, 2014), and color (Roberson, Davies, & Davidoff, 2000), as well as visual-spatial information such as direction (Levinson, 2007) and location (Majid, Bowerman, Kita, Haun, & Levinson, 2004). Given that even basic perceptual and cognitive processes vary considerably across cultural groups, we can expect social beliefs and behaviors to be particularly susceptible to cultural influences, and studies do show that cultural conventions and beliefs shape the development of children's cooperative tendencies (see Hitti, Mulvey, & Killen, 2011 for a review; Stewart & McBride-Chang, 2000), reputation management strategies (Fu, Heyman, Cameron, & Lee, 2016),

endorsement of stereotypes (Oren & Bar-Tal, 2007), and moral beliefs (Hollos, Leis, & Turiel, 1986; Yau & Smetana, 2003), in addition to many other social cognitive processes.

Despite evidence of substantial variation in the psychological processes of children from different cultural groups, some non-WEIRD groups, such as Iranian children, remain relatively absent from developmental psychology research. As I show in Chapter 2, Iranian children's group preferences are shaped by cultural and historical factors, and contrary to findings from studies of Western children (Bar-Haim, Ziv, Lamy, & Hodes, 2006; Mahajan & Wynn, 2012; McGlothlin & Killen, 2005), Iranian children do not show a clear preference for similar individuals. Further, our understanding of how adverse experiences such as poverty and socio-political conflict shape children's cooperative behaviors is limited since most research in psychology is conducted in peaceful Western societies where resources are abundant. To establish a more comprehensive understanding of how social and moral processes develop early in life, more research is needed with children residing in diverse socio-historical settings. Chapters 2 and 3 of this dissertation focus on the social cognitive development of children in non-WEIRD societies to advance this effort.

WEIRD measures pose a problem for cross-cultural research

In the last decade there has been an increase in cross-cultural research in efforts to include more non-WEIRD groups to psychology literature. Too often, however, the methods used in these studies remain WEIRD since they were primarily developed and validated in Western contexts. As a result, such measures cannot fully capture the historical, cultural, and political forces that shape the behaviors of the groups of focus. Also, when methods are developed primarily for WEIRD groups and then used to test for cross-cultural variability, the results are often biased in favor of the majority group since the measures more accurately reflect their social and cultural

experiences. For example, a standard IQ test established by Western researchers and initially validated with Western samples has been used to gauge the cognitive abilities of Argentinian, Colombian native American, Chinese, Italian, and Middle Eastern groups (Razani, Murcia, Tabares, & Wong, 2007; Weschler, 1997). The scores derived from this measure show high performance in Western groups and lower performance in some non-Western groups (Lange, 2007; Melendez, 1994)—a result which may be more reflective of differences between groups in their familiarity or understanding of the measure rather than intellectual ability (Shuttleworth-Edwards, 2016; Sunderaraman, Zahodne, & Manly, 2016).

The use of standardized measures to make group comparisons is not inherently problematic, as having a uniform method for collecting and analyzing data can be necessary for replicating studies, minimizing bias in the testing process, and producing results that are generalizable to a larger population. Such measures are valuable to identify aspects of human behavior that are universal versus specific to certain cultural groups. However, when items on a standardized measure have low cross-cultural equivalence, or carry different meaning for each culture, then the comparison of responses across different cultures can be misleading (Chen, 2008; Hambleton, Merenda, & Spielberger, 2004; Norenzayan & Heine, 2005). To increase the validity of research measures for different groups and gain more insight into why certain response patterns emerge, items should be tailored to the local customs, laws, beliefs, and social climate of the groups studied. One effective strategy for addressing issues of generalizability and cultural validity within a single study is to combine standardized measures with culturally tailored measures. In Chapter 3, I show how a two-stage model that adheres to this strategy achieves both high reliability and cultural validity in the study of children's moral development.

Current Directions

This dissertation contains three chapters detailing empirical studies conducted in Canada, India, Iran, Korea, and the United States with children between the ages of five and fourteen. Chapters 1 and 2 examine how and why children form prejudices toward different groups of people, and Chapter 3 examines how children from different cultural groups develop the ability to distinguish between violations of moral codes and violations of conventional norms. All three chapters highlight the role that cultural factors play in shaping children's social and cognitive development.

In Chapter 1, I examine whether different motives are at play when children cooperate with members of their own group versus members of a different group. There is considerable evidence that children and adults behave more cooperatively toward members of their own group over other groups (Nesdale & Flesser, 2001; Tajfel & Turner 2004) and this tendency is evident in early childhood (Moore, 2009; Yu, Zhu, & Leslie, 2016), but what remains less known is *why* these biases develop. Prior research suggests that prosociality toward outgroup members may be more driven by self-serving motives, such as reputation concerns, whereas ingroup prosociality is believed to be more motivated by empathic concerns (e.g., concern for the well-being of recipients; Levine, Prosser & Reicher, 2005; Stürmer, Snyder, Kropp, & Siem, 2006). However, there is no prior developmental work that tests whether children, like adults, have differing motives when cooperating with ingroup versus outgroup members. Further, only one other study has examined whether there are contexts for which self-serving motives also drive generosity toward ingroup members (Engelmann, Over, Herrmann, & Tomasello, 2013). In Chapter 1 of this dissertation, I answer these questions by examining whether ingroup or outgroup sharing is more driven by self-serving motives, specifically reputation concerns.

In Chapter 2 of this dissertation, I turn my attention to the question of *how* children form biases toward groups that exist in the real world and are meaningful to them. Chapter 1 shows that mere social categorization can lead to differences in how children treat ingroup and outgroup members, but this work is not informative with regards to whether children reason about different types of outgroups. Chapter 2 addresses this limitation by exploring which aspects of social groups matter most to children as they develop beliefs about real-world groups. Iranian children between the ages of 7 and 12 were asked to evaluate their ingroup and three different outgroups (Iranian children from a different school, Arab children, and American children) that contrasted from their own with regard to similarity, social and political relations, and social status. These groups were selected because they are personally significant to Iranian children, which allows us to examine how children come to form biases towards groups that have critical implications for their lives. The findings presented in Chapter 2 offer some of the only insight thus far into Iranian children's attitudes toward real-world national groups.

In Chapter 3, I highlight the problem of validity that arises when researchers use measures developed in WEIRD contexts to compare groups across diverse cultural settings. I propose a two-stage model that combines standardized measures with culturally tailored measures and test its efficacy in capturing the moral/conventional distinction in Canadian, Indian, Iranian, and Korean children. The moral/conventional distinction, or ability for children to distinguish between moral violations that cause harm to another person and violations of social conventions (Haidt & Graham, 2007; Nisan, 1987; Smetana & Braeges, 1990; Turiel, 1983), is a good case study for testing the two-stage model since it is argued to be a developmental milestone that is universally shared. I show how the use of standardized items, that have been primarily validated with Western groups, can lead to inconclusive findings when used to make cross-cultural comparisons. However, when

standardized methods are combined with items that are tailored to the cultural conventions and beliefs of the group of focus, both the reliability of the findings and cultural validity of the measures are enhanced.

Together, the studies in the following chapters further our understanding of the role that culture plays in shaping children's social and moral development. I examine children's motives for treating ingroup and outgroup members differently when the groups are novel, artificial, and minimally different from one another (Chapter 1). I then explore how children form beliefs about real-world groups that differ from their own on important cultural dimensions including status and political relations (Chapter 2). Finally, I test the extent to which children's social and moral reasoning is shaped by their cultural climate using a model that yields greater reliability of findings across cultural groups and validity of measures (Chapter 3).

References

Aboud, F.E. (1988). *Children and prejudice*. Cambridge, MA: Blackwell.

Abrams, D., & Killen, M. (2014). Social exclusion of children: Developmental origins of prejudice. *Journal of Social Issues*, *70*(1), 1-11.

Apfelbaum, E. P., Phillips, K. W., & Richeson, J. A. (2014). Rethinking the baseline in diversity research: Should we be explaining the effects of homogeneity?. *Perspectives on Psychological Science*, *9*(3), 235-244.

Arnett, J. J. (2009). The neglected 95%, a challenge to psychology's philosophy of science.

Banaji, M. R., Baron, A. S., Dunham, Y., & Olson, K. (2008). The development of intergroup social cognition: Early emergence, implicit nature, and sensitivity to group status. In S.R. Levy, M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 197-236*)*, New York: Oxford University Press.

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. *Psychological Science*, *17*(2), 159-163.

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science*, *17*(1), 53-58.

Baumeister R.F., Leary M.R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117, 497–529.

Bennett, M., Barrett, M., Karakozov, R., Kipiani, G., Lyons, E., Pavlenko, V., & Riazanova, T. (2004). Young children's evaluations of the ingroup and of outgroups: A multi-national study. *Social Development*,*13*(1), 124-141.

Berndt, T. J., & Heller, K. A. (1986). Gender stereotypes and social inferences: A developmental study. *Journal of Personality and Social Psychology*, *50*(5), 889.

Bigler, R. S. (1995). The role of classification skill in moderating environmental influences on children's gender stereotyping: A study of the functional use of gender in the classroom. *Child Development*, *66*(4), 1072-1087.

Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development*, *68*(3), 530-543.

Billig, M., & Tajfel, H. (1973). Social categorization and similarity in intergroup behavior. *European Journal of Social Psychology*, *3*(1), 27-52.

Birnbaum, D., Deeb, I., Segall, G., Ben-Eliyahu, A., & Diesendruck, G. (2010). The development of social essentialism: The case of Israeli children's inferences about Jews and Arabs. *Child Development, 81*(3), 757-777.

Brown, R. (1995). Prejudice. Its social psychology. Blackwell, Oxford, U.K.

Chalik, L., Leslie, S.-J., & Rhodes, M. (2017). Cultural context shapes essentialist beliefs about religion. *Developmental Psychology, 53*(6), 1178–1187.

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology*, *95*(5), 1005.

Del Río, M. F., & Strasser, K. (2011). Chilean children's essentialist reasoning about poverty. *British Journal of Developmental Psychology*, *29*(4), 722-743.

Diesendruck, G., & Deblinger-Tangi, R. (2014). The linguistic construction of social categories in toddlers. *Child Development, 85*(1), 114-123.

Diesendruck, G., Goldfein-Elbaz, R., Rhodes, M., Gelman, S., & Neumark, N. (2013). Cross-cultural differences in children's beliefs about the objectivity of social categories. *Child Development*, *84*(6), 1906-1917.

Diesendruck, G., & HaLevi, H. (2006). The role of language, appearance, and culture in children's social category-based induction. *Child Development*, *77*(3), 539-553.

Dukler, N., & Liberman, Z. (2022). Children use race to infer who is "in charge". *Journal of Experimental Child Psychology*, *221*, 105447.

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of ''minimal" group affiliations in childhood. Child Development, 82, 793–811.

Elashi, F. B., & Mills, C. M. (2014). Do children trust based on group membership or prior accuracy? The role of novel group membership in children's trust decisions. *Journal of Experimental Child Psychology*, *128*, 88-104.

Engelmann, J. M., Over, H., Herrmann, E., & Tomasello, M. (2013). Young children care more about their reputation with ingroup members and potential reciprocators. *Developmental Science*, *16*(6), 952-958.

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature, 454*(7208), 1079-1083.

French, D. C. (1984). Children's knowledge of the social functions of younger, older, and same-age peers. *Child Development*, *55*(4), 1429-1433.

Fu, G., Heyman, G. D., Cameron, C. A., & Lee, K. (2016). Learning to be unsung heroes: Development of reputation management in two cultures. *Child Development*, *87*(3), 689-699.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford University Press.

Gelman, S. A., Collman, P., & Maccoby, E. E. (1986). Inferring properties from categories versus inferring categories from properties: The case of gender. *Child Development*, *57*(2), 396-404.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*(1), 98-116.

Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (2004). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In *Adapting educational and psychological tests for cross-cultural assessment* (pp. 15-50). Psychology Press.

Hardin, E. E., Robitschek, C., Flores, L. Y., Navarro, R. L., & Ashton, M. W. (2014). The cultural lens approach to evaluating cultural validity of psychological theory. *American Psychologist, 69,* 656–668.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2-3), 61-83.

Hilliard, L. J., & Liben, L. S. (2010). Differing levels of gender salience in preschool classrooms: Effects on children's gender attitudes and intergroup bias. *Child Development*, *81*(6), 1787-1798.

Hirschfeld, L. A. (1988). On acquiring social categories: cognitive development and anthropological wisdom. *Man*, 611-638.

Hirschfeld, L. A. (1996). *Race in the making: Cognition, culture, and the child's construction of human kinds*. MIT Press.

Hirschfeld, L. (2005). Children's understanding of racial groups. In M. Barret & E. Buchanan-Barrow (Eds.), *Children's understanding of society* (pp. 251–285). New York: Psychology Press.

Heyman, G. D., & Gelman, S. A. (2000). Preschool children's use of trait labels to make inductive inferences. *Journal of Experimental Child Psychology*, *77*(1), 1-19.

Heyman, GD, & Giles, JW (2006). Gender and psychological essentialism. *Childhood*, *58*(3), 293-310.

Hitti, A., Mulvey, K. L., & Killen, M. (2011). Social exclusion and culture: The role of group norms, group identity and fairness. *Anales de Psicología*, *27*(3), 587-599.

Hogg, M. A., & Abrams, D. (1988). *Social identifications: A social psychology of intergroup relations and group processes*. London: Routledge.

Hollos, M., Leis, P. E., & Turiel, E. (1986). Social reasoning in Ijo children and adolescents in Nigerian communities. *Journal of Cross-Cultural Psychology*, *17*(3), 352-374.

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*(5), 587.

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, *104*, 12577-12580.

Kinzler, K. D., & Spelke, E. S. (2011). Do infants show social preferences for people differing in race?. *Cognition*, *119*(1), 1-9.

Legare, C. H., & Harris, P. L. (2016). The ontogeny of cultural learning. *Child Development*, *87*(3), 633-642.

Levine, M., Prosser, A., Evans, D., & Reicher, S. (2005). Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shape helping behavior. *Personality and Social Psychology Bulletin*, *31*(4), 443-453.

Levinson, S. C. (2007). Optimizing person reference-perspectives from usage on Rossel Island. In *Person reference in interaction: Linguistic, cultural, and social perspectives* (pp. 29-72). Cambridge University Press.

Mahajan, N., & Wynn, K. (2012). Origins of "us" versus "them": Prelinguistic infants prefer similar others. *Cognition*, *124*(2), 227-233.

Mahalingam, R. (2003). Essentialism, culture, and power: Representations of social class. *Journal of Social Issues*, *59*(4), 733-749.

Mahalingam, R. (2007). Essentialism, power, and the representation of social categories: A folk sociology perspective. *Human Development*, *50*(6), 300-319.

Majid, A., Bowerman, M., Kita, S., Haun, D. B., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, *8*(3), 108-114.

Majid, A., & Burenhult, N. (2014). Odors are expressible in language, as long as you speak the right language. *Cognition*, *130*(2), 266-270.

Mandalaywala, T. M., Tai, C., & Rhodes, M. (2020). Children's use of race and gender as cues to social status. *PloS one*, *15*(6), e0234398.

McGlothlin, H., & Killen, M. (2005). Children's perceptions of intergroup and intragroup similarity and the role of social experience. *Journal of Applied Developmental Psychology*, *26*(6), 680-698.

McGlothlin, H., Killen, M., & Edmonds, C. (2005). European-American children's intergroup attitudes about peer relationships. *British Journal of Developmental Psychology*, *23*(2), 227-250.

Meadon, M., & Spurrett, D. (2010). It's not just the subjects-there are too many WEIRD researchers. *Behavioral and Brain Sciences*, *33*(2-3), 104.

Melendez, F. (1994). The Spanish version of the WAIS: Some ethical considerations. *The Clinical Neuropsychologist*, *8*(4), 388-393.

Moore, C. (2009). Fairness in children's resource allocation depends on the recipient. *Psychological Science*, *20*(8), 944-948.

Moskowitz, H. W., Kumaraiah, V., Sharma, K. N., Jacobs, H. L., & Sharma, S. D. (1975). Cross-cultural differences in simple taste preferences. *Science*, *190*(4220), 1217-1218.

Nesdale, D., & Flesser, D. (2001). Social identity and the development of children's group attitudes. *Child Development, 72,* 506–517.

Newheiser, A. K., Dunham, Y., Merrill, A., Hoosain, L., & Olson, K. R. (2014). Preference for high status predicts implicit outgroup bias among children from low-status groups. *Developmental Psychology*, *50*(4), 1081.

Nisan, M. (1987). Moral norms and social conventions: A cross-cultural comparison. *Developmental Psychology*, *23*(5), 719.

Norenzayan, A., & Heine, S. J. (2005). Psychological universals: What are they and how can we know?. *Psychological Bulletin*, *131*(5), 763.

Oakes, P. J., Turner, J. C., & Haslam, S. A. (1991). Perceiving people as group members: The role of fit in the salience of social categorizations. *British Journal of Social Psychology*, *30*(2), 125-144.

Oren, N., & Bar-Tal, D. (2007). The detrimental dynamics of delegitimization in intractable conflicts: The Israeli–Palestinian case. *International Journal of Intercultural Relations*, *31*(1), 111-126.

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*(4), 847-860.

Pillow, B. H., & Weed, S. T. (1997). Preschool children's use of information about age and perceptual access to infer another person's knowledge. *The Journal of Genetic Psychology*, *158*(3), 365-376.

Quinn, P. C., Uttley, L., Lee, K., Gibson, A., Smith, M., Slater, A. M., & Pascalis, O. (2008). Infant preference for female faces occurs for same-but not other-race faces. *Journal of Neuropsychology*, *2*(1), 15-26.

Rabbie, J. M., & Horwitz, M. (1969). Arousal of ingroup-outgroup bias by a chance win or loss. *Journal of Personality and Social Psychology*, *13*(3), 269.

Rabbie, J. M., & Wilkens, G. (1971). Intergroup competition and its effect on intragroup and intergroup relations. *European Journal of Social Psychology*, *1*(2), 215-234.

Razani, J., Murcia, G., Tabares, J., & Wong, J. (2007). The effects of culture on WASI test performance in ethnically diverse individuals. *The Clinical Neuropsychologist*, *21*(5), 776-788.

Rhodes, M. (2012). Naïve theories of social groups. *Child Development*, *83*(6), 1900-1916.

Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology*, *59*(3), 244-274.

Rhodes, M., Leslie, S. J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences*, *109*(34), 13526-13531.

Rhodes, M., Leslie, S. J., Saunders, K., Dunham, Y., & Cimpian, A. (2018). How does social essentialism affect the development of inter-group relations?. *Developmental Science*, *21*(1), e12509.

Richter, N., Over, H., & Dunham, Y. (2016). The effects of minimal group membership on young preschoolers' social preferences, estimates of similarity, and behavioral attribution. *Collabra*, *2*(1).

Roberson, D., Davies, I., & Davidoff, J. (2000). Color categories are not universal: replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General*, *129*(3), 369.

Rotenberg, K. J., & Cerda, C. (1994). Racially based trust expectancies of Native American and Caucasian children. *The Journal of Social Psychology*, *134*(5), 621-631.

Schug, M. G., Shusterman, A., Barth, H., & Patalano, A. L. (2013). Minimal-group membership influences children's responses to novel experience with group members. *Developmental Science, 16*(1), 47-55.

Shuttleworth-Edwards, A. B. (2016). Generally representative is representative of none: commentary on the pitfalls of IQ test standardization in multicultural settings. *The Clinical Neuropsychologist, 30*(7), 975-998.

Shutts, K., Banaji, M. R., & Spelke, E. S. (2010). Social categories guide young children's preferences for novel objects. *Developmental Science, 13*(4), 599-610.

Shutts, K., Roben, C. K. P., & Spelke, E. S. (2013). Children's use of social categories in thinking about people and social relationships. *Journal of Cognition and Development, 14*(1), 35-62.

Skinner, A. L., Meltzoff, A. N., & Olson, K. R. (2017). "Catching" social bias: Exposure to biased nonverbal signals creates social biases in preschool children. *Psychological Science, 28*(2), 216-224.

Smetana, J. G., & Braeges, J. L. (1990). The development of toddlers' moral and conventional judgments. *Merrill-Palmer Quarterly, 36*(3), 329-346.

Sparks, E., Schinkel, M. G., & Moore, C. (2017). Affiliation affects generosity in young children: The roles of minimal group membership and shared interests. *Journal of Experimental Child Psychology, 159*, 242-262.

Spielman, D. A. (2000). Young children, minimal groups, and dichotomous categorization. *Personality and Social Psychology Bulletin, 26*(11), 1433-1441.

Stewart, S. M., & McBride-Chang, C. (2000). Influences on children's sharing in a multicultural setting. *Journal of Cross-Cultural Psychology, 31*(3), 333-348.

Stürmer, S., Snyder, M., Kropp, A., & Siem, B. (2006). Empathy-motivated helping: The moderating role of group membership. *Personality and Social Psychology Bulletin, 32*(7), 943-956.

Sunderaraman, P., Zahodne, L. B., & Manly, J. J. (2016). A commentary on 'generally representative is representative of none: Pitfalls of IQ test standardization in multicultural settings' by AB Shuttleworth-Edwards. *The Clinical Neuropsychologist, 30*(7), 999-1005.

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

Tajfel, H., & Turner, J. C. (1979). An integrative theory of inter-group conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of inter-group relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Tajfel, H., & Turner, J. C. (2004). The social identity theory of intergroup behavior. In *Political psychology* (pp. 276-293). Psychology Press.

Taylor, M. G. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development*, *67*(4), 1555-1571.

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

Turner, J. C., Hogg, M., Oakes, P., Reicher, S., & Wetherell, M. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford, England: Basil Blackwell.

Waxman, S. R. (2010). Names will never hurt me? Naming and the development of racial and gender categories in preschool-aged children. *European Journal of Social Psychology, 40*(4), 593-610.

Wechsler, D. (1997). *Wechsler adult intelligence scale-third edition (WAIS-III)*. San Antonio, TX: The Psychological Corporation.

Yang, X., Yang, F., Guo, C., & Dunham, Y. (2022). Which group matters more: The relative strength of minimal vs. gender and race group memberships in children's intergroup thinking. *Acta Psychologica, 229,* 103685.

Yau, J., & Smetana, J. G. (2003). Conceptions of moral, social-conventional, and personal events among Chinese preschoolers in Hong Kong. *Child Development, 74*(3), 647-658.

Yu, J., Zhu, L., & Leslie, A. M. (2016). Children's sharing behavior in mini-dictator games: The role of in-group favoritism and theory of mind. *Child Development*, *87*(6), 1747-1757.

CHAPTER 1

Children are sensitive to reputation when giving to both ingroup and outgroup members

Haleh Yazdi, Gail D. Heyman, and David Barner

Department of Psychology, University of California, San Diego

Abstract

Previous studies establish that reputation concerns play an important role in outgroup giving. However, it is unclear whether the same is true for ingroup giving, which by some accounts tends to be motivated by empathic concerns. To explore this question, we tested the extent to which 5 to 9-year-old children (Study 1: N = 164) and adults (Study 2: N = 80) shared resources with ingroup and outgroup members, either when being watched by an observer (where we expected reputation concerns to be salient) or in private (where we expected no effect of reputation concerns). We also assessed whether children and adults differ in their beliefs about which form of sharing (ingroup or outgroup giving) is nicer. Although we found that both children and adults exhibited an ingroup bias when sharing, there was no evidence in either group that reputation concerns were greater for outgroup members than for ingroup members. We also found that, in contrast to adults, children shared more resources when observed than in private. Additionally, children evaluated ingroup giving as nicer across different sharing scenarios, whereas adults identified outgroup giving as nicer when the two forms of giving were contrasted. These results are the first to suggest that reputational concerns influence children's sharing both with ingroup and outgroup members, and that children differ from adults in their reasoning about which form of group sharing is nicer.

1. Introduction

Humans often share with one another, and frequently do so without expectation of immediate reciprocation. We share food, money, time, and even ideas with other people (Bock & Kim, 2002; Fehr & Fischbacher, 2003), thereby fostering social cohesion and the achievement of goals that might otherwise be difficult to attain (Batson, 1994). However, we don't share our resources equally with all people, nor do we always share for the same reasons (Bernhard, Fischbacher, & Fehr, 2006; Stürmer, Snyder, Kropp, & Siem, 2006). Generally, individuals are more likely to share with members of their own social groups (i.e., ingroups) than with members of other groups (i.e., outgroups; Nesdale & Flesser, 2001; Tajfel & Turner, 2004), a tendency that is evident as early as the preschool years (Fehr, Bernhard, & Rockenbach, 2008; Moore, 2009; Yu, Zhu, & Leslie, 2016). Further, children and adults tend to act more generously when observed by an audience than in private— a difference that is often attributed to a heightened concern with appearing cooperative and reaping the associated benefits (e.g., reputation enhancement; Banerjee, 2002; Engelmann, Over, Herrmann, & Tomasello, 2013; Fujii, Takagishi, Koizumi & Okada, 2015; Leimgruber, Shaw, Santos, & Olson, 2012; Van Vugt & Hardy, 2010). What is less clear is whether concerns for reputation enhancement differ as a function of whether the recipient is an ingroup or outgroup member.

At least among adults, outgroup giving may be driven by qualitatively different motives than ingroup giving (Levine, Prosser, Evans & Reicher, 2005; Saucier, Miller, & Doucet, 2005; Van Leeuwen & Täuber, 2010). Specifically, outgroup giving may be more affected by the interests of the giver than those of the recipient, whereas ingroup giving may be more motivated by empathic concerns (i.e., concern for the well-being of recipients). In one study addressing this issue, Stürmer and colleagues (2005) found that although adults expressed feeling similar levels

of empathy for ingroup and outgroup members, empathy was a stronger predictor of helping when the recipient was an ingroup member, whereas feelings of attraction for the recipient more strongly predicted helping behavior when the recipient was an outgroup member (Stürmer, Snyder, & Omoto, 2005). Compatible with this, social signaling theories (Barclay, 2016; Fehr & Fischbacher, 2003; Gintis, Smith, & Bowles, 2001) and Stürmer's empathy X group membership moderation hypothesis (2005) also posit that individuals are more likely to act out of empathic concerns when helping ingroup members, and more likely to help outgroup members when it provides a benefit to the helper—e.g., by providing an opportunity to appear generous and thereby foster a favorable reputation (Batson, 1991; Silk & House, 2011). Compatible with this hypothesis, a study of Scottish adults found that telling participants that Scots are perceived as "mean" and "stingy" by the English motivated participants to behave more generously to outgroup members to restore their group's reputation (Hopkins et al., 2007). These findings, like those reviewed above, suggest that outgroup giving might stem from more strategic motives than ingroup giving. In the current work, we focus on the role of reputation concerns in driving giving behavior and investigate whether ingroup giving is also impacted by such concerns.

The claims regarding reputation management in adults raise the question of how reputation concerns emerge in childhood, and whether children's sensitivity to reputation concerns differ according to the group identity of a recipient. Like adults, children as young as five years of age exhibit a strong preference for ingroup members over outgroup individuals (Fehr et al., 2008; Nesdale, Griffith, Durkin, & Maass, 2005), whether these groups are familiar and meaningful or are so-called "minimal groups", established using arbitrary characteristics such as color or drawing ability (Dunham, Baron, & Carey, 2011; Nesdale & Flesser, 2001; Spielman, 2000; Tajfel, Billig, Bundy, & Flament, 1971). Not only do children have a more favorable attitude toward ingroup

members than outgroup members, but they also prioritize their ingroup when allocating resources (Fehr et al., 2008; Moore, 2009; Yu et al., 2016) and expect others to act in the same manner (Olson & Spelke, 2008), even if they believe that sharing equally with both groups is the nicer thing to do (DeJesus, Rhodes, & Kinzler, 2014). However, while there is considerable evidence that children favor their own group over outgroups, it remains unclear what motivates this differential treatment of group members.

Children as young as four years old are also motivated by a desire to make a favorable impression on others (Rapp, Engelmann, Herrmann & Tomasello, 2019; Fu, Heyman, Qian, Guo & Lee, 2015). For instance, 5-year-olds act more cooperatively when they believe that their behaviors will be made known to others (Leimgruber et al., 2012; Piazza, Bering, & Ingram, 2011) and share more when recipients are present than when they are absent (Engelmann, Herrmann, & Tomasello, 2012). Additionally, 6-year-olds act in ways to appear fair while unfairly favoring themselves (Shaw, Monitinari, Piovesan, Olson, Gino, & Norton, 2014). These findings suggest that children, like adults, are sensitive to social contexts and will alter their behavior to manage their reputations and present themselves more positively to others.

What is unknown is whether children and adults have differing reputation concerns depending on their relationship with the recipient. On the one hand, past research suggests that children are often more concerned with appearing "nice" to acquaintances and strangers than to close friends, and several studies show that children share equally or more with non-friends than friends when their behavior will be known to others (Buhrmester, Goldfarb, & Cantrell, 1992; Paulus, 2016). On the other hand, theories of human cooperation argue that individuals should be more concerned with how they are perceived by ingroup members than outgroup members, since they are more likely to have future interactions with, and information about their reputation

transmitted between, ingroup members (Mifune, Hashimoto, & Yamagishi, 2010; Yamagishi & Mifune, 2008). In line with this argument, one developmental study found that children gave more when observed by ingroup members than by outgroup members (Engelmann et al., 2013). However, while this literature suggests that children might be more concerned with presenting themselves positively to ingroup than outgroup members, it does not address the separate question of what motivates how they allocate resources to these groups and whether ingroup giving and outgroup giving are differentially affected by reputation concerns. One novel way to test this question is to examine how reputation concerns might affect children's behavior with ingroup and outgroup members when they are being evaluated by a neutral third party.

Critically, past studies find that understanding of prosocial motives changes in development, raising the possibility that children and adults reason differently about sharing with ingroup and outgroup members. For example, adults perceive reputation enhancement as an ulterior motive that diminishes the kindness of a prosocial act, whereas children younger than nine do not (Heyman, Barner, Heumann, & Schenck, 2013; Heyman, Fu, Barner, Zhishan, Zhou, & Lee, 2016). This may be because adults focus primarily on an actor's prosocial *intentions* when evaluating cooperative behavior (Critcher & Dunning, 2011; Ham & Vonk, 2011), and therefore discount acts of giving that are done for self-serving motives like enhancing their own reputation (Heyman et al., 2013; Padilla-Walker & Carlo, 2014). In contrast, children younger than nine judge acts of giving as *nicer* if they occur in contexts that enhance the reputation of the giver (Heyman et al., 2013; see also Butzin & Dozier, 1986). Thus, younger children differ from adults in how positively they view the efforts of others to enhance their own reputations, and in how they weigh these efforts when judging acts of ostensible generosity.

These previous studies suggest that young children may be willing to seek reputation benefits under conditions in which older children and adults are not. This raises the question of whether such concerns play differing roles when giving to ingroup and outgroup members, or if, for children, reputation concerns are always at play. Following the logic of previous studies, which operationalize reputation concerns by comparing differences between public vs. private giving (Engelmann et al., 2012; Haley & Fessler, 2005; Mifune et al., 2010; Rigdon, Ishii, Watabe, & Kitayama, 2009), Study 1 explored this question by testing the extent to which children give to members of different groups in the presence or absence of a neutral observer. Specifically, we asked whether children's reputational concerns are stronger when giving to outgroup members, or, if instead, they exhibit equal concern for self-presentation when giving to ingroup and outgroup members. Though additional factors, such as empathy, may affect children's giving toward different group members, we were most interested in the role reputation concerns might play in predicting children's ingroup and outgroup giving. Also, we assessed children's evaluations of ingroup and outgroup giving by a third party, to determine which form of giving they perceive as nicer.

We predicted that, consistent with prior literature, children would give more to ingroup members than to outgroup members (Fehr et al., 2008; Spielman, 2000, Yu et al., 2016), and that they would give more in public than in private (Buhrmester et al., 1992; Engelmann et al., 2012; Leimgruber et al., 2012). Less clear, however, was whether this preference for ingroup members would interact with children's reputation concerns. One possibility, compatible with the literature on sharing in adults, is that children's outgroup giving is more motivated by reputational concerns relative to ingroup giving. This would predict that the amount children share with outgroup members should be more affected by being observed than the amount they share with ingroup

members. On the other hand, if children believe that sharing with ingroup members is more beneficial for their reputation than sharing with outgroup members, then they may give more to their ingroup when observed than in private. Also possible is that children are mostly concerned with appearing cooperative and fair to an onlooker, and, hence, will share more equally with ingroup and outgroup members when observed than in private.

Regarding children's evaluations of third-party sharing, we didn't have a strong *a priori* hypothesis. It is possible that children perceive outgroup giving as nicer than ingroup giving since it entails helping an individual outside of one's immediate group. On the other hand, children may perceive outgroup giving as a violation of loyalty to one's own group, in which case we might expect children to perceive ingroup giving as nicer.

2. Study 1

In Study 1, we investigated children's relative concern for positive self-presentation when giving to ingroup and outgroup members, as well as their evaluations of other children's sharing with different group members. To establish group membership, we assigned children to arbitrarily defined groups using the minimal group paradigm (Dunham et al., 2011; Tajfel et al., 1971), and told them that the different groups were in competition with one another (based on previous findings that this strengthens ingroup preference; Spielman, 2000). We then tested children in two phases. In the first phase children performed a sticker allocation task in which they could share stickers with ingroup and outgroup members across trials. We manipulated whether the giving occurred privately or publicly (i.e., in which an observer stood by and watched how much children donated) between subjects, and in line with past work, we reasoned that private sharing is more indicative of a concern for the recipient, whereas public sharing is more indicative of a desire for positive self-presentation or concern for reputation management (Buhrmester et al., 1992;

Leimgruber et al., 2012; Heyman et al., 2013). Thus, as noted previously, we operationalized sensitivity to reputation management as the difference in amounts given when allocating resources publicly vs. privately (Haley & Fessler, 2005; Mifune et al., 2010; Rigdon, et al., 2009). In the second phase of the study, we asked children to judge the giving behaviors of fictitious children. All children completed the sticker allocation task prior to the evaluation task since previous studies find that children often adjust their sharing behaviors to align with those of their peers (House, Henrich, Sarnecka, & Silk, 2013; Ruggeri, Luan, Keller & Gummerum, 2017).

2.1 Method

2.1.1 Participants

We tested a total of 164 children (82 females), with 84 children in the Observer condition and 80 in the No Observer condition. This sample size was based on previous studies of children's minimal group sharing that used a similar design and task (Dunham et al., 2011; Engelmann et al., 2013), and a power analysis, which indicated that a sample size of 160 would yield medium power (d = .40), assuming an effect size of .25 and an alpha level of .05. Children's ages ranged from 5 to 9 years with 33 5-year-olds, 32 6-year-olds, 35 7-year-olds, 32 8-year-olds, and 32 9-year-olds; 82% were white, 15% were Asian American, 3% were identified as mixed/other by caregivers. Participants were students in a school district that primarily serves an upper middle-class population. The study received ethics approval from UC San Diego's Human Research Protections Program.

2.1.2 Procedures

There were two phases of the procedure. In the first phase children completed a sticker-sharing task, and in the second phase they completed an evaluation task, in which they evaluated the sharing of other children. At the start of the testing session, children randomly selected a green

or orange block hidden behind the experimenter's back to assign them to one of two possible minimal groups (the Green or Orange group). Blocks were surreptitiously manipulated to ensure that an equal number of children from each age group were assigned to each color group. Children were told that their group included other children who chose the same color block as them. They were then given a wristband matching the color of their group to wear for the duration of the task. Next, each participant heard four narratives describing a competitive relationship between the two groups (e.g., "The Green group *really* wants to win against the Orange group, and the Orange group *really* wants to win against the Green group", see Appendix A for narratives). The narratives emphasized competition between the groups without mention of rivalry over resources in order to prevent children from viewing the sticker-sharing task as a competition over resources.

Sticker-sharing task (Phase 1): All children completed the sticker-sharing task, in which they were asked to divide their stickers between themselves and either ingroup or outgroup members. We refer to ingroup and outgroup sharing as 'Group Sharing Type' and manipulated this as a within-subjects factor. The between-subjects factor was whether children were observed whilst completing the sharing-task; nearly half of the participants completed the task with an observer present (Observer condition) and the other half did so in the absence of an observer (No Observer condition).

Stimuli for the sticker-sharing task were six full-color head and shoulder photographs of white female children between the ages of 5 and 7 attached to manila envelopes. All recipients were female and of the same race in order to control for gender and race as extraneous variables in children's ingroup/outgroup sharing behavior, as previous work indicates that children favor some racial groups over others and prefer to give to recipients of their own gender, and to help prevent these effects from overriding children's minimal group preferences, since gender and race

are familiar, non-neutral social groups that children have prior experience with (Dunham et al., 2011). Controlling for gender allowed us to minimize the variability in our findings by analyzing effects due to the gender of participants, rather than the interaction of participant X recipient gender. Note that while it would be possible to double the study size to achieve the power needed to also analyze recipient gender (i.e., to an N of 328), we had no hypotheses regarding recipient gender and therefore did not pursue this question. Photographs of target children were those used by Dunham et al. (2011) and were rated by adults as being similar in attractiveness. They were edited using a photo editing software so that half of the children wore green t-shirts and the other half wore orange t-shirts.

On each trial, children were presented with an envelope featuring a picture of a target child from the contrasting outgroup (three trials) or their same group (three trials). Each time a picture of a target child was presented, the experimenter placed seven stickers on the table in a random arrangement such that stickers were not placed in rows or grouped together in sets. An unequal number of stickers was used so that children would have to choose between favoring themselves or a target recipient without defaulting to a fair distribution. This ensured that, if we observed a developmental shift in children's sharing behavior, this shift would not be primarily due to a change in children's sensitivity to fairness (which occurs between ages 6-8; Blake & McAuliffe, 2011), but rather a change in their sensitivity to group membership and intergroup biases. Children were told that the stickers belonged to them and they could distribute the stickers as they wished. They were led to believe that the envelopes would later be mailed to each target child and were told that the experimenter would not look inside of the envelopes. In each trial, children put the stickers that they wanted to donate inside of the envelope with the target child's picture on it, and the stickers that they wanted to keep in a separate envelope with their name on it. Children

completed seven trials in total including one practice trial. For counterbalancing purposes, there were two orders for the presentation of target pictures where one order was the reverse of the other.

There were two Observer conditions (Observer and No Observer) for the sticker-sharing task that differed between subjects. For both conditions, the experimenter placed a poster board between herself and the participant so that the experimenter's view of the participant was obstructed. In the Observer condition, a research assistant sat next to the child and maintained a neutral expression while watching the child complete the sticker-sharing trials. Children were told that the observer cared about how nice they were and was watching them to see how many stickers they shared. The presence of an evaluative observer was emphasized, since prior work shows that children younger than eight years old are insensitive to a passive adult observer and require verbal cues to understand that their actions are being watched and evaluated (Dutra et al., 2018; Fujii et al., 2015). In the No Observer condition children completed the trials without being watched by a research assistant.

Evaluation task (Phase 2): In Phase 2 of the study, children evaluated sharing scenarios involving third parties. Children were told that they would see images of how other children played a sticker game similar to the one they just completed—except now, they would decide if the children in the images were being nice or mean in sharing their stickers. Since we wanted to later assess how children's sharing behavior in Phase 1 might relate to their evaluations in Phase 2, we paralleled the evaluation task as closely as possible to children's experience in the sticker-sharing task with ingroup sharing and outgroup sharing manipulated within subjects, and whether or not an observer was shown manipulated between subjects.

Prior to beginning the evaluation task, examples of a child performing a "nice" act (cleaning up the classroom) and a "mean" act (pushing another child down on purpose) were read

to each participant. Children answered whether the actions were nice or mean to ensure that they understood the meaning of each term. Next, children saw eight PowerPoint slides, where each slide depicted scenarios of a target child sharing stickers with another child who was in her ingroup or outgroup in the presence or absence of a female observer. At the start of each trial, the experimenter narrated which color group the giver and recipient belonged to and how many stickers were shared (e.g., "This child is part of the Green group and she donated five of her stickers to this child in the Orange group"). Target children in the evaluation task were represented by pictures of 16 children between the ages of 5 to 7, with a mix of givers and recipients in the Orange and Green group.

Eight scenarios in total were presented to children in the evaluation task: Four scenarios consisted of the giver donating her stickers to an ingroup member and four scenarios consisted of the giver donating her stickers to an outgroup member, where the ratios of giving and keeping stickers were 7:0, 5:2, 2:5, 0:7 for both types of giving. We use the term 'Group Sharing Type' to refer to whether the giver shared with an ingroup or outgroup member, and the term 'Resource Allocation' to refer to the number of stickers donated by the giver to the recipient (e.g., 7, 5, 2, or 0 stickers). Two orders for the presentation of PowerPoint slides were used, where one order was the reverse of the other. For each trial, children judged whether the giver was nice or mean, and were probed to elaborate by answering "How nice?" or "How mean?" using a six-point pictorial Likert scale with three smiling and three frowning faces, adopted from Ng, Heyman, and Barner (2011) and transformed into values: 6 (very, very, nice) to 1 (very, very, mean) for data analysis.

2.2 Results

Our primary question in Phase 1 was whether children would be differently affected by reputation concerns when sharing with ingroup versus outgroup members. In Phase 2, we assessed

children's beliefs about different types of sharing to see whether they perceive ingroup or outgroup giving as nicer.

Sticker-sharing task (Phase 1): To examine how children shared their stickers with ingroup and outgroup members in the presence or absence of an observer, we constructed generalized linear mixed-effects regression models and compared a reduced model to fuller models.[1] No significant effects of gender were found, so it was not included in any of the models. The reduced model included Age and Group Sharing Type (ingroup, outgroup) as predictor variables, Participant/Group Sharing Type as a nested random effect, and percentage of stickers donated in a given trial as the response variable. The full, best fitting model included all predictor variables contained in the reduced model in addition to Observer condition (Observer, No Observer) and the interaction between Group Sharing Type and Observer condition.[2] Likelihood ratio tests indicated that the addition of each factor significantly improved the fit of the model in comparison to the reduced model (all *ps* < .03 for all tests), and a test of multicollinearity using variance inflation factor (VIF) revealed low collinearity among predictor variables (all VIFs < 2.25). The full model differed significantly from the basic model, $\chi^2(2) = 15.69$, $p < .001$ and a test of model fit indicated high goodness of fit ($R^2 = .73$). We found that Age predicted 28.17% unique variance, Group Sharing Type predicted 34.02%, Observer condition predicted 36.96%, and the interaction between Group Sharing Type and Observer condition predicted .85%. The degrees of freedom reported in this analysis were computed using the Satterthwaite approximation[3] and post hoc analyses were computed using Tukey's Honestly Significant Differences (HSD) with *p* set at .05.

[1] Analyses were performed in R 3.5.2 (http://www.r-project.org) using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2016).

[2] The final model specification was: Percentage of Stickers Shared ~ Group Sharing Type + Observer condition + Group Sharing Type*Observer condition + Age + (1 | Participant /Group Sharing Type).

[3] The Satterthwaite approximation is a formula used to calculate the approximate degrees of freedom for a linear combination of independent sample variances.

As expected (**Fig. 1.1**), we found a significant main effect of Observer condition ($F(1, 158) = 27.65$, $p < .001$), such that children who were observed donated significantly more stickers ($M = 54\%$, SE = 2.09), than children who were not observed (HSD, $p < .001$, $M = 38\%$, SE = 2.14). This finding is consistent with previous work showing that children are more generous when sharing in public than in private (Engelmann et al., 2012; Leimgruber et al., 2012). Also consistent with past work (Olson & Spelke, 2008; Tajfel & Turner, 2004; Yu et al., 2016), we found a main effect of Group Sharing Type ($F(1, 162) = 25.46$, $p < .001$), such that children donated significantly more stickers to ingroup members ($M = 50\%$, SE = 1.64) than to outgroup members (HSD, $p < .001$, $M = 43\%$, SE = 1.64). Most germane to the question of the present study, we did not find a significant interaction between Observer condition and Group Sharing Type ($F(1, 162) = .64$, $p = .426$), suggesting that children's giving to ingroup and outgroup members was equally affected by reputational concerns.

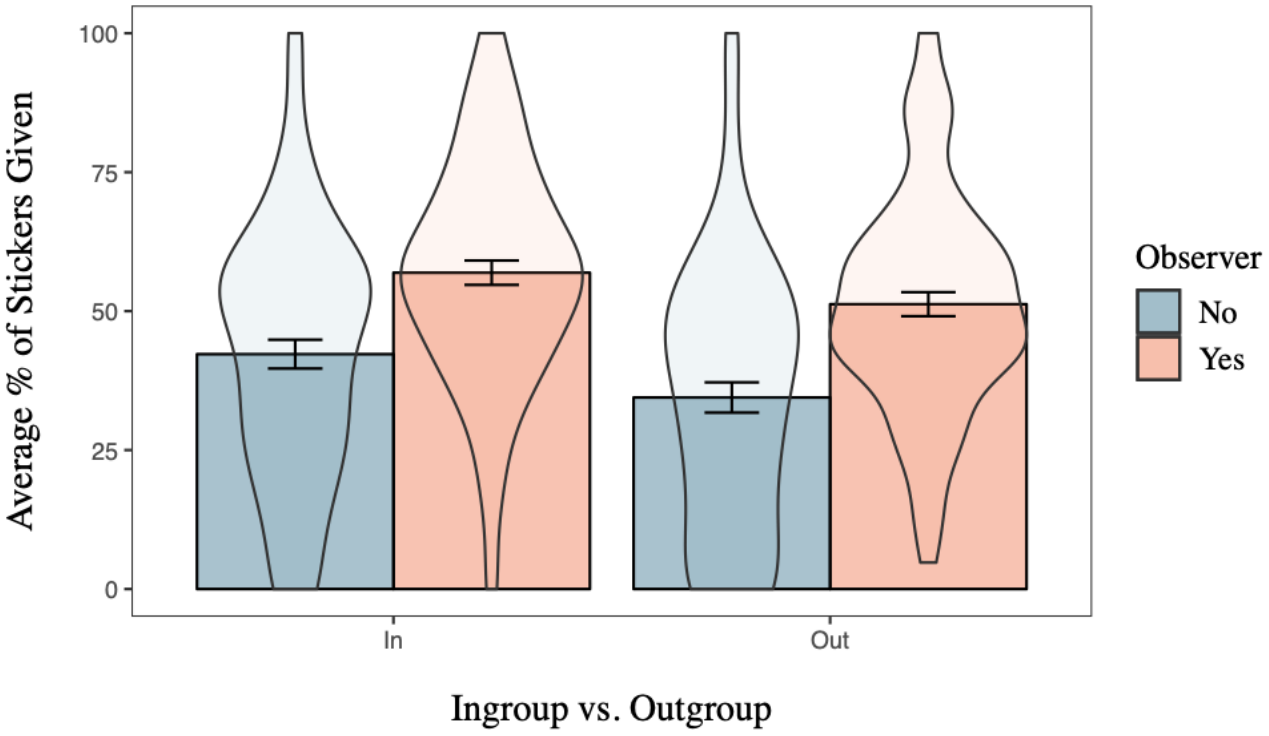

Figure 1.1. Percentage of stickers given by Observer condition (Observer, No Observer) and Group Sharing Type (ingroup, outgroup). Error bars indicate mean standard error, and violin plots show density distribution of responses.

Post-hoc analyses indicated a significant difference between the percentage of stickers shared by 5-year-olds (M = 35%, SE = 3.34) and 6-year-olds (HSD, p = .009, M = 51%, SE = 3.39), as well as between 5-year-olds and 8-year-olds (HSD, p < .001, M = 56%, SE = 3.39), demonstrating that younger children generally shared fewer stickers. Also, *post-hoc* tests revealed no interactions due to age, indicating that younger and older children were similarly sensitive to observer effects— a finding that parallels those of past studies (Aloise-Young, 1993; Banerjee, 2002; Piazza et al., 2011).

Evaluation task (Phase 2): To test whether children's sharing behavior in Phase 1 parallels their beliefs about ingroup and outgroup sharing, we assessed how children rated the niceness of another, fictitious child's sharing behavior in Phase 2, the evaluation task. We used the model

comparison method described above with linear regression models, rather than linear mixed-effects regression models, since children rated each scenario only once. No significant effects of gender were found, so this factor was not included in any of the models. Correlations computed prior to creating regression models revealed low collinearity among predictor variables (all VIFs = 1). The reduced model included Age and Group Sharing Type (ingroup, outgroup) as predictor variables, and average niceness ratings as the outcome variable. Likelihood ratio tests indicated that adding Resource Allocation to the reduced model significantly improved its goodness of fit ($p < .001$) but adding the interaction between Resource Allocation and Group Sharing Type did not ($p = .103$) and neither did adding Observer condition ($p = .156$), therefore, these variables were not included in the full model. The full model differed significantly from the reduced model, $\chi^2(3) = 694.47$, $p < .001$ and the overall model fit was $R^2 = .43$, indicating moderate goodness of fit.[4] We found that Age predicted .16% unique variance, Group Sharing Type predicted 3.13%, Resource Allocation predicted 39.75%, and 56.96% of the variance in children's niceness ratings was not explained by the model. We did not find any age effects, so have omitted participant age from our description of findings below.

Our analysis revealed a significant main effect of Group Sharing Type (see **Fig. 1.2**), such that children rated ingroup sharing ($M = 4.39$, SE = .07) as significantly nicer than outgroup sharing ($M = 3.75$, SE = .07), $F(1, 1330) = 69.23$, $p < .001$. This result is compatible with previous work showing that children endorse group cohesion and loyalty (Aboud, 2003; Abrams, Rutland, Ferrell, & Pelletier, 2008; Castelli, De Amicis, & Sherman, 2007; although see DeJesus et al., 2014 for one scenario in which children favor equal giving across groups when acts of ingroup and outgroup giving are explicitly contrasted). Our results also revealed a main effect of Resource Allocation

[4] The final model specification was: Average Niceness Ratings ~ Group Sharing Type + Resource Allocation + Age.

($F(3, 1330) = 291.32$, $p < .001$), such that children rated sharing 7 stickers as nicer than sharing 5, 2, or 0 stickers (all HSDs, $p < .001$), sharing 5 stickers as nicer than sharing 2 or 0 stickers (all HSDs, $p < .001$), and sharing 2 stickers as nicer than sharing 0 stickers (HSD, $p < .001$). As mentioned above, the interaction between Resource Allocation and Group Sharing Type was not significant and not included in the model. Consistent with DeJesus et al. (2014), we also found that children's own sharing behavior in Phase 1 was unrelated to their evaluations in Phase 2 (see Appendix B for details).

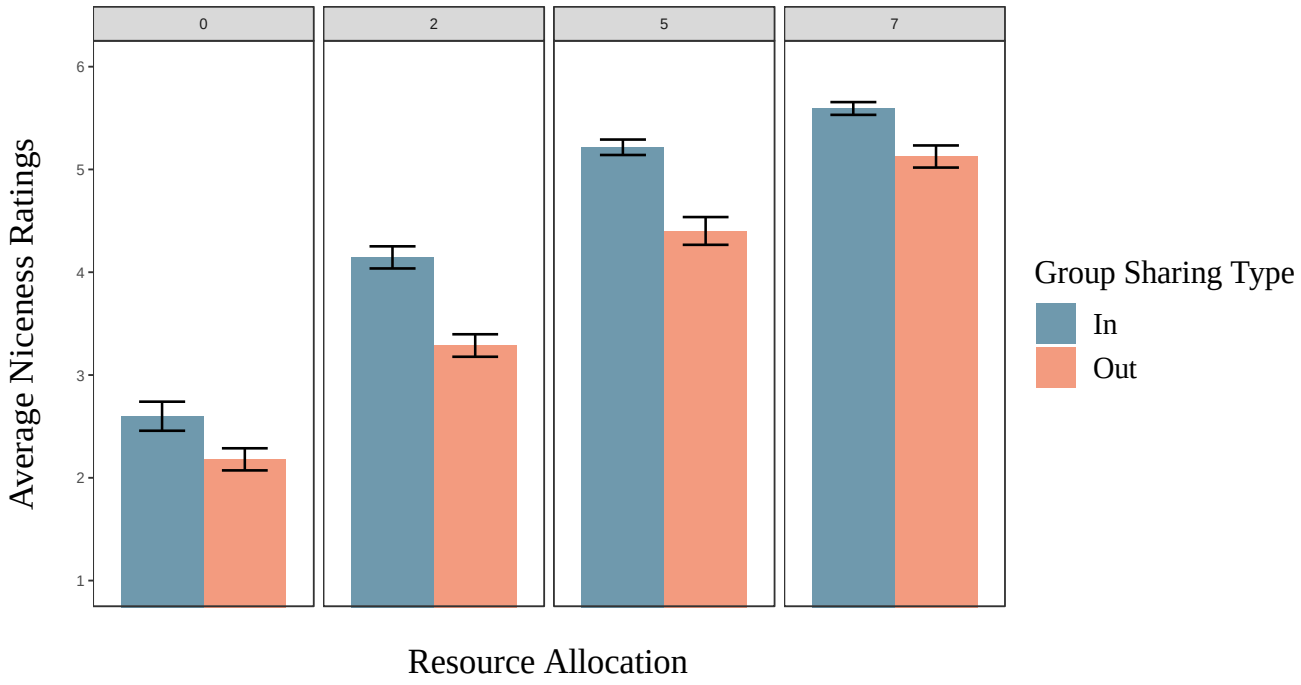

Figure 1.2. Children's average niceness ratings (1 = *very, very, mean*, 6 = *very, very, nice*) by Resource Allocation and Sharing Type (ingroup giving, outgroup giving).

2.3 Discussion

There were two main findings in Study 1. First, we found that children were equally concerned with their reputation when giving to outgroup and ingroup members; children gave more stickers overall when observed than not observed, but to the same degree regardless of the

recipient's group identity. Second, we found that children evaluated ingroup giving as nicer than outgroup giving for all instances. Both of these results raise interesting questions regarding how children view reputation management. The first result suggests that, although children view outgroup giving as an opportunity to enhance their reputation, they also view ingroup giving in this way. The second result, that children judge ingroup giving to be nicer than outgroup giving, is compatible with how they generally behaved when sharing their own stickers.

Our finding that children's ingroup and outgroup giving increased to a similar extent under observation is somewhat surprising in light of past work showing that children strategically manage their reputations depending on whether the recipient is an ingroup or outgroup member (Buhrmester et al., 1992; Engelmann et al., 2013; Paulus, 2016). It is also unexpected in the context of adult work suggesting that many forms of outgroup sharing, but not ingroup giving, are strategic in nature and arise from the desire to appear generous, or to signal status and power to outgroup members (Halabi, Dovidio, & Nadler, 2008; Klein & Azzi; 2001; Van Leeuwen & Täuber, 2010). One possible explanation is that children have unique beliefs relative to adults about ingroup and outgroup sharing. For instance, children may differ from adults in their understanding of the strategic benefits associated with outgroup giving, or how they interpret this information. However, no previous study has asked this question in adults using methods similar to those in the developmental literature. To address this issue, we used a procedure similar to the one used in Study 1 with adults.

3. Study 2

Study 1 found that, when children were asked to share with ingroup and outgroup members, they gave more when observed – a result which previous studies have interpreted as evidence for sensitivity to reputational concerns (Bateson, Nettle & Roberts, 2006; Leimgruber et al., 2012;

Van Vugt & Hardy, 2010). However, surprisingly, this effect of being observed by a neutral onlooker did not differ when children gave to ingroup versus outgroup members, suggesting that, at least in children, ingroup giving may be just as impacted by reputational concerns as outgroup giving. These results appear to be inconsistent with arguments that adults give to ingroup members chiefly out of empathic concern, and to outgroup members more out of self-serving concerns, such as impression management (Aboud & Levy, 2000; Stürmer et al., 2005, 2006). However, while these studies found more empathy-driven behavior towards ingroups, most have not tested whether reputational concerns are less important when giving to ingroup members. This is a problem because while empathy may explain more variance in ingroup giving – whether in adults or in children – it remains possible that ingroup giving is also impacted by reputation concerns, and perhaps to the same degree as for outgroup giving. Although the current study does not test empathy, our results bear on the extent to which it alone determines ingroup giving, and more importantly our results speak to alternative mechanisms.

Given these considerations, Study 2 examined whether reputational concerns differ across ingroup and outgroup giving in adults in a minimal group paradigm that paralleled the method of Study 1. Also of interest was whether adults, like children, perceive ingroup giving as nicer. Given that the tasks used in Study 2 were adapted from methods designed to test children, we also used a survey to probe adults' beliefs regarding ingroup and outgroup sharing in a format that is more familiar for their age group.

3.1 Method

3.1.1 Participants

Adult participants were undergraduate students recruited from the University of California San Diego's subject pool, who participated in exchange for course credit. A total of 80 adults (42

females) were tested, with 40 in the Observer condition and 40 in the No Observer condition. Participants' ages ranged from 18 to 29 years (M = 20.54, SD = 2.05); 36% self-identified as white, 30% self-identified as Asian, 24% self-identified as Hispanic, 5% self-identified as African American, 2.5% self-identified as Middle Eastern, and 2.5% self-identified as mixed/other. Participants were assigned to the Green and Orange group equally across genders and Observer conditions.

3.1.2 Procedures

Adult tasks paralleled those of Study 1, but were adapted to be realistic and convincing for an older age group. Participants completed three phases: a resource allocation task in Phase 1, an evaluation task in Phase 2, and a survey gauging their beliefs about ingroup and outgroup sharing in Phase 3. To make the sharing task more convincing, adult participants were led to believe that they were allocating resources (tokens) with ingroup and outgroup members in real time. To make this more plausible, participants were recruited in groups and saw other participants enter and leave the waiting room at the same time as them. To be sure that participants ascribed value to the tokens, they were told that each token would serve as an entry to a raffle that would take place weeks after the experiment ended. As in Study 1, Observer condition (Observer, No Observer) was a between-subjects factor, and Group Sharing Type (ingroup recipient, outgroup recipient) was manipulated within-subjects across trials.

At the start of the first task (Phase 1), the experimenter informed participants that they would take part in a study that was previously conducted with children. Participants were told that they would be part of the Orange or the Green group. As in Study 1, they were told that the two groups were in competition with one another and selected a green or orange block from a bag. Blocks were manipulated so that an equal number of male and female participants in each Observer

condition were assigned to each color group. Participants then wore a t-shirt matching their color group for the duration of the task.

A research assistant took a photo of the participant wearing the colored t-shirt and printed and pasted two colored copies of the participant's photo onto manila envelopes. Participants were told that they would see photos of other participants who were playing the token game and their envelope would be shared with these players. The research assistant returned to the testing room with two piles of manila envelopes. Both piles had the participant's envelope with their photo on top in direct view of the participant, and envelopes with pictures of ingroup and outgroup members underneath it. The research assistant left one pile of envelopes with the experimenter and then left the room after informing the participant that the other pile would be given to another player down the hall. For participants in the Observer condition, the research assistant returned and sat next to the participant to observe them during the token-sharing task. For both the Observer and No Observer conditions, the experimenter placed a privacy board between herself and the participant during each trial.

Token-sharing task (Phase 1): In Phase 1, all participants completed a token-sharing task similar to the sticker-sharing task in Study 1. Half of the participants completed the task with an observer present and the other half completed the task without an observer. The observer manipulation was done in a slightly different manner with adults than with children in order to prevent participants from becoming suspicious of this manipulation. Specifically, with adults the experimenter only mentioned the observer once, rather than repeatedly reminding the participant of the observer's presence.

Instead of stickers, tokens were used as the resource. Participants were told that tokens would serve as entries to a raffle that would occur in the coming weeks. To prevent participants

from viewing the task as a zero-sum game, they were told that the more people who played the game, and the more tokens there were, the more prizes there would be. Experimenters tested two participants at the same time in separate but adjacent rooms so that participants would believe that they were playing the token game in real time with others.

Similar to Study 1, there were six trials in total for the token-sharing task, and, at the start of each trial, participants were given seven tokens to distribute between themselves and a target player. Stimuli for the task were six full-color head and shoulder photographs of white females between the ages of 20 and 33 wearing orange or green t-shirts pasted onto the front of manila envelopes. We chose all female targets to mimic the sharing task in Study 1. Though we were concerned that participants might find it odd that they were presented with all female targets, none of the participants mentioned this during or after the task during debriefing. All participants were presented with the same six target recipients, but the order of targets was counterbalanced as well as Group Sharing Type (ingroup or outgroup recipient), such that half of the participants saw a target wearing an orange shirt while the other half saw the same target wearing a green shirt.

Evaluation task (Phase 2): Phase 2 paralleled the evaluation task of Study 1 with children. Participants were told that the experimenter had secured permission from past players to disclose how they distributed their tokens in Phase 1. Next, participants were presented with eight PowerPoint slides: Four of the slides portrayed ingroup sharing and four slides portrayed outgroup sharing. As in Study 1, the slides depicted various token allocations (7:0, 5:2, 2:5, 0:7) between the giver and recipient. Target givers were represented in photographs of white females between the ages of 20-33 and target recipients were portrayed as orange or green colored silhouettes. Participants who were in the Observer condition in Phase 1 saw a picture of a female observer on each PowerPoint slide and were told that players completed the task with the observer present.

Two different orders for the presentation of PowerPoint slides were used, such that one order was the reverse of the other. For each trial, participants evaluated the niceness of the giver using the same six-point pictorial Likert-type scale as used with children in Study 1.

Sharing survey (Phase 3): After Phase 1 and 2 of the study, we asked participants to complete a sharing survey. The goals of the survey were to (a) test whether participants' self-reported sharing behavior in the token-sharing task aligned with their actual sharing behavior, (b) gauge participants' explicit evaluations of ingroup and outgroup sharing, and (c) assess their reasons for why they might believe ingroup or outgroup giving is nicer. In the survey, participants were asked to report their own sharing behavior in the token-sharing task by selecting one of the four options: (1) gave more to people in my own color group, (2) gave the same to people in each color group, (3) gave more to people in the other color group, and (4) don't remember. No participant responded that they did not remember how they shared their tokens. To test which scenario participants perceived as nicer, the next item participants answered was: "Would you think someone is nicer if they gave more tokens to their own color group or to the other color group?" and were given two response options to choose from: (1) own group and (2) other group. This question was followed by a free response question: "Why do you think so?". Due to time constraints, not all participants could stay to complete the survey, so we collected responses from 60 participants (30 females) and eliminated 1 participant from the analysis due to incomplete answers.

3.2 Results

Token-sharing task (Phase 1): Similar to Study 1, we examined how participants shared their tokens with ingroup and outgroup members when observed and not observed. In order to allow a direct comparison of results across studies, we report results from a general linear mixed-

effects regression model based on the final model structure used in Study 1.[5] Tests run prior to creating regression models revealed low collinearity among predictor variables (all VIFs < 2.3). No significant effect of gender was observed, and following the model used in Study 1, it was not included. Model comparisons indicated no difference between a reduced model with only Group Sharing Type as a predictor variable and the full model ($\chi^2(7) = .41$, $p = .813$). The overall model fit was $R^2 = .89$, indicating substantial goodness of fit. An analysis of each predictor variable indicated that Group Sharing Type predicted 99.02% unique variance, Observer condition predicted .23% and the interaction between Group Sharing Type and Observer condition predicted .75%.

As in Study 1, we found a significant effect of Group Sharing Type, with participants donating significantly more tokens to ingroup members ($M = 49\%$, $SD = .23$) than to outgroup members ($M = 33\%$, $SD = .24$), $F(1, 78.58) = 40.55$, $p < .001$. However, unlike our finding for children in Study 1, the presence of an observer did not appear to affect adults' sharing (see **Fig. 1.3**). Participants who were observed during the task ($M = 41\%$, $SD = .26$) shared similarly to participants who were not observed ($M = 42\%$, $SD = .24$), $F(1, 79.08) = .09$, $p = .759$. Also, as in Study 1, we did not find an interaction between Observer condition and Group Sharing Type, $F(1, 78.58) = .31$, $p = .581$. Overall, while we found a strong and significant effect of Group Sharing Type on adult giving, we did not find evidence that outgroup giving was influenced more by reputational concerns than ingroup giving. More generally, in this paradigm, adults appeared to be less concerned than children with reputation management, despite exhibiting other well-attested effects, such as a robust ingroup preference.

[5] The final model specification was: Percentage of Tokens Shared ~ Group Sharing Type + Observer condition + Group Sharing Type*Observer condition + (1 | Participant /Group Sharing Type).

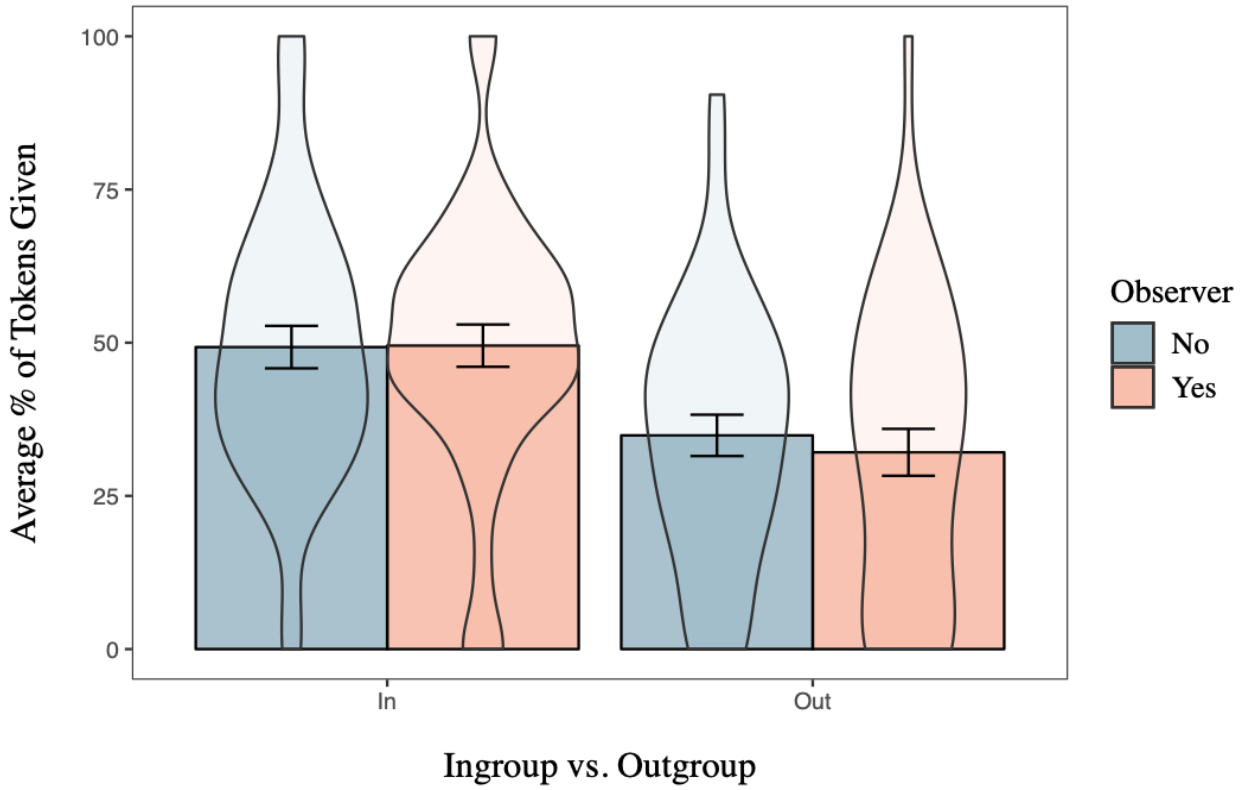

Figure 1.3. Percentage of tokens given by Observer condition (Observer, No Observer) and Group Sharing Type (ingroup, outgroup). Error bars indicate mean standard error, and violin plots show density distribution of responses.

Evaluation task (Phase 2): Similar to Study 1, participants' niceness ratings (1 = very, very, mean…6 = very, very, nice) in the evaluation task were analyzed using model comparisons between reduced and fuller models.[6] Likelihood ratio tests indicated that adding Resource Allocation to the reduced model significantly increased its goodness of fit $\chi^2(3) = 764.8$, $p < .001$, but adding the interaction between Resource Allocation and Group Sharing Type did not ($p = .929$) nor did adding Observer condition ($p = .978$). Correlations computed prior to generating the full regression model revealed low collinearity among predictor variables (all VIFs = 1). In the final model, Resource Allocation predicted 69.72% unique variance, Group Sharing Type predicted

[6] The final model specification was: Average Niceness Ratings ~ Group Sharing Type + Resource Allocation.

.01%, and 30.27% of the variance in participants' niceness ratings was not explained by the model. The overall model fit was R^2 = .70, indicating moderately high goodness of fit.

Unlike findings from Study 1 with children, no effect of Group Sharing Type was observed among adults—participants rated scenarios of ingroup sharing (M = 4.27, SE = .05) similarly to scenarios of outgroup sharing (M = 4.29, SE = .05), $F(1, 635) = .13$, $p = .720$. Similar to Study 1, and as expected, a significant effect of Resource Allocation was found ($F(3, 635) = 377.43$, $p < .001$), such that participants rated sharing 7 tokens as nicer than sharing 5, 2, or 0 tokens (all HSDs, $p < .001$), sharing 5 tokens as nicer than sharing 2 or 0 tokens (all HSDs, $p < .001$), and sharing 2 tokens as nicer than sharing 0 tokens (HSD, $p < .001$) (see **Fig. 1.4**). Thus, we found a strong and significant effect of Resource Allocation on niceness ratings, but no effect of Observer condition, and no preference for ingroup vs. outgroup giving.

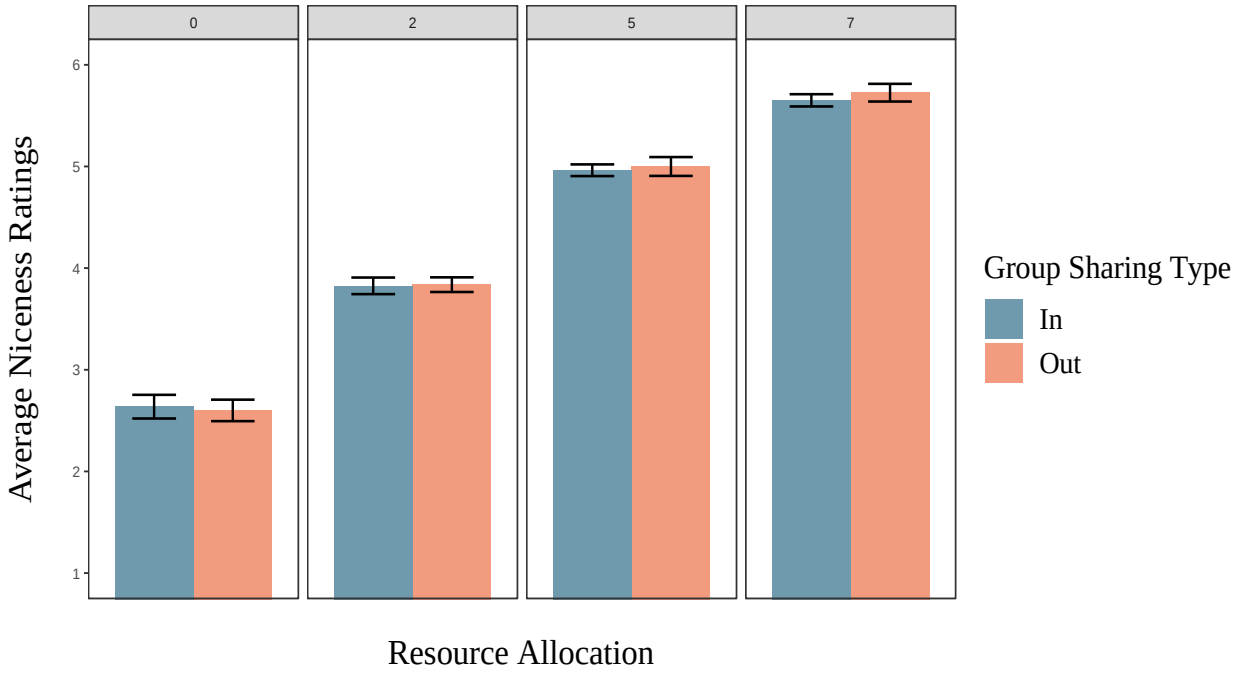

Figure 1.4. Adults' average niceness ratings (1 = *very, very, mean*, 6 = *very, very, nice*) by Resource Allocation and Sharing Type (ingroup giving, outgroup giving).

Sharing Survey (Phase 3): To further explore the question of whether adults perceive ingroup giving as nicer than outgroup giving, we analyzed their responses to the sharing survey. Results are reported in **Table 1.1**. A total of 59 participants (29 females) were included in the analysis. Of those participants 52% reported that they gave more tokens to ingroup members than outgroup members, while 46% reported that they distributed their tokens equally between groups, and 2% reported that they gave more tokens to outgroup members. However, interestingly, 39% of participants behaved differently from their self-reported allocations, with most (57%) incorrectly reporting that they distributed their tokens equally between groups, whereas they actually exhibited an ingroup bias. Also, when asked to explicitly contrast ingroup and outgroup giving, 81% of respondents indicated that giving to outgroup members is nicer than giving to ingroup members—a result that is compatible with participants' exaggeration of the extent to which they shared with outgroup members. The most common explanation that participants gave for this judgment was that it is more selfless to give to another group since it decreases the chances of one's own group benefitting and thus requires greater sacrifice. Many other participants mentioned that it is more difficult, and less common, to give to people who share no similarities or direct relation with you, thereby making outgroup giving the more generous act.

Table 1.1. Adult Evaluations of Ingroup and Outgroup Giving Organized by Self-reported Token Distributions

How did you share your tokens?	Which is nicer?			
	Ingroup giving		Outgroup giving	
	Frequency	Percent	Frequency	Percent
Gave more to ingroup	8	72.7	23	47.92
Distributed equally between groups	3	27.3	24	50
Gave more to outgroup	0	0	1	2.08
Total	11	100	48	100

3.3 Discussion

In Study 2 we tested whether reputation concerns motivate giving to outgroup and ingroup members in adults, and how adults evaluate these different forms of giving when judging other people. We found that adults, like children, showed comparable observer effects when giving to both outgroup and ingroup members. However, adults were less affected by the presence of an observer and showed a stronger bias to share with ingroup members than children. Further, in the evaluation task, adults based their niceness ratings on the number of tokens shared rather than the type of sharing that occurred (ingroup vs. outgroup sharing), which is likely why we observed no differences between how adults rated scenarios where 0 stickers were shared with ingroup versus outgroup members. Another possibility is that the format of the evaluation task was better suited for gauging children's moral evaluations rather than those of adults, which is why we included an additional survey in Study 2. Thus, the difference we observed between children and adults in their niceness perceptions of ingroup and outgroup giving could be partly due to having only asked adults to make a forced choice between the two forms of giving. Interestingly, despite giving more to ingroup members, adults subsequently reported in the survey that they shared equally across groups and indicated that giving to outgroup members is nicer than giving to ingroup members when asked about this explicitly – a possible explanation for why they under-reported their own ingroup bias.

4. General Discussion

Theories of prosocial behavior in adults argue that ingroup giving is guided more by empathic concerns, and less by concerns related to reputation, relative to outgroup giving (Aboud & Levy, 2000; Stürmer et al., 2005, 2006). Based on this literature, we asked whether children might reason differently about reputation concerns for ingroup and outgroup giving. Following

previous studies, Study 1 operationalized reputation concerns by asking children to share resources in a minimal-group paradigm, either while being observed or privately. We also asked children to judge the giving behaviors of others. Surprisingly, we found that, while children shared more resources when being observed, this effect was just as large for ingroup giving as for outgroup giving. In addition, we found that children judged others to be nicer when they shared with ingroup members as opposed to outgroup members. In Study 2, we found that adults also showed a clear ingroup bias when sharing resources between groups, but, unlike children, they exhibited a general indifference to observer effects. Also, unlike children, adults identified outgroup giving as nicer than ingroup giving when explicitly asked to make a choice between the two. Together, these results suggest that, at least in the context of giving to members of minimal groups, reputation concerns do not differ significantly between ingroup and outgroup giving in children or adults.

In Study 1, we found that, though children generally favor members of their ingroup, when a neutral observer is present, children's sharing behavior increases similarly towards ingroup and outgroup members. This finding is notable because, although many studies have examined cases where children favor their ingroup and disfavor outgroup members (e.g., Aboud, 2003; Spielman, 2000; Tajfel & Turner, 2004), few have explored how children manage their reputations when interacting with different group members and when observed by a neutral onlooker. Here, we provide evidence that children are equally sensitive to reputation concerns when sharing with outgroup and ingroup members.

Although children in our study were sensitive to the presence of an observer when sharing their own resources, we found that their evaluations of third-party giving did not differ when conducted in the presence or absence of an observer. One reason that children were insensitive to this factor may be that it was less salient than in previous work that reported observer effects. In

our study, children were assigned to evaluate givers who were observed or not observed between subjects (which corresponded with whether the child was themselves observed when performing the sharing task). In contrast, other studies (e.g., Heyman et al., 2013; 2016) directly contrasted public versus private giving within subjects, making explicit that the givers shared either publicly or privately with the intention of revealing or concealing their act of sharing. Thus, it is possible that children in our study were not sensitive to the observer condition when making their evaluations.

Relatedly, our finding that observer effects were less robust in adults than in children may be due to emphasizing the presence of an observer more to children than adults in an effort to make the task instructions age appropriate for each group. For example, stronger observer effects might have been found if we had reminded adults that they were being watched and evaluated at the start of each trial. Also possible is that children generally have greater concerns with how they will be judged by others (Ruggeri et al., 2017; Wolf, Bazargani, Kilford, Dumontheil, Blakemore, 2015). Finally, our study is not unique in failing to find observer effects in adults. Although such effects have been found in some previous studies (Bateson et al., 2006; Haley & Fessler, 2005; Powell, Roberts, & Nettle, 2012), others do not find such effects (Carbon & Hesslinger, 2011; Lamba & Mace, 2010; Raihani & Bshary, 2012), suggesting that they are not robust across different experimental contexts (Fehr & Schneider, 2010; Rigdon et al., 2009).

In addition to examining sharing behavior, Study 1 examined children's evaluations of how others share resources. Here, we found that children rated ingroup sharing as significantly nicer than outgroup sharing. This was in contrast to the evaluations of adults, who viewed outgroup giving as nicer, sometimes noting that giving to outgroup members is less likely to be driven by self-serving ulterior motives. We considered two possible interpretations of this effect. First,

children may evaluate ingroup giving as nicer because, unlike adults, they fail to discount self-serving motives when judging other people's giving (Heyman et al., 2013), and therefore fail to discount ingroup giving. A second possibility is that children view ingroup giving as nicer because they feel a strong obligation to their group. Compatible with this possibility, past studies show that children expect individuals to prioritize ingroup members over outgroup members (DeJesus et al., 2014, Olson & Spelke, 2008), and evaluate loyal group members as nicer, more likeable, and more trustworthy than disloyal group members (Castelli et al., 2007; Misch, Over, & Carpenter, 2014).

Unlike children, when adults in Study 2 were asked to judge other people's acts of giving in the evaluation task, they based their judgments on the overall quantity of resources shared and did not rate ingroup giving differently from outgroup giving. However, when asked to directly compare the two forms of giving in a follow-up survey, adults identified outgroup giving as nicer than ingroup giving. Adults also exaggerated the extent to which they gave to the two groups equitably. One possibility is that adults forgot how they distributed their tokens across all trials. Another possibility is that adults knowingly exaggerated how much they gave to outgroup members in order to present a favorable image of themselves: participants seemed to be aware that other individuals also favor their own group when sharing resources, since many expressed in survey responses that outgroup giving is nicer precisely because it is more unexpected than ingroup giving. A third possibility is that adults were biased to recall their decisions or actions in a way that maintains their desired self-views (e.g., I am a fair and generous person; Alicke & Sedikides, 2009; Guenther & Alicke, 2008).

This returns us to the question of why adults think that it is nicer to direct acts of giving to individuals who are not in one's own group? One possibility is that this effect is due to a prior expectation that people should share more with ingroup members, making it exceptionally nice to

give to members of an outgroup. For example, if a parent were to throw their child a birthday party, we might not perceive this act as unusually generous, whereas if that same parent funded a birthday party for a child that was not their own, we might view this act as more deserving of praise, since it exceeds the expectations that we have for that giver-recipient relationship. Similarly, we might negatively judge a parent who throws birthday parties for other children, but not their own, because we expect that parents should first fulfill their obligation to their own child before extending their generosity to an outgroup member. Compatible with this line of reasoning, research regarding charitable giving and aid indicates that adults feel an obligation to first help domestic victims, who they perceive as belonging to their ingroup, before extending help to international victims, who they perceive as outgroup members (Cuddy, Rock, and Norton, 2007; Kogut & Ritov, 2007). The statistics for charitable giving reflect these group differences: Americans donated nearly 4 times more to Hurricane Katrina efforts than to Indian Ocean tsunami relief, though the death toll for the tsunami was estimated to be 190 times more than that of Hurricane Katrina (Kost, Tran, Tuntideelert, Kulrattanamaneeporn, & Peungposop, 2006), and in 2018 the bulk of private charity in the U.S. was directed to local causes, with only 5% of total donations going to international causes (Giving USA Foundation, 2019). In the current study, it is possible that participants rated outgroup giving as nicer on the assumption that anyone who shares with outgroup members has already given to their own group to some extent, thus making that person exceptionally generous.

Our findings leave open a number of important questions. First, while our results indicate that children are concerned with reputation whether when sharing with ingroup or outgroup members, it is likely that their motives for each form of giving nevertheless differ importantly, and that empathic concerns may play a greater role in ingroup giving than in outgroup giving – a possibility that we did not directly test in this study. Second, we cannot rule out the possibility that

children in the Observer condition gave more resources out of a general concern with wanting to behave nicely rather than concerns with being directly evaluated (although see Fujii et al., 2015 for an account where young children's sharing behavior was not affected by indirect cues of evaluation). Third, we used a neutral observer in our experimental design since we were most interested in the relationship between reputation concerns more generally and their effect on ingroup and outgroup giving; thus, our findings speak only to sharing patterns that occur in the presence of a neutral audience. It is possible that we would have found a stronger observer effect if the group membership of the observer (e.g., Engelmann et al., 2013) or the observer's relationship to the participant (e.g., Buhrmester et al., 1992) had been manipulated. Finally, further work is needed to investigate how our results generalize to different targets and populations. We used only female targets who were white, so it will be important to examine how our results generalize to male targets and other races. It is also unclear how our results will generalize to familiar social or cultural groups (Nesdale et al., 2005; Simon, Stürmer, & Steffens, 2000) or to close friends and family members versus strangers (Buhrmester et al., 1992; Costin & Jones, 1992; Moore, 2009; Olson & Spelke, 2008).

The current study is the first to show that children are equally affected by reputation concerns when giving to ingroup and outgroup members. Here we found that even the youngest children shared more when the strategic, and potentially selfish opportunity for reputation enhancement was present. This raises the question of whether activating more prosocial motives, such as empathic concern, could encourage children to behave more cooperatively in an intergroup context. Our findings are also the first to show that children, unlike adults, evaluate ingroup sharing as nicer than outgroup sharing, indicating a developmental shift between childhood and adulthood in how individuals reason about ingroup and outgroup giving.

Acknowledgments

We thank Del Mar Heights Elementary School for their involvement with recruiting participants, and Miriam Rubenson, Sophie Standen, Nestor Tulagan, Vennisia Mo, Vaishnavi Nandanam, Sydney Ghazarian and Kaitlyn Peck for their help in testing participants. We also thank Tim Lew, Katie Wagner, and Rose Schneider for their help with data analysis.

Chapter 1, in full, is a reprint of the material as it appears in: Yazdi, H., Heyman, G.D., & Barner, D. (2020) Children are sensitive to reputation when giving to both ingroup and outgroup members. Journal of Experimental Child Psychology, 194, 104814. The dissertation author was the primary investigator and author of this paper.

References

Aboud, F. E. (2003). The formation of ingroup favoritism and outgroup prejudice in young children: Are they distinct attitudes?. *Developmental Psychology*, *39*, 48-60. doi:10.1037/0012-1649.39.1.48.

Aboud, F. E., & Levy, S. R. (2000). Interventions to reduce prejudice and discrimination in children and adolescents. In S. Oskamp (Ed.), *"The Claremont Symposium on Applied Social Psychology" Reducing prejudice and discrimination* (pp. 269-293). Mahwah, NJ: Lawrence Erlbaum Associates.

Abrams, D., Rutland, A., Ferrell, J., & Pelletier, J. (2008). Children's judgments of disloyal and immoral peer behavior: Subjective group dynamics in minimal intergroup contexts. *Child Development, 79*, 444-461. doi:10.1111/j.1467-8624.2007.01135.x.

Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology*, *20*, 1-48. doi:10.1080/10463280802613866.

Aloise-Young, P.A. (1993). The development of self-presentation: Self-promotion in 6-to-10-year-old children. *Social Cognition*, *11*, 201-222. doi:10.1521/soco.1993.11.2.201

Banerjee, R. (2002). Audience effects on self-presentation in childhood. *Social Development, 11*, 487-507. doi:10.1111/1467-9507.00212.

Barclay, P. (2016). Biological markets and the effects of partner choice on cooperation and friendship. *Current Opinion in Psychology*, *7*, 33-38. doi:10.1016/j.copsyc.2015.07.012.

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, *2*, 412-414. doi:10.1098/rsbl.2006.0509

Batson, C.D. (1991). The altruism question: Toward a social-psychological answer. Hillsdale, NJ: Erlbaum.

Batson, C. D. (1994). Why act for the public good? Four answers. *Personality and Social Psychology Bulletin*, *20*, 603-610. doi:10.1177/0146167294205016.

Bernhard, H., Fischbacher, U. & Fehr, E. (2006). Parochial altruism in humans. *Nature*, *442*, 912-915. doi:10.1038/nature04981.

Blake, P. R., & McAuliffe, K. (2011). "I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition*, 120, 215-224. doi:10.1016/j.cognition.2011.04.006.

Bock, G. W., & Kim, Y. G. (2002). Breaking the myths of rewards: An exploratory study of attitudes about knowledge sharing. *Information Resources Management Journal (IRMJ)*, *15*, 14-21. doi:10.4018.irmj.2002040102.

Buhrmester, D., Goldfarb, J., & Cantrell, D. (1992). Self- presentation when sharing with friends and non-friends. *Journal of Early Adolescence, 12,* 61-79. doi:10.1177/0272431692012001004.

Butzin, C. A., & Dozier, M. (1986). Children's use of ulterior motive information. *Child Development,* 1375-1385. doi:10.2307/1130417.

Carbon, C. C., & Hesslinger, V. M. (2011). Bateson et al.'s (2006) cues-of-being-watched paradigm revisited. *Swiss Journal of Psychology, 70,* 203-210. doi:10.1024/1421-0185/a000058.

Castelli, L., De Amicis, L., & Sherman, S. J. (2007). The loyal member effect: On the preference for ingroup members who engage in exclusive relations with the ingroup. *Developmental Psychology, 43,* 1347-1359. doi:10.1037/0012-1649.43.6.1347.

Costin, S. E., & Jones, D. C. (1992). Friendship as a facilitator of emotional responsiveness and prosocial interventions among young children. *Developmental Psychology, 28,* 941-947. doi:10.1037/0012-1649.28.5.941.

Critcher, C. R., & Dunning, D. (2011). No good deed goes unquestioned: Cynical reconstruals maintain belief in the power of self-interest. *Journal of Experimental Social Psychology, 47,* 1207-1213. doi:10.1016/j.jesp.2011.05.001.

Cuddy, A. J., Rock, M. S., & Norton, M. I. (2007). Aid in the aftermath of Hurricane Katrina: Inferences of secondary emotions and intergroup helping. *Group Processes & Intergroup Relations, 10,* 107-118. doi:10.1177/1368430207071344.

DeJesus, J.M., Rhodes, M. & Kinzler, K.D. (2014). Evaluations versus expectations: Children's divergent beliefs about resource distribution. *Cognitive Science, 38,* 178-193. doi:10.1111/cogs.12093.

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in childhood. *Child Development, 82,* 793-811. doi:10.1111/j.1467-8624.2011.01577.x.

Dutra, N. B., Boccardi, N. C., Silva, P. R. R., de Oliveira Siqueira, J., Hattori, W. T., Yamamoto, M. E., & de Alencar, A. I. (2018). Adult criticism and vigilance diminish free riding by children in a social dilemma. *Journal of Experimental Child Psychology, 167,* 1-9. doi:10.1016/j.jecp.2017.10.007.

Engelmann, J.M., Herrmann, E., & Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. *PLoS One, 7*. e48433. doi:10.1371/journal.pone.0048433.

Engelmann, J.M., Over, H., Herrmann, E. & Tomasello, M. (2013). Young children care more about their reputation with ingroup members and potential reciprocators. *Developmental Science, 16,* 952-958. doi:10.1111/desc.12086.

Fehr, E., Bernhard, H. & Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, *454*, 1079-1084. doi:10.1038/nature07155.

Fehr, E. & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*, 785-791. doi:10.1038/nature02043.

Fehr, E., & Schneider, F. (2010). Eyes are on us, but nobody cares: are eye cues relevant for strong reciprocity?. *Proceedings of the Royal Society of London B: Biological Sciences*, *277*, 1315-1323. doi:10.1098/rspb.2009.1900.

Fu, G., Heyman, G. D., Qian, M., Guo, T., & Lee, K. (2016). Young children with a positive reputation to maintain are less likely to cheat. *Developmental Science*, *19*, 275-283. doi:10.1111/desc.12304.

Fujii, T., Takagishi, H., Koizumi, M., & Okada, H. (2015). The effect of direct and indirect monitoring on generosity among preschoolers. *Scientific Reports*, *5*, 9025. doi:10.1038/srep09025.

Gintis, H., Smith, E. A. & Bowles, S. (2001). Costly signaling and cooperation. *Journal of Theoretical Biology*, *213*, 103-119. doi:10.1006/jtbi.2001.2406.

Giving USA Foundation (2019). The annual report on philanthropy for the year 2018. Indianapolis: The Center on Philanthropy at Indiana University.

Guenther, C. L., & Alicke, M. D. (2008). Self-enhancement and belief perseverance. *Journal of Experimental Social Psychology*, *44*, 706-712. doi:10.1016/j.jesp.2007.04.010

Halabi, S., Dovidio, J. F., & Nadler, A. (2008). When and how do high status group members offer help: Effects of social dominance orientation and status threat. *Political Psychology*, *29*, 841-858. doi:10.1111/j.1467-9221.2008.00669.x.

Haley, K. J., & Fessler, D. M. (2005). Nobody's watching?: Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior*, *26*, 245-256. doi:10.1016/j.evolhumbehav.2005.01.002.

Ham, J., & Vonk, R. (2011). Impressions of impression management: Evidence of spontaneous suspicion of ulterior motivation. *Journal of Experimental Social Psychology*, *47*, 466-471. doi:10.1016/j.jesp.2010.12.008.

Heyman, G. D., Barner, D., Heumann, J., & Schenck, L. (2013). Children's sensitivity to ulterior motives when evaluating prosocial behavior. *Cognitive Science*, *38*, 683-700. doi:10.1111/cogs.12089.

Heyman, G. D., Fu, G., Barner, D., Zhishan, H., Zhou, L., & Lee, K. (2016). Children's evaluation of public and private generosity and its relation to behavior: Evidence from China. *Journal of experimental child psychology, 150*, 16-30. doi:10.1016/j.jecp.2016.05.001.

Hopkins, N., Reicher, S., Harrison, K., Cassidy, C., Bull, R., & Levine, M. (2007). Helping to improve the group stereotype: On the strategic dimension of prosocial behavior. *Personality and Social Psychology Bulletin, 33*, 776-788. doi:10.1177/0146167207301023.

House, B., Henrich, J., Sarnecka, B., & Silk, J. B. (2013). The development of contingent reciprocity in children. *Evolution and Human Behavior, 34*, 86-93. doi:10.1016/j.evolhumbehav.2012.10.001.

Klein, O., & Azzi, A. E. (2001). The strategic confirmation of meta-stereotypes: How group members attempt to tailor an out-group's representation of themselves. *British Journal of Social Psychology, 40*, 279-293. doi:10.1348/014466601164759

Kogut, T., & Ritov, I. (2007). "One of us": Outstanding willingness to help save a single identified compatriot. *Organizational Behavior and Human Decision Processes, 104*, 150-157. doi:10.1016/j.obhdp.2007.04.006.

Kost, G. J., Tran, N. K., Tuntideelert, M., Kulrattanamaneeporn, S., & Peungposop, N. (2006). Katrina, the tsunami, and point-of-care testing: Optimizing rapid response diagnosis in disasters. *American Journal of Clinical Pathology, 126*, 513-520. doi:10.1309/NWU5E6T0L4PFCBD9.

Kuznetsova, A., Brockhoff, P.B., & Christensen, R.H.B. (2016). lmerTest: Tests in linear mixed effects models. *R package version, 2*, 0-252015.

Lamba, S., & Mace, R. (2010). People recognize when they are really anonymous in an economic game. *Evolution and Human Behavior, 31*, 271-278. doi:10.1016/j.evolhumbehav.2010.02.002.

Leimgruber, K.L., Shaw, A., Santos, L.R. & Olson, K.R. (2012). Young children are more generous when others are aware of their actions. *PLoS One, 7*, e48292. doi:10.1371/journal.pone.0048292.

Levine, M., Prosser, A., Evans, D. & Reicher, S. (2005). Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shapes helping behavior. *Personality and Social Psychology Bulletin, 31*, 443-453. doi:10.1177/0146167204271651.

Mifune, N., Hashimoto, H., & Yamagishi, T. (2010). Altruism toward in-group members as a reputation mechanism. *Evolution and Human Behavior, 31*, 109-117. doi:10.1016/j.evolhumbehav.2009.09.004.

Misch, A., Over, H., & Carpenter, M. (2014). Stick with your group: Young children's attitudes about group loyalty. *Journal of Experimental Child Psychology, 126,* 19-36. doi:10.1016/j.jecp.2014.02.008.

Moore, C. (2009). Fairness in children's resource allocation depends on the recipient. *Psychological Science, 20,* 944-948. doi:10.1111/j.1467-9280.2009.02378.x.

Nesdale, D., & Flesser, D. (2001). Social identity and the development of children's group attitudes. *Child Development, 72,* 506-517. doi:10.1111/1467-8624.00293.

Nesdale, D., Griffith, J., Durkin, K., & Maass, A. (2005). Empathy, group norms and children's ethnic attitudes. *Journal of Applied Developmental Psychology, 26,* 623-637. doi:10.1016/j.appdev.2005.08.003.

Ng, R., Heyman, G.D., & Barner, D. (2011). Collaboration promotes proportional reasoning about resource distribution in young children. *Developmental Psychology, 47,* 1230-1238. doi:10.1037/a0024923.

Olson, K. R. & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition, 108,* 222-231. doi:10.1016/j.cognition.2007.12.003.

Padilla-Walker, L.M., & Carlo, G. (2014). The study of prosocial behavior: Past, present, and future. In L.M. Padilla-Walker & G. Carlo (Eds.), *Prosocial development: A multidimensional approach,* (pp.3-16). New York, NY: Oxford University Press.

Paulus, M. (2016). Friendship trumps neediness: The impact of social relations and others' wealth on preschool children's sharing. *Journal of Experimental Child Psychology, 146,* 106-120. doi:10.1016/j.jecp.2016.02.001.

Piazza, J., Bering, J.M., & Ingram, G. (2011). "Princess Alice is watching you": Children's belief in an invisible person inhibits cheating. *Journal of Experimental Child Psychology, 109,* 311-320. doi:10.1016/j.jecp.2011.02.003.

Powell, K. L., Roberts, G., & Nettle, D. (2012). Eye images increase charitable donations: Evidence from an opportunistic field experiment in a supermarket. *Ethology, 118,* 1096-1101. doi:10.1111/eth.12011.

Raihani, N. J., & Bshary, R. (2012). A positive effect of flowers rather than eye images in a large-scale, cross-cultural dictator game. In *Proceedings of the Royal Society B: Biological Science*. doi:10.1098/rspb.2012.0758.

Rapp, D. J., Engelmann, J. M., Herrmann, E., & Tomasello, M. (2019). Young children's reputational strategies in a peer group context. *Developmental Psychology, 55,* 329-336. doi:10.1037/dev0000639.

Rigdon, M., Ishii, K., Watabe, M., & Kitayama, S. (2009). Minimal social cues in the dictator game. *Journal of Economic Psychology*, *30*, 358-367. doi:10.1016/j.joep.2009.02.002

Ruggeri, A., Luan, S., Keller, M., & Gummerum, M. (2017). The influence of adult and peer role models on children' and adolescents' sharing decisions. *Child Development*, *89*, 1589-1598. doi:10.111/cdev.12916.

Saucier, D. A., Miller, C. T., & Doucet, N. (2005). Differences in helping whites and blacks: A meta-analysis. *Personality and Social Psychology Review*, *9*, 2-16. doi:10.1207/s15327957pspr0901_1.

Simon, B., Stürmer, S., & Steffens, K. (2000). Helping individuals or group members? The role of individual and collective identification in AIDS volunteerism. *Personality and Social Psychology Bulletin*, *26*, 497–506. doi:10.1177/0146167200266008.

Shaw, A., Monitinari, N., Piovesan, M., Olson, K.R., Gino, F. & Norton, M.I. (2014). Children develop a veil of fairness. *Journal of Experimental Psychology*, *143*, 363-375. doi:10.1037/a0031247.

Silk, J. B., & House, B. R. (2011). Evolutionary foundations of human prosocial sentiments. *Proceedings of the National Academy of Sciences*, *108*, 10910-10917. doi:10.1073/pnas.1100305108.

Spielman, D. A. (2000). Young children, minimal groups and dichotomous categorization. *Personality and Social Psychology Bulletin*, *26*, 1433-1441. doi:10.1177/0146167200263010.

Stürmer, S., Snyder, M., Kropp, A., & Siem, B. (2006). Empathy-motivated helping: The moderating role of group membership. *Personality and Social Psychology Bulletin*, *32*, 943-956. doi:10.1177/0146167206287363.

Stürmer, S., Snyder, M., & Omoto, A. M. (2005). Prosocial emotions and helping: the moderating role of group membership. *Journal of Personality and Social Psychology*, *88*, 532-546. doi:10.1037/0022-3514.88.3.532.

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

Tajfel, H., & Turner, J. C. (2004). The social identity theory of intergroup behavior. In J. T. Jost & J. Sidanius (Eds.), *Political psychology: Key readings* (pp. 276-293). London: Psychology Press.

Van Leeuwen, E., & Täuber, S. (2010). The strategic side of outgroup helping. In S. Stürmer & M. Snyder (Eds.), *The psychology of prosocial behavior: Group processes, intergroup relations, and helping*, (pp. 81-99). Chichester, UK: Wiley-Black.

Van Vugt, M., & Hardy, C. L. (2010). Cooperation for reputation: Wasteful contributions as costly signals in public goods. *Group Processes & Intergroup Relations*, *13*, 101-111. doi:10.1177/1368430209342258.

Wolf, L. K., Bazargani, N., Kilford, E. J., Dumontheil, I., & Blakemore, S. J. (2015). The audience effect in adolescence depends on who's looking over your shoulder. *Journal of Adolescence*, *43*, 5-14. doi:10.1016/j.adolescence,2015.05.003.

Yamagishi, T., & Mifune, N. (2008). Does shared group membership promote altruism? Fear, greed, and reputation. *Rationality and Society*, *20*, 5-30. doi:10.1177/1043463107085442.

Yu, J., Zhu, L., & Leslie, A. M. (2016). Children's sharing behavior in mini-dictator games: The role of in-group favoritism and theory of mind. *Child Development*, *87*(6), 1747-1757. doi:10.111/cdev.12635.

Chapter 1 Appendix A: Competitive Stories of Minimal Groups in Study 1, Phase 1

The Orange group really wants to win against the Green group, and the Green group really wants to win against the Orange group. The Green group wants to show everyone that they're the best group, and the Orange group wants to show everyone that they're the best group.

At lunchtime, the Orange group always wants to sit at the table first, and the Green group always wants to sit at the table first.

When it's time to play games, the Green group always wants to be the winner and the Orange group always wants to be the winner.

At school, the Orange group always wants to be the first group to get on the bus, and the Green group always wants to be the first group to get on the bus.

*Note: Experimenter counterbalanced whether Orange group or Green group was mentioned first across participants.

Chapter 1 Appendix B: Additional Analyses from Study 1, Phase 2

Past work shows that children's judgments of how individuals *should* share resources with ingroup and outgroup members diverge from how they expect individuals to actually behave (DeJesus et al., 2014). Here, we tested whether children behaved according to their own normative standards, or whether their behaviors diverged from what they judged to be nice. To do this, we compared children's sharing in Phase 1 with their evaluations in Phase 2. We computed each participant's' average number of stickers shared in Phase 1 ('Mean Stickers Donated') as well as their 'Ingroup Sharing Bias', which was calculated by subtracting the mean number of stickers donated to outgroup members from the mean number of stickers donated to ingroup members. We then tested whether these factors predicted children's niceness ratings in Phase 2 using model comparisons between a reduced linear regression model—the final model used to evaluate children's niceness ratings in Phase 2[7]—and a fuller model which included Age, Mean Stickers Donated and Ingroup Sharing Bias as predictor variables and participants' average niceness ratings as the response variable.[8] A likelihood ratio test revealed no difference between the full model and the reduced model $\chi^2(2) = 3.33$, $p = .189$, indicating that children's own sharing behavior in Phase 1 was not a robust predictor for how they evaluated the sharing behavior of others in Phase 2.

[7] Average Niceness Ratings ~ Group Sharing Type + Resource Allocation + Age.
[8] The full model specification was: Average Niceness Ratings ~ Mean Stickers Donated + Ingroup Sharing Bias + Group Sharing Type + Resource Allocation + Age.

CHAPTER 2

Children's intergroup attitudes: Insights from Iran

Haleh Yazdi, David Barner, Gail D. Heyman

Department of Psychology, University of California, San Diego

Abstract

Children generally favor individuals in their own group over others, but it is unclear which dimensions of the out-group affect this bias. This issue was investigated among 7- to 8-year-old and 11- to 12-year-old Iranian children (N = 71). Participants evaluated in-group members and three different out-groups: Iranian children from another school, Arab children, and children from the United States. Children's evaluations closely aligned with the perceived social status of the groups, with Americans viewed as positively as in-group members and Arabs viewed negatively. These patterns were evident on measures of affiliation, trust, and loyalty. These findings, which provide some of the first insights into the social cognition of Iranian children, point to the role of social status in the formation of intergroup attitudes.

1. Introduction

People tend to favor individuals from their own group over individuals outside of their group (Aboud, 1993; Brown, 1995; Nesdale & Flesser, 2001). These biases develop early in life and can result in prejudiced beliefs and discriminatory behavior toward outgroup members. Prior research shows that merely being a member of the outgroup can be sufficient to induce a negative bias (Misch, Over, & Carpenter, 2016; Richter, Over, & Dunham, 2016; Schug, Shusterman, Barth & Patalano, 2013). However, it may still be that different dimensions of the specific groups in question have important implications for intergroup dynamics. In the present work, we address this question by examining how Iranian children between the ages of 7 and 12 evaluate groups that contrast along three possible dimensions: a *similarity* dimension, which refers to perceived similarity to the ingroup, an *intergroup relations* dimension, which refers to social and political relations between the groups, and a *social status* dimension, which refers to the relative social status of the ingroup and outgroups.

According to social identity theory (Tajfel & Turner, 1986) and self-categorization theory (Turner, Hogg, Oakes, Reicher & Wetherell, 1987), children often derive their sense of identity and feelings of self-worth from their group membership. This can lead children to adopt an "us versus them" mentality (Aboud, 1988) in which they accept or reject one another on the basis of social category membership. Support for the idea that creating an outgroup is sufficient to induce bias comes from studies in which children are randomly assigned to "minimal groups" that differ along an arbitrary dimension, which produces strong ingroup preferences (Dunham, Baron, & Carey, 2011; Misch, et al., 2016; Schug et al., 2013). Notably, this work establishes that mere social categorization can lead to biases and provides a way to examine intergroup attitudes in a systematic, controlled environment. However, this work is not informative about whether children make distinctions in how they reason about different types of outgroups.

If children do make distinctions between different outgroups, one possibility is that they make distinctions between outgroups with reference to the similarity dimension. This possibility is supported by evidence that similarity-based preferences are evident very early in life: even infants prefer native speakers of their language (Kinzler, Dupoux, & Spelke, 2007) and individuals who share their food preferences (Mahajan & Wynn, 2012). Additionally, infants raised in a racially homogenous environment generally prefer faces of their own race (Bar-Haim, Ziv, Lamy, & Hodes, 2006). This finding is consistent with evidence that older children tend to make more positive evaluations of individuals who are similar to them along a wide range of dimensions such as physical appearance, national identity, beliefs, and personality (Bennett et al., 2004; Nesdale & Flesser, 2001; Powlishta, 1995). Further support for the possibility that children attend to the similarity of outgroups comes from evidence that children evaluate outgroup members more favorably when they are of the same race. In one such study with Taiwanese 3- to 11-year-olds, children selected other non-Taiwanese Asian children as potential play partners significantly more often than White or Black children, even when the other Asian children were identified as Japanese, a group that within Taiwanese society is historically associated with conflict (Kowalski & Lo, 2001). Moreover, children prefer outgroup members who share their psychological characteristics, such as attitudes and social interaction style (Howes & Philipsen, 1992). These findings suggest that children's liking of others may be a function of their perceived similarity.

A second possibility is that children make distinctions between outgroups with reference to the intergroup relations dimension (Oppenheimer & Barrett, 2011). Consistent with this possibility are findings that Jewish Israeli children, who are familiar with the sociohistorical Israeli-Arab conflict, ascribe more negative stereotypes to Arabs than to Jews (Bar-Tal, 1996; Teichman, 2001). Similar findings have been observed among children residing in Bosnia after the

Bosnian war, with Bosniak and Serbian children rating their own group more positively than members of the enemy outgroup (Oppenheimer & Midzic, 2011). These findings suggest that historical relations and prevailing attitudes toward outgroups may help shape children's feelings toward individuals in those groups. It is also possible that intergroup relations could have an increased effect once children begin to grasp complex sociohistorical meanings that relate to intergroup perceptions (see Dunham, Baron, & Carey, 2011).

A final possibility is that children make distinctions between outgroups with reference to the social status dimension. This possibility is consistent with evidence that as early as three years old, children are aware of status and sometimes prefer to affiliate and share resources with members of high-status groups (Horwitz, Shutts, & Olson, 2014; Newheiser, Dunham, Merrill, Hoosain, & Olson, 2014). It is also consistent with evidence that, when given the opportunity, children of a low-status group will sometimes leave their group to join a higher status group (Bigler, Brown, & Markell, 2001), and that members of higher status groups exhibit more ingroup favoritism and loyalty than members of lower status groups (Horwitz et al., 2014; Mullen, Brown & Smith, 1992). For example, in the U.S., where European Americans have higher levels of income, education, and occupational prestige than African Americans, African American children often show no ingroup favoritism and sometimes even show anti-Black bias (Newheiser & Olson, 2012; Spencer & Markstrom-Adams, 1990). Finally, this possibility is consistent with work suggesting that a group's status may have implications for implicit bias (Dunham et al., 2007; Qian et al., 2016). For example, Qian et al. (2016) examined the development of implicit racial bias among preschool children and adults in China and Cameroon by assessing relative preferences for Black, White, and Chinese faces in both populations. Young children's biases were unaffected by

social status, with both populations showing strong ingroup preferences. In contrast, adults' biases differed based on their perceived social status of the groups.

The current study examines whether the nature of outgroups matters to children, and if so, how this might change with development. Our primary focus was on children's interest in affiliating with four different targets: Iranian ingroup members, and three different outgroup members: Iranian children from a different local school, Arab children, and American children. We tested Iranian children both because there has been little research on their social-cognitive development, and because of their particular relation to the outgroups we tested. Specifically, if the similarity dimension is of primary importance in children's judgments of outgroups, one would expect that Iranian children would rather affiliate with other Iranian children (whether from their ingroup or from another school) than Arab children or American children. This should be especially true given that Iran's population is racially and religiously homogenous (99.4% of the population identifies as Muslim; United Nations Iran Census Results; 2011). One would also expect to see the lowest level of desire to associate with American children who are the least similar outgroup in terms of culture, language, religion, and a variety of other dimensions (Ahmadi, 2013; Riazi, 2005; Versteegh, 2001).

However, if the intergroup relations dimension is of primary importance, one would expect Iranian children to prefer to affiliate with Iranians over Americans and Arabs due to Iran's history of social and political conflict with Arab countries and the United States (Ahmadi, 2013; Fürtig, 2002; Murray, 2009). Although the U.S.-Iran conflict didn't become especially intense until 1979, following the seizure of American hostages in Tehran (Houghton, 2001; although see Bakhshandeh, 2015 for an argument that tensions began in 1925), Arab countries have posed a longstanding threat to Iran beginning with the first Arab invasion in 633 (Akram, 2009), which

resulted in the Islamization of Iran and adoption of Arabic script and words into the Farsi language (Kia, 1998; Lorentz, 2010; Mehran, 2010; Tavakoli-Targhi, 1990). These invasions were followed centuries later by a rise in Persian nationalism and resistance to the "Arabization of Iranian culture", such that nationalists took great strides to transform the written history of Iran, and purge Farsi of foreign words and terminologies (Kia, 1998; Tavakoli-Targhi, 1990). Iranian children are taught political, cultural, and religious ideologies through their school textbooks, which are assigned by government authorities and centralized throughout the country (Mehran, 2010). A prominent theme throughout these textbooks is an emphasis on "Iranian-ness" and the need to protect Iranian land, culture, language, and religious ideologies from foreign influences and enemies, particularly from Arab regimes and the U.S. who are illustrated as the "other" (KhosraviNik & Zia, 2014; Mehran, 2010).

Finally, if the social status dimension is of primary importance, one would expect to see the strongest preference for American children, who rank the highest of all groups in GDP (World Bank, 2017) and on the Social Progress Index (Social Progress Index, 2018), which assesses the social and environmental opportunities of each country, such as to technology and higher education. Although it is very unlikely that Iranian children are aware of the GDP standing of the U.S., they commonly hear messages relating to the status and prosperity of Americans from adults, relatives residing abroad, and media portrayals of Americans as consumerists of luxury brands, large houses, transnational vacations, and expensive cars (Alikhah, 2008).

We tested children between the ages of 7 and 12 because previous research shows that children's racial and ethnic prejudices undergo a developmental shift during these years, with young children often exhibiting higher levels of ingroup favoritism than older children (Aboud, 1988; Doyle & Aboud, 1995). There is also evidence that as children reach middle childhood, they

seek membership in a positively distinctive group and give increasingly more weight to factors such as their nation's social representations of national outgroups, the presence of intergroup conflict, and perceived status of the outgroup (Brown, 1995; Hewstone, Rubin, & Willis, 2002; Nesdale, 2017; Oakes, Haslam & Turner, 1994). However, other research shows that children's negative attitudes toward ethnic outgroups peaks at age 6 and then decreases until age 12 (Aboud, 1988; Barrett, Wilson, & Lyons, 2003; Teichman, 2001). We also chose an age range in which even the youngest children would have some basic knowledge of intergroup relations and social status (Barrett, 2007; Shutts, Brey, Dornbusch, Slywotzky, & Olson, 2016), as well as the cognitive and linguistic capacities to understand our measures.

Although our focus was on children's interest in affiliating with different groups, we also included measures of trust and loyalty for exploratory purposes. This was of interest given debates about the extent to which children differentiate between traits of the same valence (Lane, Wellman, & Gelman, 2013). Measures of trust and loyalty were of particular interest because they assess emotional connection, and it is possible that children want to affiliate with people while still keeping emotional distance from them. For example, they might want to spend time with high status individuals as way to enhance their own status, but still view them as untrustworthy and unworthy of loyalty.

2. Method

2.1 Participants

We tested 71 Iranian children from three upper-middle class single-sex elementary schools in Mehrshahr, a suburb of Karaj, Iran located 20 miles west of Tehran. Female participants were tested at an elementary school and middle school, and male participants were tested at a K-12th grade school. Children were enrolled in 2nd and 6th grade and categorized into a younger age group

with 35 children (19 females) (*M* = 7.59 years, *SD* = 0.34, range = 7.09 – 8.22) and an older age group with 36 children (16 females) (*M* = 11.64 years, *SD* = 0.35, range = 11.10 - 12.78). All participants were native speakers of Farsi, and all materials and instructions were administered in this language. Informed consent was obtained from all parents or legal guardians and assent was obtained from all child participants. All consent forms were signed by parents in the principal's office prior to testing. The study was approved by the Research Ethics Review Committee both in the U.S. and Iran. Testing sessions were led by the primary experimenter, an Iranian-American female fluent in Farsi, along with a teacher from the participants' schools. Instructions were given by the primary experimenter and study questions were read aloud by children's teacher. Children were not informed that this study was run in collaboration with an American institution, so as not to influence their responses.

2.2 Overview of Procedure

Children completed questionnaires in which they were asked about four different social groups: Iranian children from their own school whom they did not know, Iranian children from a nearby school, Arab children, and children from the United States. Each question was read out loud and participants were instructed to respond as if language and travel were not barriers to interacting with members of different social groups. Children responded to six questions for each outgroup: two questions pertaining to their desire for affiliation, two questions assessing their feelings of trust, and two questions pertaining to their feelings of loyalty. In addition, a manipulation check was included in which participants ranked the social status of all four groups.

Participants were first familiarized with the 4-point scale that they used to report how much they wanted to interact with members of different social groups by answering practice questions about how much they would like to eat ice cream, ride a roller coaster, and complete extra

homework. Response options ranged from 0= *definitely no* to 3= *definitely yes*. Next, participants were read an introductory statement telling them to imagine a group of peers that they had never met before. For example, when asked to evaluate American children, the introductory statement was as follows: "Imagine you had the chance to get to know a group of American children your age that you haven't met before". Next, children were asked our primary dependent measures which assessed interest in affiliation. On these *affiliation measures* children were asked: "If you had the chance, how much would you want to be friends with these children?" and "How much would you want to play with them?". Next, children were asked the *trust measures*. On these measures, they were asked, "If you asked these children to keep your secret, how much do you think they would do it?" and "If these children promised to help you, how much do you think they would do it?". Finally, children were asked the *loyalty measures,* in which they were asked how likely they would be to tell on members of each social group if they cheated in a sports competition ("If you found out that these children cheated in a sports competition, would you tell a teacher that they cheated?") and in a music competition ("If you found out that these children cheated in a music competition, would you tell a teacher that they cheated?"). Responses to the loyalty questions were reverse coded such that a response of *definitely no* was scored as indicating the highest level of loyalty.

To test our assumption that children would view Americans as having higher status than both Iranians and Arabs, we used a ladder that has reliably been used with children and adults to assess their beliefs about the social status of different individuals (Goodman et al., 2001; Kraus & Tan, 2015; Mistry, Brown, White, Chow & Gillen-O'Neel, 2015). Children were asked to rank the status of each of the groups on a 10-rung ladder, with 1 being the lowest rank and 10 being the highest, to confirm. Participants first heard an introductory statement as follows: "Imagine that

this ladder is for status. At the top are people who have a lot of things. They have a lot of money, live in big houses, and have good education and jobs. At the bottom are people who don't have very much. They don't have much money, live in small houses and don't have the best jobs. In the middle are people who have some money, live in middle-sized houses and have some things but not so many, and have OK jobs. Now think about these four groups of children: Iranian children from your school, Iranian children from another school, Arab children and American children. Where would you put them on the ladder if the top is the people who have the most and the bottom is the people who have the least?" Participants then assigned each group to one rung of the ladder with higher rungs indicating higher status.

3. Results

Analysis of children's social status rankings of the different groups confirmed that Iranian children view Americans as the highest status group (see **Fig. 2.1**). Pairwise comparisons using the Wilcoxon rank sum test showed that all groups were ranked significantly different from one another (all p values < .02), with participants ranking American children (Mdn = 10, SD = 2.44) higher than the Iranian ingroup (Mdn = 8, SD = 2.22), Iranian children from another school (Mdn = 6, SD = 3.0), and Arab children (Mdn = 4, SD = 3.42).

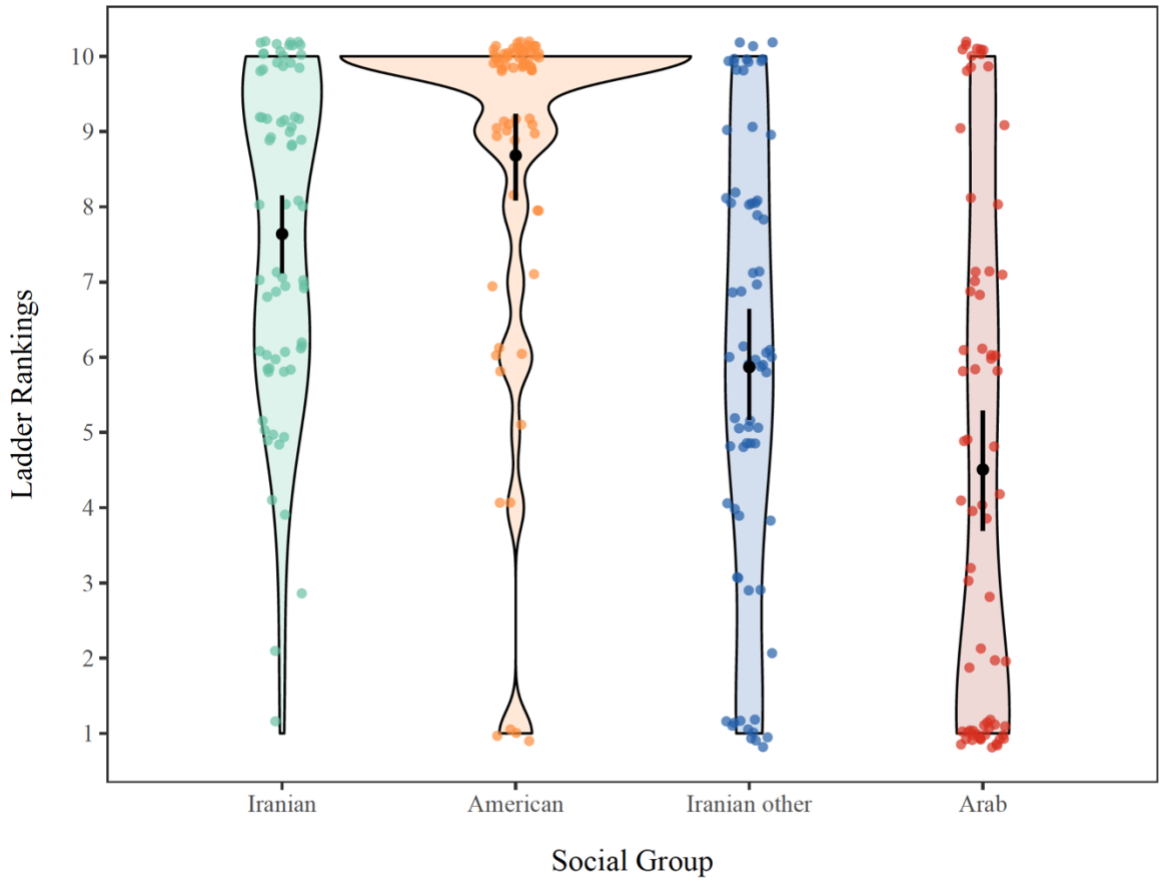

Figure 2.1. Boxplot of children's median rankings of each Social Group in regard to social status from 1 (lowest status) to 10 (highest status). Violin plots show density distribution of responses. Black dots represent means, vertical black lines represent 95% confidence intervals, and each colored dot represents one participant response. Dots are jittered so that they do not overlap.

Data from children's responses to affiliation, trust, and loyalty questions were analyzed using a linear mixed model with participants treated as a random factor. By treating participants as a random factor, the linear mixed model accounts for variance in children's ratings of different social groups and variance in how each child uses the rating scale. We compared a reduced model to fuller models that included one additional parameter and report the difference between the reduced model and the full model with the best fit. The reduced model included Age Group (young, old), Gender (male, female) and Social Group (Iranian children from the same school, Iranian children from a different school, Arab children, American children) as predictor variables, and

participants' ratings from 0 (definitely no) to 3 (definitely yes) as the outcome variable. The intermediate model included all variables in the reduced model in addition to social status rankings of groups. The addition of children's social status rankings significantly improved the model's fit ($\chi^2(17) = 39.84$, $p < .001$). The full model included all variables in the intermediate model in addition to the interaction between Social Group and Question Type (affiliation, trust, and loyalty measures). The full model differed significantly from the basic model, $\chi^2(10) = 193.04$, $p < .001$, and a test of model fit indicated moderate goodness of fit ($R^2 = .40$). The degrees of freedom reported in the analysis were computed using the Satterthwaite approximation and post hoc analyses were computed using Tukey's Honestly Significant Differences (HSD) with p set at .05.

As predicted, and evident in **Figure 2.2**, a main effect of Social Group was observed ($F(3,1619) = 52.46$, $p < .001$), and post-hoc comparisons revealed that, overall, Iranian children gave significantly higher ratings to their ingroup ($M = 1.67$, SE = .08) than Iranian children from another school (HSD, $p < .001$, $M = 1.36$, SE = .08) and Arab children (HSD, $p < .001$, $M = .99$, SE = .08), but not American children (HSD, p = n.s.; $M = 1.76$, SE = .08)—a group which they rated as favorably as their own group. Also in line with our predictions, there was a main effect of children's social status rankings of groups ($F(1,1629) = 39.91$, $p < .001$), such that higher rankings predicted more desired interaction with the group in question.

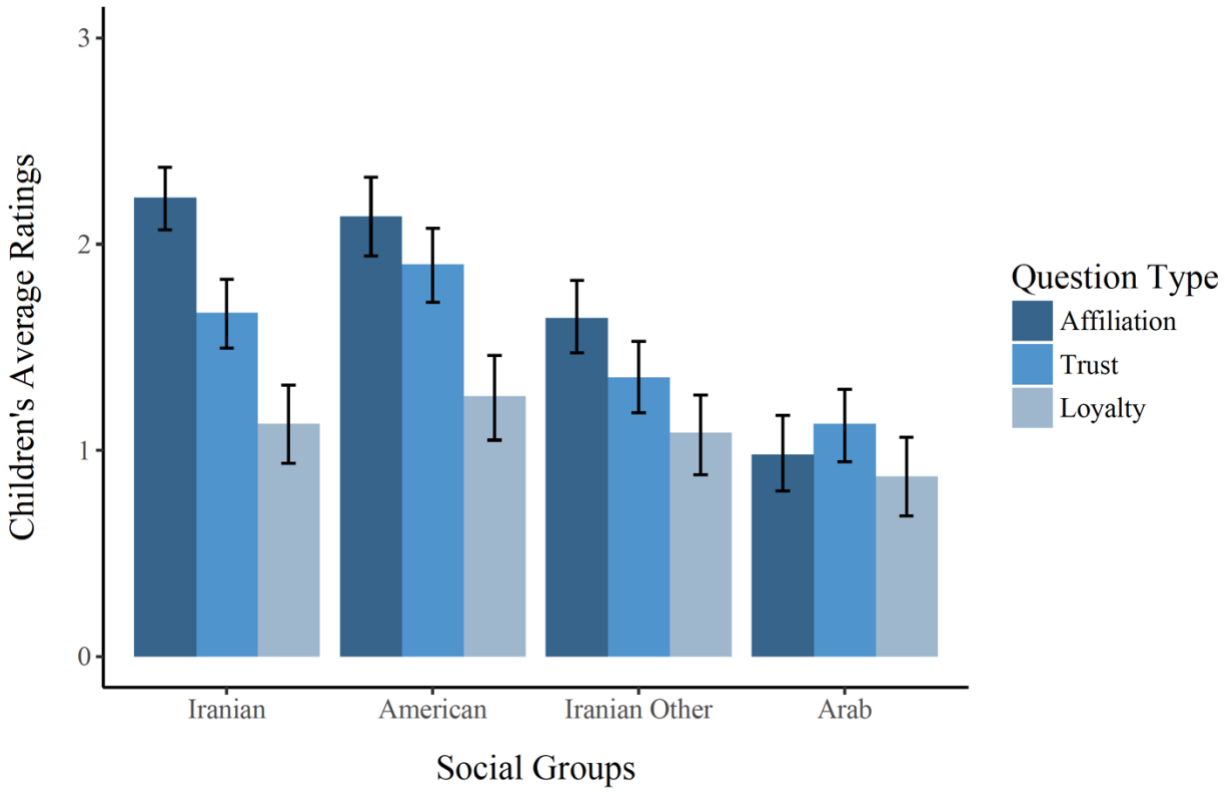

Figure 2.2. Children's average ratings (0 = *definitely no*, 3 = *definitely yes*) of different social groups across three question types (affiliation, trust, and loyalty measures).

A main effect of Question Type was also found $F(2,1573) = 66.59$, $p < .001$, such that Iranian children gave higher ratings on the affiliation measures ($M = 1.74$, SE = .08) than the trust ($M = 1.51$, SE = .08) and loyalty measures ($M = 1.08$, SE = .08) (all HSDs $p < .01$). A significant interaction between Question Type and Group was also observed ($F(6,1573) = 6.66$, $p < .001$) and post-hoc analyses indicated that, for the affiliation measures, children rated both their ingroup ($M = 2.23$, SE = .10) and American children ($M = 2.13$, SE = .10) significantly higher than Iranian children from another school (HSD, $p < .001$, $M = 1.64$, SE = .10) and Arab children (HSD, $p < .001$, $M = .98$, SE = .10). For the trust measures, children rated their ingroup significantly worthier of trust than Arab children (HSD, $p < .001$ $M = 1.64$, SE = .10), but not significantly different from Iranian children from another school (HSD, p = n.s., $M = 1.35$, SE = .10) or American children

(HSD, p = n.s., M = 1.90, SE = .10). Interestingly, children rated American children as worthier of trust than Iranian children from another school (HSD, $p < .001$). For the loyalty measures, children rated their own group (M = 1.13, SE = .12) and others similarly (HSDs, p = n.s., $M_{\text{Iranianother}}$ = 1.08, SE = .12; M_{American} = 1.26, SE = .12; M_{Arab} = .87, SE = .12), with only a marginally significant difference between how they rated Americans and Arabs (HSD, p =.048). Additionally, we found a significant main effect of gender, such that females gave higher ratings than males across all measures, $F(1,66) = 25$, $p < .001$, as well as a main effect of age, such that older children gave higher ratings overall than younger children, $F(1,66) = 4.87$, p = .031.We did not find any significant interactions between Social Group and Question Type with gender or age, demonstrating that boys and girls and older and younger children showed the same general pattern of rating social groups on the affiliation, trust, and loyalty measures. Together, these findings show that children exhibited a negative outgroup bias towards Iranian children from another school and Arab children on some measures, but not all. Our results also indicate that Iranian children rated American children similar to their own group on all measures.

Overall, these results reveal robust differences in children's ratings of outgroups, with children rating American children the highest in social status and the most favorably on measures of affiliation, trust and loyalty, and Arab children the least favorably on both dimensions. As we describe in the Discussion, these results are most compatible with the pattern of findings expected if children judge outgroups primarily on the status dimension.

4. Discussion

We investigated the potential effects of three dimensions on intergroup attitudes: the similarity dimension (the extent to which members are seen as similar to ingroup members), the intergroup relations dimension (the social and political relations between the groups), and the

social status dimension (the relative status of outgroups and ingroups). Our findings provided the strongest evidence for the importance of the social status dimension. Specifically, we found that participants showed a greater desire to affiliate with American children, whom they identified as having the highest social status of all the groups. Interestingly, this desire was as great as it was for Iranian children at their own school. We also found that children showed the least desire to affiliate with Arab children, whom they rated as having the lowest social status of the groups. These general patterns not only held for affiliation measures, but also for measures of trust and loyalty.

Our findings suggest that children give less weight to the other two dimensions when judging outgroups. The fact that Iranian children evaluated American children much more favorably than other outgroups is inconsistent with the possibility that children judge outgroups primarily on their perceived similarity, given that American children were the outgroup that differed the most in terms of ethnicity, culture, language, religion, and proximity; however, it is still plausible that they viewed themselves as more similar to Americans than the other outgroups on other dimensions. The fact that Iranian children evaluated American children so favorably is also inconsistent with the possibility that children attend most to the intergroup relations dimension, given the high level of political tension between Iran and the U.S.

The fact that we did not observe significant age differences is somewhat surprising given that previous studies show a reduction in outgroup bias between 6 and 12 years of age (Aboud, 1988; Barrett et al., 2003; Doyle & Aboud, 1995). Our results also contrast with findings showing that as children get older, they give increasingly more weight to factors such as global status and sociohistorical conflict between groups (Brown, 1995; Hewstone et al., 2002; Oakes et al., 1994; though see Neto, 2016). This may be because Iranian children hear messages emphasizing the high

social status of Americans early in life, which continue to be emphasized throughout their childhood. However, it is also possible that we could have detected age differences with more statistical power.

It is also somewhat surprising that, despite the political conflict between the U.S. and Iran in recent years, American children were viewed favorably relative to other groups. This is not because of a lack of exposure to the conflict. Tensions between Iran and the U.S. are regularly discussed in media that children are exposed to (Bakhshandeh, 2015; Shaghasemi & Heisey, 2009), and adults commonly talk about them around children, including the effects of the U.S. imposed sanctions on Iran, which had caused Iran's currency to decrease by 30% in value several months before this study was conducted (Drezner, 2015; Ianchovichina, Devarajan, & Lakatos, 2016). These negative views are even promoted in children's educational materials. For example, the history textbook used for all fifth graders in public schools describes the U.S. as the "hidden enemy" behind the Iran-Iraq war and an aggressive force that continues to conspire against Iran (Mehran, 2010). Despite these strong anti-U.S. messages that children are exposed to, it may be that children do not understand these conflicts to the extent that would be needed to affect their judgments, or that they do not weigh this information very heavily. It may also be the case that the positive information Iranian children hear about life in the U.S., such as about access to educational and work opportunities (Shaghasemi & Heisey, 2009), serves to counteract the negative messages they hear. Yet another possibility is that Iranian children distinguish between the U.S. government and American children. There is some evidence for a similar distinction among adults: A 2014 poll found that many Iranian adults (50%) held favorable views of American people despite having unfavorable views of the U.S government (87% of respondents; IranPoll, 2014). Of note, however, is that data for the current study were collected in September of 2016, and relations between the

U.S. and Iran have become more hostile since that time. For example, the Trump administration has carried out policies that have raised tensions between the U.S. and Iran, such as terminating the Iran nuclear deal, reinstating sanctions on Iran, and issuing a travel ban that suspends the entry of Iranians to the United States (Brands, 2017; Fullerton, 2017; Sherman, 2018). Such policies may be contributing to more negative views towards Americans in more recent years (IranPoll, 2018).

Interestingly, Iranian children's favorable ratings of American children also extended to their desire to trust and be loyal to them. This finding is in line with other studies showing that children generally prefer high status groups and attribute a wide range of positive attributes to group members, such as being hardworking, intelligent, and popular (Shutts et al., 2016; Sigelman, 2012). It would not be surprising if children's expressions of trust and loyalty reflect general positive attitudes toward such individuals. However, it should be noted that children and adults sometimes hold negative stereotypes of wealthy people, sometimes judging them to be selfish (Elenbaas & Killen, 2018) and untrustworthy (Durante & Fiske, 2017). It is likely that the extent to which children feel emotionally connected to high status others is affected by the kinds of comments they hear people say about rich or poor people. For instance, they might feel more negative toward individuals if they hear critical comments about their group (see Lane, Conder, & Rottman, 2019). Future research is needed to address this possibility.

Yet another somewhat surprising aspect of our findings is that Iranian children rated Arab children much more negatively than the other outgroups and identified them as having low social status, despite the fact that several Arab countries have a higher GDP per capita than Iran (World Bank, 2017). Iranian children's views of Arab children may arise from their desire to maintain a cultural identity that is distinct from that of Arabs, since they often hear about threats posed by

Arab influence. These messages are often explicit in children's textbooks, including discussions of specific controversies such as the use of the name Persian Gulf versus Arabian Gulf (Mehran, 2010). In line with these teachings, an analysis of online posts made on a Facebook Page called *Persian Gulf* found that in opposing the use of the name Arabian Gulf, Iranian users expressed anti-Arab sentiments, as well as glorification of Iran's pre-Islamic history and distinct cultural identity from an Arab identity (Khosravi, 2008; KhosraviNik & Zia, 2014). Further, children are exposed to anti-Arab messages through popular satellite television networks—the majority of which are broadcast from the U.S. and are owned and operated by Iranian nationalist groups that promote anti-Arab biases (Alikhah, 2007). The view that Iranians need to actively maintain a cultural, religious, and political identity distinct from Arabs is still widely held by Iranians and may be passed on to children by parents, or in school as children learn about the impact of Arab culture on Iran's history, religion, and language (Mehran, 2010). For children, maintaining a distinct cultural identity becomes increasingly important during middle childhood, as they learn the socio-cultural and political history and achievements of their national group and begin to place greater importance on acceptance and identification with their national or cultural ingroup (Nesdale, 2017; Oppenheimer & Barrett, 2011). Iranian children may also be echoing anti-Arab sentiments they learned from elders who were exposed to anti-Arab propaganda during the time of the Pahlavi dynasty and the Iran-Iraq war (Tavakoli-Targhi, 1990). The fact that our participants rated Arabs much more negatively than Iranian children from both their own school and from another school also provides an interesting contrast with the views of American adults, who often see no distinction between Iranians and Arabs: in a survey we conducted with 82 undergraduate American students, over half of our sample (56%) agreed with the statement that "most Iranians are Arab".

The current study offers some of the only data regarding children's intergroup attitudes in Iran, and the only data we know of regarding their attitudes toward Americans. As noted by Legare and Harris (2016), the vast majority of what we know about development comes from Western, educated, industrialized, rich, democratic (WEIRD) samples. As a consequence, it remains unclear to what extent many major developmental psychology findings reflect fundamental psychological processes versus culture-specific learning. Research in Iran is especially underrepresented in the developmental psychology literature, with only a small body of research examining Iranian children's beliefs (e.g., Davoodi, Corriveau, & Harris, 2016, Novin, Banerjee, Dadkhah, & Rieffe, 2009, and Shahaeian, Nielsen, Peterson, & Slaughter, 2013) and none that we know of examining perceptions of social groups.

Further research will be needed to examine the extent to which our findings generalize to other populations in and outside of Iran. Although Iran is racially and religiously homogenous, there may well be differences in socio-evaluative reasoning based on income and where children live, as has been found in other countries (e.g., Caravita, Giardino, Lenzi, Salvaterra & Antonietti, 2012; Chen, Wang, & Wang, 2009). It is also unknown whether children in wealthy Western countries emphasize social status to the same extent given recent findings suggesting that children sometimes attach negative stereotypes to upper class groups (e.g., Burkholder, Elenbaas, Killen, 2019; Elenbaas & Killen, 2018; Horwitz & Dovidio, 2017).

One significant limitation of our research is that, though we propose three different dimensions of social groups that might shape children's intergroup attitudes, our design did not allow us to test the effects of each dimension directly and independently. For instance, it is possible that children viewed themselves as more similar to members of high-status groups than members of lower status groups (e.g., by comparing themselves on factors such as education and resources

rather than language and culture to maximize perceived similarity). Likewise, children may have been motivated to perceive their intergroup relations as more positive with higher status groups than lower status groups. Thus, the dimensions we examined may overlap in complex and nuanced ways. Future work can help address this limitation by systematically manipulating one dimension at a time in the descriptions of individuals children are asked about. For example, children could be asked to evaluate low versus high status Iranian children as well as low versus high status American children.

Another limitation of the present research is that although we used an adapted version of a well-established status scale (Goodman et al., 2001; Kraus & Tan, 2015; Mistry et al., 2015), there are potentially important aspects of status that we did not tap into. For example, we did not assess how influential different groups were perceived to be, as this could be as important or even more important than what we did assess.

Our findings point to some important implications for Iranian children's peer relationships, some of which may apply to peer relationships more broadly. Iranian children's focus on status concerns means that children may avoid Arabs and other children considered to have lower social status even when given opportunities to interact with them. As a consequence, it will likely take special efforts to humanize members of low status groups, to allow them to be viewed as distinct individuals rather than as part of an undifferentiated group (Heyman & Yazdi, 2019). Our findings also provide hope that Iranian children are likely to welcome intergroup contact opportunities with some dissimilar others, and that such interactions have the potential to promote better Iranian-American relations. Finally, our findings demonstrate that it cannot be assumed that children will always show a preference to affiliate with ingroup members than with outgroup members.

Our use of real-world outgroups that are personally significant to the people being asked about them is an important contribution of our study because it allows us to examine ecologically-valid intergroup conflicts that have important implications for children's lives. However, real-world social groups also vary along many other dimensions that we cannot control for, which makes it difficult to know which aspects of the group children are focusing on most when they make their evaluations. For this reason, it would be beneficial to complement work on real-world groups with methodologies, such as the minimal group paradigm, that allow for closer experimental control (Dunham et al., 2011; Spielman, 2000).

In summary, our work shows that children do not view all social outgroups as interchangeable. Instead, their preferences closely align with the social status of outgroups, and in some cases, they view outgroup members as favorably as ingroup members. Our findings extend to children's ideas about who is trustworthy and deserving of loyalty, which suggests that these judgments have broad implications. Finally, and perhaps most importantly, the findings from the current study broaden our understanding of social cognitive development to include a population that has been fairly absent from the developmental literature.

Acknowledgments

We are grateful for the parents, children, teachers and school directors who made it possible for us to conduct this research in Iran. We thank Tim Lew and Rose Schneider for their help with data analysis, and Telli Davoodi, Drew Nesdale, and Majid Yazdi for their expertise and helpful feedback regarding this work.

Chapter 2, in full, is a reprint of the material as it appears in: Yazdi, H., Barner, D., & Heyman, G.D. (2020) Children's intergroup attitudes: Insights from Iran. *Child Development, 91* (5), 1733-1744. The dissertation author was the primary investigator and author of this paper.

References

Aboud, F. E. (1988). *Children and prejudice*. New York, NY: Basil Blackwell.

Aboud, F.E. (1993). The developmental psychology of racial prejudice. *Transcultural Psychiatric Research Review, 30,* 229-242. doi:10.1177/136346159303000303

Ahmadi, H. (2013). Iran and the Arab Spring: Why haven't Iranians followed the Arabs in waging revolution? *Asian Politics & Policy, 5,* 407-420. doi:10.1111/aspp.12038

Akram, A. I. (2009). *The Muslim conquest of Persia*. Birmingham, UK: Maktabah Publications.

Alikhah, F. (2008). The politics of satellite television in Iran. In M. Semati (Ed.), *Media, Culture and Society in Iran: Living with Globalization and the Islamic State* (pp. 94-110). New York, NY: Routledge.

Bakhshandeh, E. (2015). *Occidentalism in Iran: Representations of the West in the Iranian media*. London, UK: IB Tauris.

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. *Psychological Science, 17,*159-163. doi:10.1111/j.1467-9280.2006.01679.x

Barrett, M. (2007). *Children's knowledge, beliefs and feelings about nations and national groups*. Hove, UK: Psychology Press.

Barrett, M., Wilson, H., & Lyons, E. (2003). The development of national in-group bias: English children's attributions of characteristics to English, American and German people. *British Journal of Developmental Psychology, 21,* 193-220. doi:10.1348/026151003765264048

Bar-Tal, D. (1996). Development of social categories and stereotypes in early childhood: The case of "the Arab" concept formation, stereotype and attitudes by Jewish children in Israel. *International Journal of Intercultural Relations, 20,* 341-370.

Bennett, M., Barrett, M., Karakozov, R., Kipiani, G., Lyons, E., Pavlenko, V., & Riazanova, T. (2004). Young children's evaluations of the ingroup and of outgroups: A multi-national study. *Social Development, 13,*124-141. doi:10.1046/j.1467-9507.2004.00260.x

Bigler, R. S., Spears Brown, C., & Markell, M. (2001). When groups are not created equal: Effects of group status on the formation of intergroup attitudes in children. *Child Development, 72,* 1151-1162. doi:10.1111/1467-8624.00339

Brands, H. (2017). The unexceptional superpower: American grand strategy in the age of Trump. *Survival, 59,* 7-40. doi:10.1080/00396338.2017.1399722

Brown, R. (1995). *Prejudice: Its social psychology*. Cambridge, MA: Blackwell.

Burkholder, A., Elenbaas, L., & Killen, M. (2019). Children's and adolescents' evaluations of intergroup exclusion in interracial and inter-wealth peer contexts. *Child Development.* doi: 10.1111/cdev.13249

Caravita, S. C. S., Giardino, S., Lenzi, L., Salvaterra, M., & Antonietti, A. (2012). Socio-economic factors related to moral reasoning in childhood and adolescence: The missing link between brain and behavior. *Frontiers in Human Neuroscience*, *6*, 262.

Chen, X., Wang, L., & Wang, Z. (2009). Shyness-sensitivity and social, school, and psychological adjustment in rural migrant and urban children in China. *Child Development*, *80*, 1499-1513. doi:10.1111/j.1467-8624.2009.01347.x

Davoodi, T., Corriveau, K. H., & Harris, P. L. (2016). Distinguishing between realistic and fantastical figures in Iran. *Developmental Psychology*, *52*(2), 221. doi:10.1037/dev0000079

Doyle, A. B., & Aboud, F. E. (1995). A longitudinal study of white children's racial prejudice as a social-cognitive development. *Merrill-Palmer Quarterly, 41*, 209-228.

Drezner, D. W. (2015). Targeted sanctions in a world of global finance. *International Interactions*, *41*, 755-764. doi:10.1080/03050629.2015.1041297

Dunham, Y., Baron, A. S., & Banaji, M. R. (2007). Children and social groups: A developmental analysis of implicit consistency in Hispanic Americans. *Self and Identity, 6*, 238-255.

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child Development, 82*(3), 793-811. doi:10.1111/j.1467-8624.2011.01577.x

Durante, F., & Fiske, S. T. (2017). How social-class stereotypes maintain inequality. *Current opinion in psychology*, *18*, 43-48. doi:10.1016/j.copsyc.2017.07.033

Elenbaas, L., & Killen, M. (2018). Children's perceptions of economic groups in a context of limited access to opportunities. *Child Development*. doi:10.1111/cdev.13024

Elenbaas, L., & Killen, M. (2017). Children's perceptions of social resource inequality. *Journal of Applied Developmental Psychology*, *48*, 49-58. doi:10.1016/j.appdev.2016.11.006

Fullerton, M. (2017). Trump, Turmoil, and Terrorism: The US Immigration and Refugee Ban. *International Journal of Refugee Law*, *29*, 327-338. doi:10.1093/ijrl/eex021

Fürtig, H. (2002). *Iran's Rivalry with Saudi Arabia between the Gulf Wars*. Reading, UK: Garnet and Ithaca Press.

Goodman, E., Adler, N. E., Kawachi, I., Frazier, A. L., Huang, B., & Colditz, G. A. (2001). Adolescents' perceptions of social status: Development and evaluation of a new indicator. *Pediatrics, 108*, 31-38. doi:10.1542/peds.108.2.e31

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464-1480. doi:10.1037/0022-3514.74.6.1464

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53,* 575-604. doi:10.1146/annurev.psych.53.100901.135109

Heyman, G.D., & Yazdi, H. (2019). The role of individuation in the development of ingroup preferences. *Infant and Child Development, 28,* e2121. doi:10.1002/icd.2121

Horwitz, S. R., & Dovidio, J. F. (2017). The rich—love them or hate them? Divergent implicit and explicit attitudes toward the wealthy. *Group Processes & Intergroup Relations*, *20*, 3-31. doi:10.1177/1368430215596075

Horwitz, S. R., Shutts, K., & Olson, K. R. (2014). Social class differences produce social group preferences. *Developmental Science*, *17*, 991-1002. doi:10.1111/desc.12181

Houghton, D. P. (2001). *US foreign policy and the Iran hostage crisis* (Vol. 75). Cambridge, UK: Cambridge University Press.

Howes, C., & Philipsen, L. (1992). Gender and friendship: Relationships within peer groups of young children. *Social Development, 1,* 230-242. doi:10.1111/j.1467-9507.1992.tb00126.x

Ianchovichina, E., Devarajan, S., & Lakatos, C. (2016). Lifting economic sanctions on Iran: global effects and strategic responses. *Policy Research Working Papers.* doi:10.1596/1813-9450-7549

IranPoll. (2014, 2018). *Iranian Public Opinion After the Protests Questionnaire*. Retrieved from http://www.cissm.umd.edu/publications/iranian-public-opinion-after-protests

Khosravi, S. (2008). *Young and defiant in Tehran*. Philadelphia, PA: University of Pennsylvania Press.

KhosraviNik, M., & Zia, M. (2014). Persian nationalism, identity and anti-Arab sentiments in Iranian Facebook discourses: Critical discourse analysis and social media communication. *Journal of Language and Politics*, *13*, 755-780.

Kia, M. (1998). Persian nationalism and the campaign for language purification. *Middle Eastern Studies*, *34*, 9-36. doi:10.1080/00263209808701220

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, *104*, 12577-12580.

Kowalski, K., & Lo, Y. F. (2001). The influence of perceptual features, ethnic labels, and sociocultural information on the development of ethnic/racial bias in young children. *Journal of Cross-Cultural Psychology, 32*, 444-455. doi:10.1177/0022022101032004005

Kraus, M. W., & Tan, J. J. X. (2015). Americans overestimate social class mobility. *Journal of Experimental Social Psychology, 58,* 101-111. doi: 10.1016/j.jesp.2015.01.005

Lane, J. D., Wellman, H. M., & Gelman, S. A. (2013). Informants' traits weigh heavily in young children's trust in testimony and in their epistemic inferences. *Child Development, 84,* 1253-1268. doi:10.1111/cdev.12029

Lane, J.D., Conder, E.B., & Rottman, J. (2019). The influence of direct and overheard messages on children's attitudes toward novel social groups. *Child Development,* doi:10.1111/cdev.13238

Legare, C. H., & Harris, P. L. (2016). The ontogeny of cultural learning. *Child Development, 87,* 633-642. doi:10.1111/cdev.12542

Mahajan, N., & Wynn, K. (2012). Origins of "us" versus "them": Prelinguistic infants prefer similar others. *Cognition, 124,* 227-233. doi:10.1016/j.cognition.2012.05.003

McConnell, A. R., & Leibold, J. M. (2001). Relations among the implicit association test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37,* 435-442. doi:10.1006/jesp.2000.1470

Mehran, G. (2010). Representations of Arabs in Iranian Elementary School Textbooks. In A.E. Mazawi & R.G. Sultana (Eds.), *World Yearbook of Education 2010: Education and the Arab 'World': Political Projects, Struggles, and Geometries of Power* (pp. 361-381). New York, NY: Routledge.

Misch, A., Over, H., & Carpenter, M. (2016). I won't tell: Young children show loyalty to their group by keeping group secrets. Journal of Experimental Child Psychology, *142,* 96-106.

Mistry, R. S., Brown, C. S., White, E. S., Chow, K. A., & Gillen-O'Neel, C. (2015). Elementary school children's reasoning about social class: A mixed-methods study. *Child Development, 86,* 1653-1671. doi:10.1111/cdev.12407

Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22,* 103-122.

Murray, D. (2009). *U.S. foreign policy and Iran: American-Iranian relations since the Islamic revolution*. London, UK: Routledge.

Nesdale, D. (2017). Children and social groups: A social identity approach. In Rutland, A., Nesdale, D., & Spears Brown, C. (Eds.), The Wiley Handbook of group processes (pp. 1-22). Chichester, UK: John Wiley & Sons. doi:10.1002/9781118773123

Nesdale, D., & Flesser, D. (2001). Social identity and the development of children's group attitudes. *Child Development, 72,* 506-517. doi:10.1111/1467-8624.00293

Neto, F. (2016). Evaluations of the in-group and of out-groups among Portuguese children. *European Journal of Developmental Psychology, 13*, 472-487.

Newheiser, A. K., Dunham, Y., Merrill, A., Hoosain, L., & Olson, K. R. (2014). Preference for high status predicts implicit outgroup bias among children from low-status groups. *Developmental Psychology, 50*, 1081-1090. doi:10.1037/a0035054

Newheiser, A. K., & Olson, K. R. (2012). White and Black American children's implicit intergroup bias. *Journal of Experimental Social Psychology, 48*, 264-270. doi:10.1016/j.jesp.2011.08.011

Novin, S., Banerjee, R., Dadkhah, A., & Rieffe, C. (2009). Self-reported use of emotional display rules in the Netherlands and Iran: Evidence for sociocultural influence. *Social Development, 18*, 397-411. doi:10.1111/j.1467-9507.2008.00485.x

Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). *Stereotyping and social reality*. Oxford, UK: Blackwell.

Oppenheimer, L., & Barrett, M. (2011). National identity and in-group/out-group attitudes in children: The role of sociohistorical settings. An introduction to the special issue. *European Journal of Developmental Psychology, 8*, 1-4. doi:10.1080/17405629.2010.533948

Oppenheimer, L., & Midzic, E. (2011). National identification and in-group/out-group attitudes with Bosniak and Serbian children in Bosnia. *European Journal of Developmental Psychology, 8*, 43-57. doi:10.1080/17405629.2010.533974

Powlishta, K. K. (1995). Gender bias in children's perceptions of personality traits. *Sex Roles, 32*, 17-28. doi:10.1007/bf01544755

Qian, M. K., Heyman, G. D., Quinn, P. C., Messi, F. A., Fu, G., & Lee, K. (2016). Implicit racial biases in preschool children and adults from Asia and Africa. *Child Development, 87*, 285-296. doi:10.1111/cdev.12442

Riazi, A. (2005). The four language stages in the history of Iran. In A. Lin & P. Martin (Eds.), Decolonization, globalization: Language-in-education policy and practice (pp. 98-114). Clevedon, Buffalo: Multilingual Matters.

Richter, N., Over, H., & Dunham, Y. (2016). The effects of minimal group membership on young preschoolers' social preferences, estimates of similarity, and behavioral attribution. *Collabra, 2*, 1-8. doi:10.1525/collabra.44

Schug, M. G., Shusterman, A., Barth, H., & Patalano, A. L. (2013). Minimal-group membership influences children's responses to novel experience with group members. *Developmental Science, 16*, 47-55. doi:10.1111/j.1467-7687.2012.01193.x

Shaghasemi, E., & Heisey, D. R. (2009). The cross-cultural schemata of Iranian-American people toward each other: A qualitative approach. *Intercultural Communication Studies*, *18*, 143-160.

Shahaeian, A., Nielsen, M., Peterson, C. C., & Slaughter, V. (2014). Cultural and family influences on children's theory of mind development: A comparison of Australian and Iranian school-age children. *Journal of Cross-Cultural Psychology*, *45*, 555–568.

Sherman, W. R. (2018). How we got the Iran deal: And why we'll miss it. *Foreign Affairs*, *97*, 186-197.

Shutts, K., Brey, E. L., Dornbusch, L. A., Slywotzky, N., & Olson, K. R. (2016). Children use wealth cues to evaluate others. *PLoS One, 11*, e0149360.

Sigelman, C. K. (2012). Rich man, poor man: Developmental differences in attributions and perceptions. *Journal of Experimental Child Psychology, 113*, 415-429. doi:10.1016/j.jecp.2012.06.011

Simon, B., & Pettigrew, T. F. (1990). Social identity and perceived group homogeneity: Evidence for the ingroup homogeneity effect. *European Journal of Social Psychology*, *20*, 269-286. doi:10.1002/ejsp.2420200402

Social Progress Index. (2018). Retrieved from http://www.socialprogress.org/

Spencer, M. B., & Markstrom-Adams, C. (1990). Identity processes among racial and ethnic minority children in America. *Child Development, 61*, 290-310. doi:10.2307/1131095

Spielman, D. A. (2000). Young children, minimal groups, and dichotomous categorization. *Personality and Social Psychology Bulletin*, *26*(11),1433-1441.

Tajfel, H., & Turner, J.T. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. Austin (Eds.), *Psychology of intergroup relations* (pp.7-24). Chicago, IL: Nelson-Hall.

Tavakoli-Targhi, M. (1990). Refashioning Iran: Language and culture during the constitutional revolution. *Iranian Studies, 23*, 77-101. doi:10.1080/00210869008701750

Teichman, Y. (2001). The development of Israeli children's images of Jews and Arabs and their expression in human figure drawings. *Developmental Psychology, 37*, 749-761.

Turner, J.C., Hogg, M.A., Oakes, P.J., Reicher, S.D., & Wetherell, M.S. (1987). *Rediscovering the social group: A self-categorization theory*. Cambridge, MA: *Basil Blackwell.*

United Nations World Population and Housing Census Program (2011). *National Population and Housing Census for Iran*. Retrieved from http://unstats.un.org/unsd/demographic-social/census/

Versteegh, K. (2001). Linguistic contacts between Arabic and other languages. *Arabica*, *48*, 470-508. doi:10.1163/157005801323163825

World Bank. (2017). *Gross Domestic Product Ranking Table*. Retrieved from http://datacatalog.worldbank.org/dataset/gdp-ranking

CHAPTER 3

The development of morality and conventionality across cultures: Implementing a two-stage model for cross-cultural research

Haleh Yazdi,[1] Rose M. Schneider, [1] Erica J. Yoon,[2] Michael C. Frank, [2] Mahesh Srinivasan,[3] Yarrow Dunham,[4] David Barner[1]

[1] Department of Psychology, University of California, San Diego

[2] Department of Psychology, Stanford University

[3] Department of Psychology, University of California, Berkeley

[4] Department of Psychology, Yale University

Abstract

Establishing a shared sense of right and wrong is an essential milestone for human cooperation, raising the question of whether humans believe that a universal set of moral intuitions exist. Like elsewhere in the psychological sciences, tests of universality in the domain of human prosocial behavior are hindered by the overrepresentation of participants from Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies. Equally threatening, however, is the problem of validity that arises when researchers use measures developed in WEIRD contexts to compare groups across diverse cultural settings. Here we address this tension between cross-cultural generalizability and validity by deploying a two-stage model to investigate moral cognition, which exploits both the power of large-scale cross-cultural comparison and culture-specific measures. Specifically, we test a classic case study in which strong universality claims have previously been made–the "moral/conventional" distinction–in four societies. Using a standard test of this distinction, we find the strongest evidence of the moral/conventional distinction in Canadian children who rated transgressions defined as "moral" by prior research more severely than "conventional" transgressions. In contrast, our pattern of findings from children in Korea, India, and Iran is more variable, with Iranian children showing the weakest evidence for a moral/conventional distinction–an apparent challenge to universalist claims. Critically, however, when experimental stimuli are tailored to reflect culture-specific norms in a second stage of testing, Iranian children also exhibit a moral/conventional distinction. These findings support the use of a two-stage model in conducting cross-cultural research that achieves both high reliability and cultural validity.

1. Introduction

Why do humans think some behaviors are bad and others are good? Do we have a universal belief that some violations, like murder, are morally wrong independent of specific cultural conventions, but that other violations, such as not adhering to school dress codes, depend on culture-specific social consensus and norms? Or are all moral intuitions shaped by cultural exposure? A challenge to answering this question, and other questions of human universality, is that most research in the psychological sciences has been restricted to Western, educated, industrialized, rich, and democratic (WEIRD) societies (Apfelbaum, Phillips, & Richeson, 2014; Henrich, Heine, & Norenzayan, 2010; Legare & Harris, 2016; Meadon & Spurret, 2010), which represent 16% of the world's population yet constitute 96% of the samples in psychological research (Arnett, 2008; Hardin et al., 2014). Recent work has sought to diversify psychology–and the study of human social and moral cognition–by studying participants from a wider range of socio-economic groups, religious groups, and countries. For example, several high-profile studies have administered batteries of classic behavioral measures of social and moral cognition cross-culturally, including Australia, the Central African Republic, China, Ecuador, Fiji, India, Mexico, Namibia, Peru, Senegal, Uganda, the U.S., and others (Barrett, Peterson, & Frankenhuis, 2016; Dunham, Srinivasan, Dotsch, & Barner, 2014; Fu, Heyman, & Lee, 2016; Kanngiesser, Rossano, Zeidler, Haun, & Tomasello, 2019; Nielsen, M., Mushin, I., Tomaselli, K., & Whiten, A., 2016). These studies, like studies of basic human cognitive processes such as memory, emotions, and problem solving (Halberstadt & Lozada, 2011; Goh et al., 2013; Kitayama, Duffy, Kawamura, & Larsen, 2003; Wang, 2016), have documented considerable variability across different groups, exacerbating the concern that testing only WEIRD participants leads to conclusions that are not generalizable to human psychology more broadly.

Morality and Conventionality

A central construct in moral cognition — the "moral/conventional" distinction — posits that, from early in life, children distinguish between moral norms that are determined by universal principles of justice because they have intrinsic consequences for the welfare and rights of others (e.g. not hitting a victim), and conventional norms that are rooted in concerns with authorities, consensus, or tradition (e.g., not putting one's shoes on the dinner table; Dahl & Killen, 2018; Haidt & Graham, 2007; Smetana, Jambon, & Ball, 2014; Nisan, 1987; Nucci & Turiel, 1993; Nucci, 2001; Turiel, 1983, 2006). Research on the moral/conventional distinction posits that, from early in life, children distinguish between moral norms that are determined by universal principles of justice because they have intrinsic consequences for the welfare and rights of others (e.g. not hitting a victim), and conventional norms that are rooted in concerns with authorities, consensus, or tradition (e.g., not putting one's shoes on the dinner table; Dahl & Killen, 2018; Haidt & Graham, 2007; Smetana et al., 2014; Nisan, 1987; Nucci & Turiel, 1993; Nucci, 2001; Turiel, 1983, 2006). Previous studies find that by three and a half years children judge violations classified as "moral" by researchers as more serious, generalizable across contexts, and independent of rules or authority figures than violations of norms that are classified as "conventional", and that this distinction becomes more pronounced as children age and internalize social norms (Nucci, Turiel & Encarnacion-Gawrych; 1983; Smetana & Braeges, 1990; Smetana, Rote, Jambon, Tasopoulos-Chan, Villalobos, & Comer, 2012).

The moral/conventional distinction is often argued to be a human universal, and several studies comparing groups of Western children with non-Western children find similar developmental patterns in both groups (Nucci et al., 1983; Nucci & Turiel, 1993). In contrast, other studies have challenged the idea of a shared universal moral intuition, either on the basis of

experimental data that finds variability across cultures (Haidt, Koller, & Dias, 1993; Rai & Fiske, 2011; Shweder, Mahaptra, & Miller, 1987), or on theoretical grounds, arguing either that universal moral intuitions imbue all cultural norms, or that cultural norms impact how moral intuitions are expressed (Hauser, Lee, & Huebner, 2010; Turiel, Killen, & Helwig, 1987).

Although most prior work on the moral/conventional distinction has focused on WEIRD populations (Crane & Tisak, 1995; Davidson, Turiel & Black, 1983; Helwig, Tisak, Turiel, 1990; Smetana, 1985; Smetana & Braeges, 1990; Smetana, Ball, Jambon & Yoo, 2018), a number of studies with non-Western children have found variability in how children from different cultural groups judge religious, conventional and moral acts (Chernyak, Kang, & Kushnir, 2019; Chernyak, Kushnir, Sullivan, & Wang, 2013; Hollos, Leis, & Turiel, 1986; Nisan, 1987; Song, Smetana, & Kim 1987; Yau & Smetana, 2003). However, also like other work in psychology, many of these studies have compared groups using measures originally designed to test WEIRD groups, and therefore focus on transgressions that may not be interpreted similarly across cultures. This is potentially a problem when taking a cross-cultural approach if there are moral implications for violating social norms that vary across cultures. For instance, in Iran, some of the items that are considered social conventions in commonly used tests of the moral/conventional distinction are linked to religious practices that are enforced by Iran's judicial system and interpreted as acts against the will of Allah (Kazemipur & Rezaei, 2003). Meanwhile, actions that might contravene the will of Allah are not considered moral violations among non-Muslim children in India (Srinivasan, Kaplan & Dahl, 2018). Although some studies have tested non-WEIRD groups using items tailored to the local culture of participants (Hollos et al., 1986; Nisan, 1987; Nucci, Camino & Sapiro, 1996; Song et al.,1987; Yau & Smetana, 2003; Zimba 1994), they do not directly

compare children across groups, leaving open the question of how children in these groups compare to those from other cultural groups.

A two-stage approach

Cross-cultural work on the moral/conventional distinction highlights a persisting problem in cross-cultural work on social cognition. Although there has been an increased effort to include more non-WEIRD participants in research on human morality and prosocial behavior, the measures used to assess cognition and behavior have often remained anchored in items that were developed for WEIRD populations. In particular, many large-scale cross-cultural studies in psychology use methods developed and validated in a western context to test participants and make intergroup comparisons across cultures (Blake et al., 2015; Callaghan et al., 2005; Ekman et al., 1987; Henrich et al., 2006; House et al., 2013; Robbins & Rochat, 2011; Rochat et al., 2009; van Leeuwen et al., 2018). The merit of these studies is that using a single measure allows researchers to deploy a uniform system for collecting and analyzing data (Moles et al., 1977), which permits an apples-to-apples comparison of groups. In principle, this allows researchers to draw general conclusions about human behavior, and to ask whether certain behaviors or cognitive mechanisms are human universals, or instead vary across groups.

On the other hand, as in other areas of psychology, measures of cognitive development often undergo relatively little formal evaluation for their psychometric properties (Fried & Flake, 2020). They rarely are evaluated for their external validity (e.g., via their relations to other constructs) and they are infrequently evaluated for their reliability (e.g., whether they produce stable measurements within individuals). In the context of cognitive development research, a "standardized" measure then often simply means one that has been used frequently, not one that has been normed with a large population or evaluated systematically. This lack of psychometric

evidence creates a problem for cross-cultural research. If a measure reveals cross-cultural variation, is it the measurement properties that vary cross-culturally, or the underlying construct? Concretely, if we fail to observe a moral/conventional distinction in some cultures, is it because some humans fail to make such a distinction, or because the items used to elicit the distinction are not universally valid measures of the construct?

The issue of cross-cultural validity is not abstract: when measures are not well-calibrated to local cultural practices or when they load heavily on differences between groups that are unrelated to the theoretical questions at hand, they can lead to incorrect inferences (Cheng, 2008; Kline, Shamsudheen & Broesch, 2018; Mason, 2005; Norenzayan & Heine, 2005; Wang, 2016). For example, if groups differ in their access to nutrition or education, findings may not reflect differences in cultural values so much as differences in a participant's ability to attend to a complex task. Thus, while a significant step toward generalizable research in psychology is to include diverse participants, there remains an important problem of how to compare groups while remaining sensitive to the social, political, and cultural realities of participants across cultures.

In the present study, we employed a two-stage approach to conducting cross-cultural work on human moral cognition: we first assess variability across cultural groups on a single set of measures, and then investigate the nature of putative differences between groups using culturally tailored materials. In a first stage, we used a measure of the moral/conventional distinction, which we refer to as a "standard" or "standardized" measure in the sense that has been used widely used in prior developmental (Lahat et al., 2012; Nucci, 1981; Nucci & Turiel, 1978; Sander son & Siegal, 1988; Siegal & Storey, 1985; Smetena 1981, 1985, 1986; Smetana & Braeges, 1990; Turiel, 1983; 2008) to assess and compare children's moral judgments across groups from four different countries: Canada, Iran, India, and Korea. Our findings revealed that children from Canada made

a clear distinction in their evaluations of moral and conventional violations, but children from non-WEIRD societies showed a more variable pattern in their evaluations, with children in Iran showing an especially weak distinction between moral and conventional violations. These findings seem to challenge the claim that the moral/conventional distinction is universal, but without further investigation it is unclear whether the division between moral and conventional acts is culturally dependent, or if the standard measures used were unable to capture the distinction in the non-WEIRD groups.

In the second stage of our study, we focused on the one non-WEIRD group that exhibited the weakest evidence for the moral/conventional distinction, Iranian children, and calibrated our measures to reflect the cultural and social practices of that group. Items for our culturally tailored measure were established by consulting local Iranian residents of varying socioeconomic status and educational backgrounds regarding commonly practiced cultural, religious, and political norms of Iranian society. Iran is a unique context for investigating the moral/conventional distinction because the government imposes legal punishments for violations that overlap with religious codes. Notably, many such religious transgressions do not typically qualify as moral violations (i.e., they do not harm another person) in the moral/conventional literature. Consequently, transgressions which might typically be classified as "conventional" are associated with both religious and legal strictures in Iran, which may explain why they are judged by children to be similar to items that are typically classified as "moral" violations.

In Study 2, we sought to differentiate two possibilities. First, compatible with the results of Study 1, it is possible that Iranian children do not make a robust moral/conventional distinction. A second possibility, however, is that Iranian children are just as likely to make a robust moral/conventional distinction as children in other cultures, but that the items in Study 1 are not a

valid test of this distinction in Iran, because the conventional items, in particular, are associated with severe religious and legal consequences. Therefore, in Study 2 we added additional measures to examine how Iranian children understand moral violations and conventional violations that do not have legal implications but go against social norms and religious practices. The new items also assessed whether children consider the cultural and religious identity of the perpetrator in their judgments, believe that some transgressions *should* be illegal or religiously forbidden, and how they justify each of their ratings. With the addition of these culturally tailored measures, Study 2 found that Iranian children, like children in other cultures, made a robust distinction between moral and conventional transgressions.

2. Study 1

The main goal of Study 1 was to ask whether a standardized battery, used primarily in research conducted in WEIRD societies, would reveal a moral and conventional distinction among children from a range of non-WEIRD cultures.

2.1 Method

2.1.1 Participants

Participants were children recruited from countries representing a diverse range of cultural groups (see **Table 3.1** in Supplemental Materials). A total of 268 children between 5 to 10 years old (M_{age} = 8.36 yrs, SD_{age} = 1.67 yrs) were tested across four research sites where one society fit the WEIRD categorization and the other three were non-WEIRD: Canada (n = 80; n female = 38, n male = 42; M_{age} = 7.81, SD_{age} = 1.59), India (n = 73; n female = 21, n male = 40, n sex unreported = 12; M_{age} = 9.52, SD_{age} = 1.32), Iran (n = 84; n female = 42, n male = 42; M_{age} = 7.65, SD_{age} = 1.61), and Korea (n = 31; n female = 12, n male = 19; M_{age} = 8.98, SD_{age} = 1.04). Sample sizes were limited by the number of children who we could access at each site. The age range of 5 to 10

years was selected based on prior research showing age-related changes in children's ability to distinguish between morality and social conventions within these years (Killen & Smetana, 1999; Nucci, Camino, & Sapiro, 1996; Smetana & Braeges, 1990; Yau & Smetana, 2003).

2.1.2 Procedures

Participants were presented with narratives of child characters engaging in three actions classified as "moral transgressions" and three actions classified as "conventional transgressions" (see Supplemental Materials for details), based on prior developmental research in WEIRD settings with children as young as 3 and adults (Lahat et al., 2012; Nucci, 1981; Nucci & Turiel, 1978; Sander son & Siegal, 1988; Siegal & Storey, 1985; Smetena 1981, 1985, 1986; Turiel, 1983; 2008). Moral items were: hitting another child, calling another child a bad name, and tearing up another child's drawing. Conventional items were: putting shoes on the lunch table, wearing a bathing suit to school, calling a teacher by their first name, and not cleaning up toys in the classroom. After each narrative was read aloud, children were asked to indicate whether the transgression was "good" or "bad," followed by a rating of the acceptability of each transgression using a 6-point Likert-scale depicted by frowning and smiling faces. Responses ranged from 1-6 with '1' indicating that the action was *very very bad* and '6' indicating that the action was *very very good*. Children were then asked to make judgments about the generalizability, rule contingency, and dependence on conformity of each transgression. These judgments were: "Would the transgression be okay if" (a) "It took place in a faraway country?" (b) "There were no rules against it?" and (c) "Everyone was doing it?" Responses to these questions were either "yes" or "no". These questions have been used by prior studies to assess the permissibility of an act, generalizability of an act's wrongness, and the act's contingency on rules (Davidson, Turiel, &

Black, 1983; Kelly, Stich, Haley, Eng, & Fessler, 2007; Nichols & Folds-Bennett, 2003; Nucci & Turiel, 1978; Smetana, 1981; Weston & Turiel, 1980; Yau & Smetana, 2003).

2.2 Results

The primary question of Study 1 was whether children make different judgments for moral and conventional violations across 4 different cultures, using measures adapted from prior moral development research.

Acceptability Ratings: To see if children from all four groups exhibit a moral/conventional distinction, we compared children's ratings of moral and conventional transgressions across the different cultural groups. For each cultural group, we constructed a linear mixed effects model with the formula: Acceptability rating (1-6) ~ Transgression(moral/conventional) * Age + (Transgression | Participant) + (Age | Item), with age centered and scaled to facilitate model fit.[1,2]

Overall, these models indicated that although children in Canada, a WEIRD group, appeared to make a robust distinction between moral and conventional transgressions, this distinction was not as robust in non-WEIRD groups (**Fig. 3.1**).

[1] All models were fit in R (version 3.6.1) using the 'lme4' package (Bates et al., 2015).

[2] For this and all models, we pre-registered the maximal effects structure (Barr et al., 2013) and iteratively pruned random effects until the model converged. For each cultural group, we report the final maximal model. For models with pre-registered interactions we also tested whether the interaction significantly improved the fit of the model by comparing it to a reduced model without that interaction via a Likelihood Ratio Test. If the interaction did not explain additional variance, we removed it from the model. Final models and outputs are reported in Supplemental Materials.

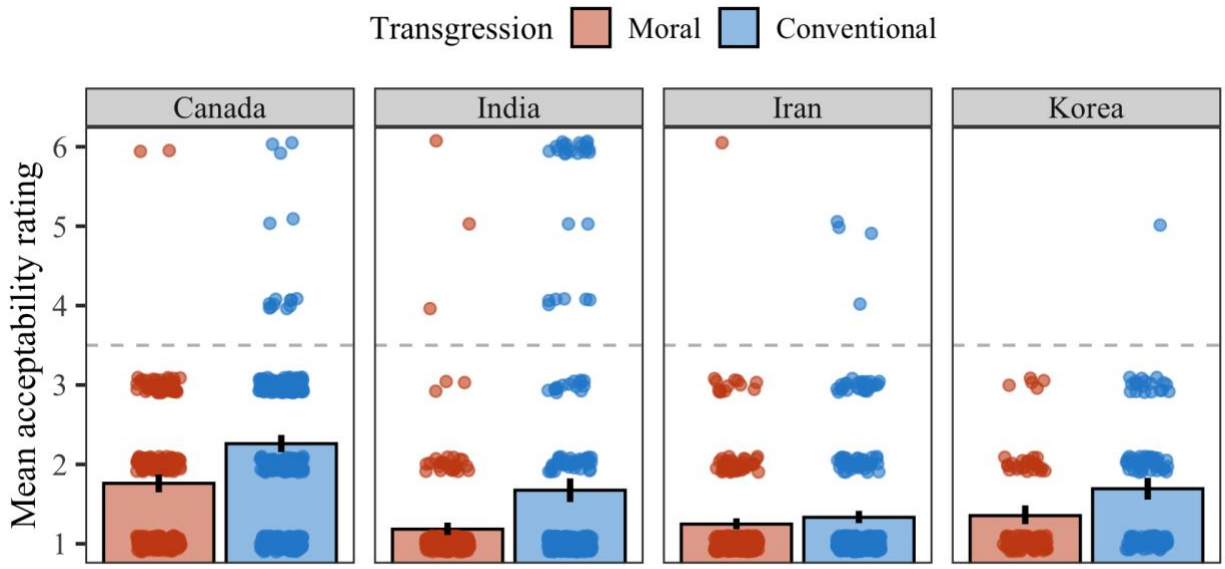

Figure 3.1. Mean acceptability ratings (1 = *very very bad*, 6 = *very very good*) for moral and conventional transgressions by testing site. Points are ratings given by individual participants for each item. Error bars indicate 95% confidence intervals computed by nonparametric bootstrap.

Canadian children rated conventional transgressions as being significantly more acceptable than moral transgressions (β = 0.50, p = .02), with older children rating all transgressions more leniently than younger children (β = 0.14, p = .02) –findings that are consistent with prior studies of WEIRD groups (Killen & Smetana, 1999; Smetana & Braeges, 1990; Smetana, Schlagman, & Adams, 1993; Tisak & Turiel, 1988). For non-WEIRD groups, this moral/conventional distinction was either smaller or nonexistent (See **Fig. 3.2** for main effects of age, transgression type, and cultural group and **Table 3.4** in Supplemental Materials for interactions between variables). Among Indian children, we found a significant interaction between transgression and age ($\chi^2_{(1)}$ = 7.71, p = .006), such that older children rated moral transgressions more severely than conventional transgressions in comparison to younger children, with no main effects of transgression type (β = 0.28, p = .11), or age (β = 0.02, p = .85). Korean children, on the other hand, did *not* rate moral transgressions more harshly than conventional transgressions (β = 0.34, p = .25), and acceptability

ratings did not increase significantly with age (β = 0.20, p = .06).[3] Finally, in Iranian children we found no significant difference between how children rated moral and conventional transgressions (β = 0.09, p = .41), although acceptability ratings for all transgressions increased significantly with age (β = 0.14, p = .002). In summary, of the four groups, only Canadian children showed a consistent pattern in rating moral transgressions more severely than conventional transgressions.

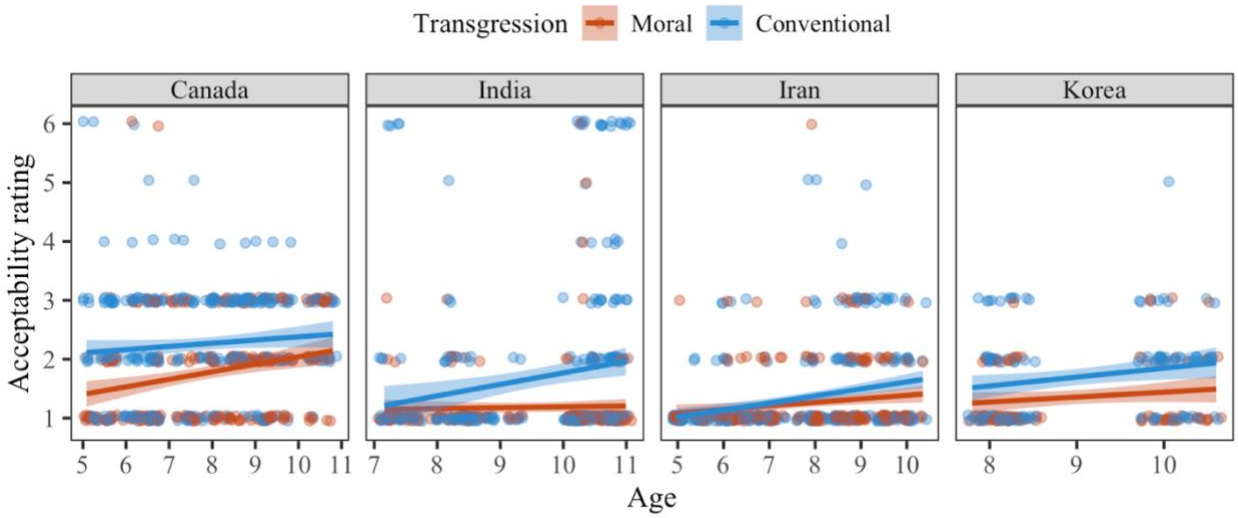

Figure 3.2. Mean acceptability ratings (1 = *very very bad*, 6 = *very very good*) for moral and conventional transgressions by age and testing site.

Analysis of Circumstances: We next assessed the evidence for a moral/conventional distinction within each culture by examining whether the circumstances surrounding a transgression affected its permissibility – a key measure of whether items labeled as "conventional" were indeed understood by children to be variable across cultures and contexts,

[3] Although Korean children did not appear to make a distinction between moral and conventional transgressions, a *post-hoc* exploration revealed that this was due to variance in children's ratings at the item-level: Korean children rated one conventional item (calling teacher by their first name) much lower (M = 1.06) in comparison to other conventional items (M range = 1.87 - 1.97), potentially reflecting Korean culture's emphasis on politeness and respect for elders (Stadler, 2011; Song et al., 1987). When the item-level random intercept was removed from this model we found a significant main effect of transgression, with Korean children rating conventional items lower in severity than moral items (β = 0.34, p < .001). Removing the random intercept for items did not significantly impact results for any other cultural group.

and not governed by a universal moral code. Similar to prior work (e.g., Hollos et al., 1986; Song et al., 1987; Yau & Smetana, 2003; Zimba, 1994), we asked questions to probe the generalizability (e.g., "Would it be okay in a faraway country?"), rule contingency ("Would it be okay if there were no rules against it?"), and dependence on conformity ("Would it be okay if everyone was doing it?") for each moral and conventional item. If children make a moral/conventional distinction, they should find moral items unacceptable regardless of the circumstance under which they occur, but judge that conventional items may be acceptable in some contexts, or for some groups. To test this question, within each culture we constructed a generalized linear mixed effects model with the formula: Acceptability (1/0) ~ Transgression*Age + (Transgression | Participant) + (Age | Item).

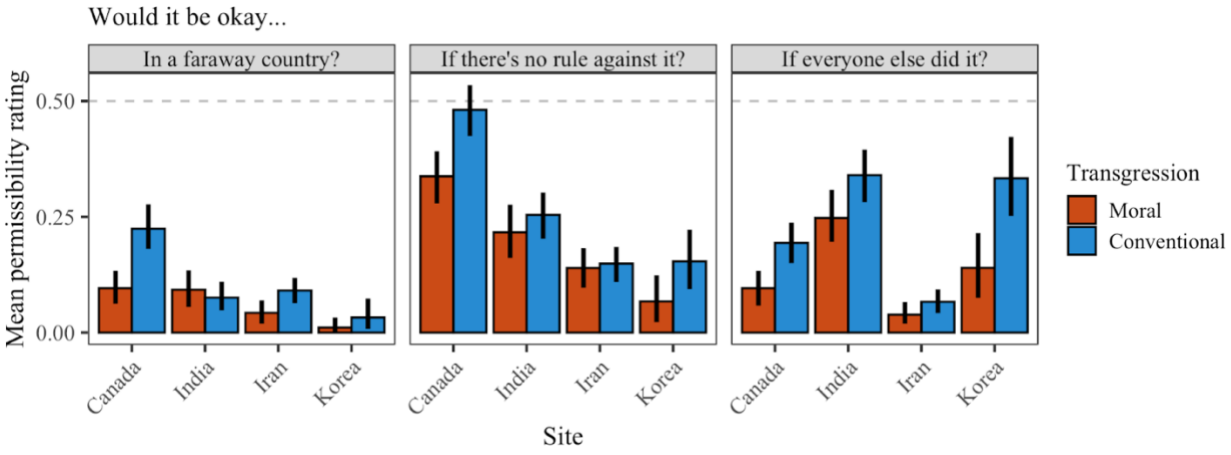

Figure 3.3. Mean permissibility ratings for each transgression type by circumstance. Error bars indicate 95% confidence intervals computed by nonparametric bootstrap.

Mean permissibility ratings for moral and conventional transgressions are shown in **Figure 3.3**. Here we found that Western children viewed the permissibility of moral transgressions as less flexible than that of conventional transgressions regardless of the circumstances under which they

occurred, but for children from non-WEIRD groups this pattern was less straightforward. Canadian children rated moral transgressions as significantly less permissible than conventional transgressions across all circumstances (all *p*s < .03).[4] In contrast, Indian children evaluated moral and conventional transgressions similarly in permissibility across all circumstances, indicating that they did not view the permissibility of either type of transgression as context dependent (all *p*s > .10; unlike the severity models, there was no significant interaction between transgression type and age in any of these models, with all *p*s > .5). In Korea, children only judged conventional transgressions as more permissible if everyone else was doing them (β = 1.73, *p* = .006), but judged moral and conventional transgressions as similarly impermissible under other circumstances (both *p*s > .10). Finally, Iranian children showed a mixed pattern, viewing moral and conventional transgressions similarly in permissibility if there was no rule against them (β = 0.64, *p* = .15), but rating conventional transgressions as marginally more permissible than moral transgressions under the circumstances that they occurred in a faraway country (β = 2.74, *p* = .05) or if everyone else was doing them (β = 1.03, *p* = .06). These findings show that Canadian children consistently view the permissibility of conventional transgressions as more context dependent than moral transgressions, but children from other cultural groups either view both types of transgressions as wrong regardless of the circumstances under which they occur, or evaluate conventional transgressions as more permissible under some, but not all, circumstances.

To complement these within-culture analyses that assessed the existence of the moral/conventional distinction within each group, we also constructed a between-culture model that treated culture as a variable. It is important to note that, though this between-culture approach is widely used in human research to make cross-cultural comparisons, it addresses a different

[4] The transgression*age interaction did not significantly improve the fit of any of the acceptability models (all *p*s > .10)

question than tests of universality. In particular, our within-culture models examine whether the moral/conventional distinction exists in all four groups, whereas the between-culture approach examines whether the distinction is stronger in some groups than others. Using a single linear mixed effects model with the formula: Acceptability rating (1-6) ~ Culture*Transgression*Age + (1 | Participant), we found a three-way interaction between site, transgression type, and age ($\chi^2_{(3)}$ = 14.71, p = .002), and *post-hoc* pairwise tests revealed that Canadian children rated both moral and conventional items higher in acceptability than Indian, Iranian, and Korean children (all ps < .001). In contrast, there was no significant difference across Indian, Iranian, and Korean children's overall acceptability ratings for either moral or conventional items (all ps > .3). These findings, in conjunction with our within-culture models, reveal that part of the group differences observed can be attributed to children from non-Western groups using the acceptability scale similarly to one another, but differently from Canadian children. These results provide further support for the notion that the standardized measures used to assess children's ratings are more suitable for capturing the moral/conventional distinction in WEIRD groups than non-WEIRD groups.

2.3 Study 1 Discussion

The results of Study 1 revealed that in our one Western society (Canada), children rated moral transgressions as significantly less acceptable and permissible under different circumstances than conventional violations—a clear pattern that replicates those of prior developmental studies with WEIRD populations (Lahat et al., 2012; Nucci, 1981; Nucci & Nucci, 1982; Sanderson & Siegal, 1988; Schmidt, Rakoczy, & Tomasello, 2012; Smetana, 1981, 1985; Smetana & Braeges, 1990; Nucci & Turiel, 1978). In contrast, the findings from our non-WEIRD societies were less straightforward: Indian and Korean children rated moral transgressions more severely than conventional transgressions; however, Indian children failed to make a moral/conventional

distinction when asked about the permissibility of transgressions under different circumstances, while Korean children only made this distinction in instances where everyone else was committing the transgression. Notably, Iranian children did not appear to make a moral/conventional distinction in their acceptability ratings and judged conventional transgressions as only slightly more permissible if they occurred in a faraway country or if everyone else was doing them. These findings suggest that the standardized battery that we used to test for the moral/conventional distinction reveals the clearest distinction in Canadian children, but less straightforward evidence of the distinction in children from non-WEIRD groups.

These results are compatible with at least two interpretations. First, it is possible that children from non-WEIRD societies do not make as clear a moral/conventional distinction as children from WEIRD groups, contrary to universalist claims. Also possible, however, is that children in non-WEIRD groups make an equally strong moral/conventional distinction, but the measures used in this study are not valid for capturing this distinction. To probe these possibilities, Study 2 investigated whether children from Iran, the group where we found the weakest evidence for the moral/conventional distinction, show the same pattern of behavior as observed in Study 1 with a measure that is tailored directly to their cultural practices and beliefs.

3. Study 2

To test whether the findings of Study 1 challenge universalist claims, or instead reflect the poor validity of the measures used in non-WEIRD groups, in Study 2 we focused on Iranian children, who showed weak evidence of the moral/conventional distinction. In Iran, moral codes and cultural practices are often linked to religious rules, which are taught to children in schools, are legally imposed by Iran's Islamic government, and are enforced day to day in each neighborhood by a religious police force known as the 'morality police' (Golkar, 2011; Mehran,

1989; 2010; Mirhadi, 1997). Consequently, children may interpret the violation of some cultural and religious conventions as having moral implications. Related to this, Study 1 included "conventional" items that could have legal repercussions in Iranian society (e.g., wearing a bathing suit to school) or could be interpreted as violating the welfare and respect of an elder (e.g., calling a teacher by their first name), which potentially influenced how children evaluated conventional items.

To examine whether children make a distinction between moral and conventional transgressions when they do not overlap with additional concerns like religious or legal codes, Study 2 contrasted items from Study 1 (standard moral and conventional items) with two new sets of items tailored to Iran: The first new set were violations or religious conventions that did not have legal implications (e.g., telling a joke during prayer), which we refer to as religious transgressions, and the second set were violations of Iranian conventions (e.g., turning one's back to an elder) that had neither religious nor legal implications, which we refer to as violations of Iran-specific conventions. We expected that if Iranian children make a moral/conventional distinction, then the culture-specific conventional items should be treated differently from the conventional items from Study 1.

We also included five additional follow up questions to each item to see whether children distinguish the permissibility of the different transgression types based on the circumstances under which they occur and the identity of the violator. In particular, children were asked whether it would be okay for (1) a foreigner and (2) a non-Muslim to commit the transgression in order to see whether violations of cultural and religious norms are viewed as more permissible for foreigners and non-Muslims, whereas moral violations might be viewed as wrong regardless of the violator's identity. Also, children were asked whether a transgression (3) is a sin (4) should be

illegal, and (5) is rude/socially unacceptable, to probe whether children believed the items were related to religious codes, should carry legal consequences, or if they viewed the behaviors as what Iranians might classify as *zesht* (roughly, bad because they occurred in the presence of others).

3.1 Method

3.1.1 Participants

A total of 63 children between 5 and 14 years of age (M_{age} = 10.13 years, SD_{age} = 2.28 years; *n* female = 30, *n* male = 33) were recruited from Tehran and Karaj, Iran (See Supplemental Materials for additional details).

3.1.2 Procedures

Children in Study 2 were presented with all of the moral and conventional items from Study 1 (standard moral and conventional items) and additional items that were tailored to the conventions and religious practices of Iranian society. These included four new items that described violations of religious practices that had no legal repercussions (e.g., not washing up for prayer, eating in front of individuals who are fasting, wearing red during a holy month of mourning, and telling a joke during prayer) and four that involved non-religious violations of conventions specific to Iranian society that had neither religious nor legal implications (e.g., turning one's back to an elder, wearing house slippers to a party, wearing shoes inside another person's home, and drinking water in front of guests without offering any to them). Similar to Study 1, after children rated these items in acceptability on a scale of 1-6 (1 = *very very bad* to *6 = very very good*), they judged the generalizability, rule contingency, and dependence on conformity for each item by responding with a "yes" or "no," which were coded as 1 or 0 respectively.

Children were asked to respond "yes" or "no" to five additional questions about each transgression: Would it be permissible for (1) foreigners or (2) non-Muslims, and if the transgression was (3) a sin, (4) should be legally punishable, and (5) rude or socially unacceptable? These follow up questions allowed us to ask whether children differentiate between cultural and religious items and if they evaluate different items on the basis of rudeness, potential legality, or both. Finally, participants were asked to explain why they rated an item as good or bad in an open response format–a method that has been used in prior development studies to further understand how children reason about different transgressions and to identify common themes in children's responses (Davidson, Turiel, & Black, 1983; Hollos et al., 1986; Kahn, 1992; Nucci & Weber, 1995; Song et al., 1987; Weston & Turiel, 1980). These justifications were recorded in the form of short statements and categorized into common themes by four raters who achieved high inter-rater agreement (see Supplemental Materials for additional information).

3.2 Results

Our primary question in this study was whether Iranian children differentiate between different transgression types when these transgressions more closely reflect cultural norms.

Acceptability Ratings: We first examined whether children's acceptability ratings of standard conventional, Iran-specific conventional, and religious transgressions differed from their ratings of standard moral transgressions by constructing a linear mixed effects model with the following formula: Acceptability rating (1-6) ~ Transgression type + Age + (1 | Participant). We used a maximal random effects structure, and iteratively removed coefficients until the model converged (see Supplemental Materials for details). Children's ratings for these four transgression types are shown in **Figure 3.4**. We found that once again, Iranian children did not rate the standard conventional transgressions from Study 1 differently from moral transgressions ($\beta = 0.06$, $p = .41$).

However, children *did* rate Iran-specific conventional transgressions as significantly more acceptable than moral transgressions (β = 0.44, p < .001). The model also revealed that Iranian children did not rate religious transgressions differently from moral transgressions (β = 0.06, p = .41), suggesting that they perceive violations of religious conventions to be on par with moral violations.

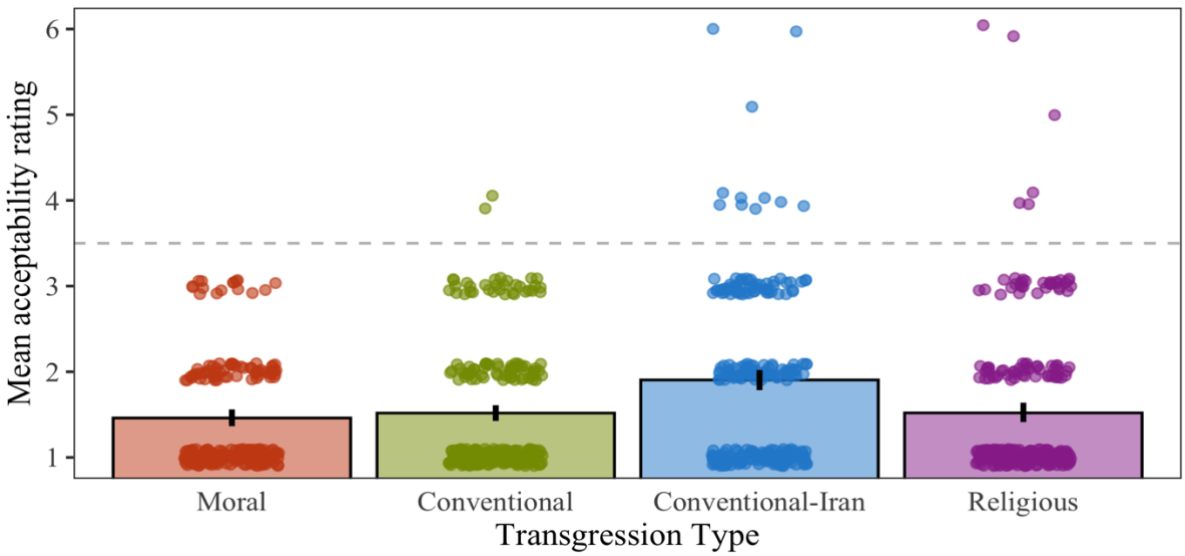

Figure 3.4. Mean acceptability ratings for each transgression type (1 = *very very bad*, 6 = *very very good*). Points are ratings given by individual participants for each item. Error bars indicate 95% confidence intervals computed by nonparametric bootstrap.

Analysis of Circumstances: As in Study 1, we next tested whether children's judgments of different transgressions were affected by the hypothetical circumstances under which they took place, another measure of whether children make a moral/conventional distinction. We analyzed children's responses to the three circumstances that were included in Study 1 regarding the generalizability of an act ("Is it okay in a faraway country?"), rule contingency ("Is it okay if there is no rule against it?"), and whether an act involves conformity ("Is it okay if everyone else is doing it?") using a generalized linear mixed effects model with the formula: Permissibility rating

(0/1) ~ Transgression*Age + (Transgression | Participant) + (Age | Item). Children's mean permissibility ratings for these questions are shown in **Figure 3.5**.

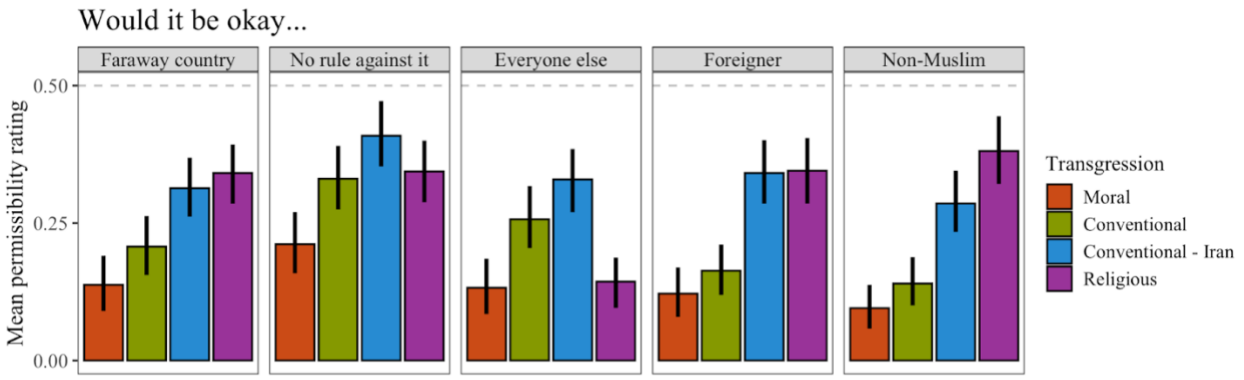

Figure 3.5. Mean permissibility ratings (1 = *yes*, 0 = *no*) for each transgression type by circumstance (Faraway country = *In a faraway country*, Everyone else = *Everyone else committing it*, Foreigner = *Foreigner committing it*, Non-Muslim = *Non-Muslim committing it*). Error bars indicate 95% confidence intervals computed by nonparametric bootstrap.

We found that children rated transgressions of Iran-specific conventions as significantly more permissible than moral transgressions under all circumstances (all *p*s < .001). However, in contrast to Study 1, we found that children rated standard conventional transgressions as more permissible than moral transgressions under all circumstances (all *p*s < .03). This finding suggests that, although children in Study 2 rated standard moral and conventional transgressions similar in severity, they do make some distinction in the permissibility of these violations and view standard conventional transgressions as more context dependent than moral transgressions. This could be because children in Study 2 were asked about a wider range of transgressions and circumstances, which could have shaped the relative permissibility of different transgressions. Finally, although children rated religious transgressions equal in severity to moral transgressions, they judged religious transgressions as more permissible compared to moral transgressions if they occurred in a faraway country or if there was no rule against them (both *p*s < .002), but not if everyone else

was committing them (p = .70) – indicating some awareness that religious violations are more contingent on the laws, norms, and religious practices of the society in which they occur, whereas the permissibility of moral violations is less context dependent . In each scenario, children's permissibility ratings increased with age (all ps < .002). Together, these findings provide some indication of how children distinguish between the permissibility of standard items and both culturally tailored conventional and religious transgressions.

We further examined whether children make a distinction between different transgression types by asking if they considered the violator's cultural identity in their ratings. To do this, we used the previous model to test whether children's permissibility judgments differed if the transgression was committed by either a foreigner or a non-Muslim. Interestingly, we found that the transgression*age interaction was significant in both circumstances (Foreigner: $\chi^2_{(3)}$= 12.32, p = .006; Non-Muslim: $\chi^2_{(3)}$= 15.06, p = .002) suggesting that, with age, children become increasingly aware that some transgressions are more permissible if they are committed by members outside of their cultural and religious group. In circumstances where the transgression was committed by a foreigner, children judged both religious and Iran-specific conventional (but not standard conventional, β = 0.54, p = .13) transgressions as more permissible than moral transgressions overall (both ps < .001). The significant transgression*age interaction in this model was driven by religious transgressions, which older children judged as significantly more permissible than moral items in comparison to younger children (β = 0.78, p = .02). In circumstances where the transgression was committed by non-Muslims, children once again determined that Iran-specific conventional and religious transgressions (but not standard conventional transgressions, β = 0.38, p = .30) were more permissible overall in comparison to moral transgressions (both ps < .001). In this model, the transgression*age interaction was

significant for all transgression types, with older children judging standard conventional, Iran-specific conventional, and religious transgressions as more permissible than moral transgressions in comparison to younger children (all *p*s < .03). Together, the findings from children's judgments about the permissibility of different transgression types when committed by foreigners and non-Muslims suggest that children restrict the scope of these norms to Iranians and Muslims, which aligns with past work showing that children do not generalize religious norms to foreigners or, more generally, followers of another religion (Srinivasan et al., 2019).

Judgments of transgression types: To further understand the rationale behind children's judgments of particular transgression types and to test whether our culture-specific and religious items were interpreted by children as belonging to these distinct categories, we adapted the previous model to analyze children's responses (yes = 1, no = 0) to questions of whether a transgression is a sin, should be illegal, or is rude. Our findings reveal that children did differentiate between different transgression types: First, children were far more likely to say that religious transgressions were a sin in comparison to all other transgression types (all *p*s < .001; see **Fig. 3.6**). Second, children were more likely to state that moral transgressions should be illegal compared to Iran-specific conventional transgressions (β = 1.50, p < .001). However, we *also* found that children were just as likely to say that standard conventional transgressions should be illegal as moral transgressions (β = -0.03, p = .90). This finding bolsters the hypothesis that children in both Study 1 and Study 2 failed to uniformly distinguish between standard moral and conventional transgressions because these conventional transgressions may have more severe ramifications in non-WEIRD cultures.

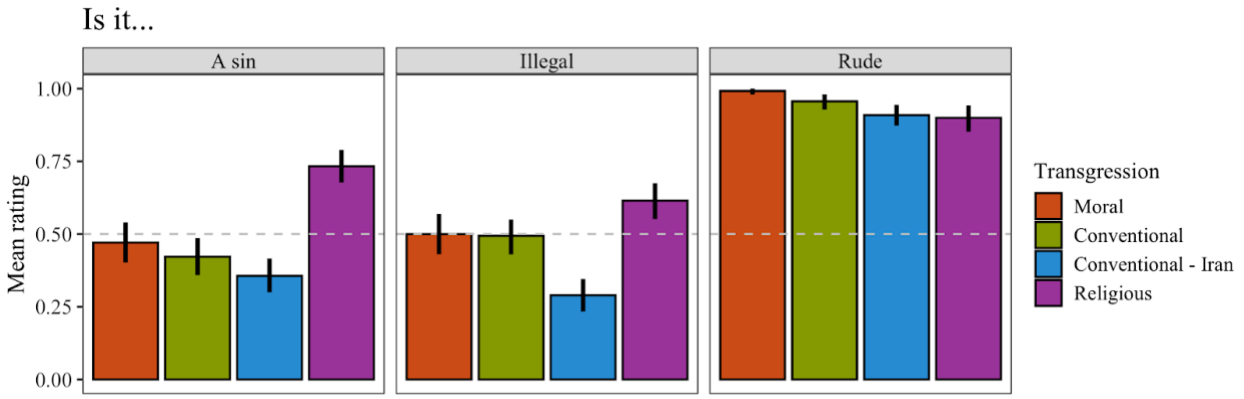

Figure 3.6. Mean ratings for each transgression type by question (1 = *yes*, 0 = *no*). Error bars indicate 95% confidence intervals computed by nonparametric bootstrap.

Interestingly, however, we also found that children were significantly more likely to state that religious transgressions should be illegal in comparison to moral transgressions (β = 0.78, p = .002). The finding that children found moral transgressions less deserving of illegality than religious transgressions likely reflects both that religious transgressions frequently have legal ramifications in Iran, as well as the fact that the moral transgressions, while causing injury to another individual, do not carry similar legal consequences. Finally, contrary to our predictions, children judged Iran-specific conventional transgressions as significantly *less* rude than both moral transgressions (β = -2.85, p < .001) and standard conventional transgressions (β = -0.98, p = .02), but equally rude in comparison to religious transgression (β = 0.07, p = .83), suggesting that they based their perception of "rudeness" on factors other than the extent to which an act violated social norms, expectations and order. In all these models, the likelihood of judging that a transgression was a sin, should be illegal, or was rude decreased with age (all ps < .03). Together, these findings provide an initial indication that even though children may treat some transgressions as equally bad in terms of severity, their reasons for doing so differ by transgression types. To better understand children's reasoning behind the different transgression types, we asked children to provide justifications beyond the yes/no responses to these three questions.

Justifications of ratings: Our previous analyses reveal some indication that Iranian children make distinctions between certain transgression types, but there are still some unresolved questions as to why children make these distinctions. In our final analysis, we pursued this question by using a thematic analysis to identify themes in children's justifications for their ratings of the different transgression types (See **Table 3.12** in Supplemental Materials for category descriptions). Three raters coded all justifications into seven categories, similar to those found in prior research (Davidson et al., 1983; Hollos et al., 1986; Nucci, 1981; Smetana, 1985; Song et al.,1987), and the category that two thirds of raters selected as the best fit for each justification was used. Interrater reliability was 97.53% (see Supplemental Materials for details). For all transgression types, children referenced disrespect at a high rate, and this was especially the case for culture-specific conventions (see **Fig. 3.7**). Apart from disrespect, children's justifications differed by transgression type, such that they most frequently cited injury as the justification for their ratings of moral transgressions, disruption to social order for both standard and Iran-specific conventional transgressions, and religious reasons for religious transgressions. These findings further show that when culturally tailored items were used, children in Study 2 made a clear distinction between different transgression types, the circumstances under which they are permissible and the reasons for why they rated the distinct transgression types differently.

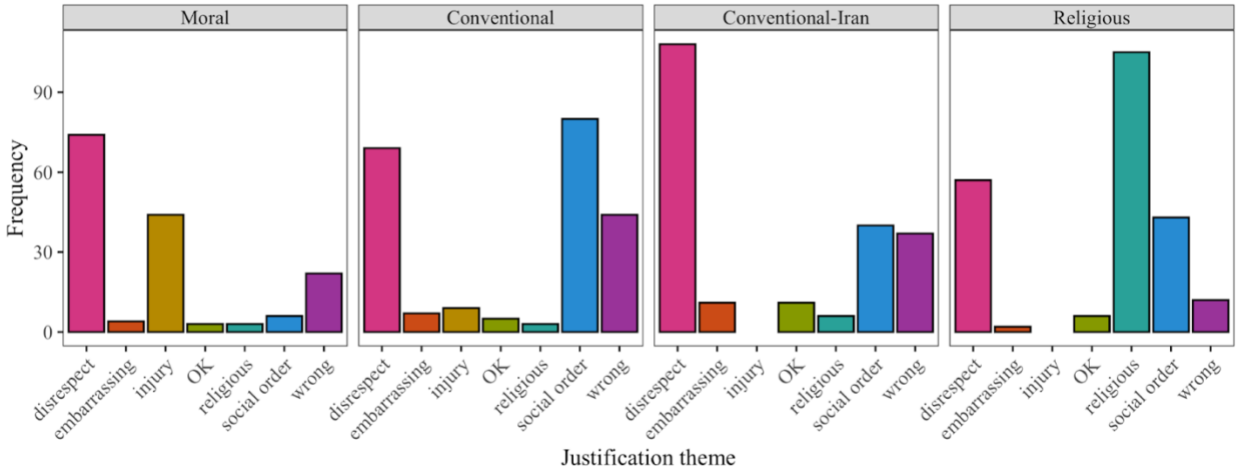

Figure 3.7. Frequency of themes referenced in children's justifications for their ratings by transgression type. See Table 3.12 in Supplemental Materials for explanation of each justification theme.

3.3. Study 2 Discussion

The goal of Study 2 was to ask whether Iranian children, who did not show evidence of a moral/conventional distinction in Study 1, exhibit this distinction when items are tailored to local cultural practices. Our results revealed that when conventional and religious items did not overlap with legal concerns, as they did in Study 1, Iranian children differentiated between the different transgression types. We also found that children evaluated the standardized conventional items from Study 1 similarly to moral items with regard to whether these violations should be illegal. This finding highlights how the conventional items taken from the standardized measure can overlap with other concerns for Iranian children, such as having more severe consequences in Iranian society.

4. General Discussion

A central question to the field of psychology is how culture impacts cognition, and cross-cultural studies have been essential to answering this question. Also crucial is the inclusion of

more non-WEIRD groups to this work, since most psychological research has oversampled WEIRD populations. In the current study we tested whether the moral/conventional distinction is present in children who belong to both WEIRD and non-WEIRD groups. To achieve methodological consistency and allow apples-to-apples comparison of behaviors across groups, it is common for researchers to export experimental measures developed with Western groups to study non-Western groups. However, a problem with this approach is that it hinges on the assumption that methods that are developed in one culture measure the same constructs when deployed in another culture. The results from the current study reveal that when children across cultures are testing using a single set of items, we find significant variability between groups, but that when items are tailored to a specific culture, evidence for a moral/conventional distinction emerges even in groups where it is otherwise not detectable. These findings provide important context for evaluating past studies that find variability in measures of social cognition and raise the possibility that the use of a single, untailored, method across cultures may inflate the appearance of cross-cultural differences.

In Study 1, using standard measures that have been widely used in prior developmental work, we found clear evidence of the moral/conventional distinction in Canadian children, but not in children from non-Western groups. Compatible with past studies (Nucci, 2001; Smetana & Braeges, 1990; Smetana et al., 2012; Tisak & Turiel, 1988), Canadian children consistently rated moral transgressions more severely and as less permissible under circumstances having to do with rules, authority, and conformity than conventional transgressions. In contrast, Korean, Indian, and Iranian children did not exhibit a clear pattern of the moral/conventional distinction. For instance, only older Indian children rated moral transgressions more severely than conventional transgressions, whereas Korean and Iranian children did not. For Korean children, the conventional

violation of speaking impolitely to one's teacher was rated to be as unacceptable as moral violations These results highlight how some violations that have been deemed by prior work as "conventional" can be interpreted as having moral consequences in non-WEIRD cultures.

We also did not find a clear developmental pattern among children from the non-WEIRD groups. While Canadian children showed similar patterns to those found in prior developmental studies, with older children rating transgressions more leniently than younger children (Ardila-Rey & Killen, 2001; Killen & Smetana, 1999; Midgette et al., 2016; Turiel, 1983), Indian children showed no age-related changes in their ratings, and Korean and Iranian children's acceptability ratings decreased with age (unlike the Canadian sample). Further developmental work is needed to understand why children from some cultural groups, but not others, evaluate violations more harshly as they approach adolescence and whether this changes in adulthood. Altogether, our results from Study 1 suggest that children's ratings of moral and conventional violations are shaped by cultural factors, and violations that have been defined as "conventional" in Western literature can be interpreted as having moral implications in non-Western societies. In Study 2, we further examine how cultural factors such as societal structure and religious laws influence how children from one non-Western society reason about the moral and conventional domains.

In Study 2 we found that when items were tailored to the cultural beliefs and practices of Iranian society and did not overlap with other concerns (as they did in Study 1), children showed a robust moral/conventional distinction. Children rated Iran-specific conventions as more acceptable and permissible under various circumstances than moral and religious transgressions. Also, children's ratings and justifications of the culturally tailored items revealed that these were interpreted by children as belonging to the distinct categories we had anticipated (e.g., moral, conventional, and religious). In agreement with theories of the moral/conventional distinction,

children's justifications revealed that their judgments of moral violations were based on concerns for the welfare of others, whereas their judgements of violations of Iran-specific conventions were based on concerns for preserving social order. One interesting finding to note is that, though we selected items that do not have legal ramifications in Iran, children still viewed violations of religious practices as more deserving of illegality than moral transgressions, suggesting that they extend their knowledge of Iran's religious laws to religious practices more generally. Past work with Muslim Arab children shows a similar pattern, with children indicating that a law should exist for some social violations (Nisan, 1987). Further work is needed to understand why Iranian children believe that religious violations (that do not impinge on the welfare of others) are worthy of legal sanctions and the potential role that residing under a religious government plays in shaping these beliefs.

The findings from Study 1 and 2 provide evidence of the moral/conventional distinction in WEIRD and non-WEIRD children, which supports universality claims of the distinction, but our cross-cultural findings also suggest that social and cultural factors shape how children reason about the different transgression types. Though our findings do not speak to *why* children from India, Iran and Korea did not make as clear of a distinction between moral and conventional violations as Canadian children, they do provide insight into why the standard used measure of the moral/conventional distinction fell short for Iranian children and was not able to sufficiently detect the distinction in all groups. Future work should examine whether other factors that were not assessed in the current work (e.g., influence of family and institutions, socioeconomic status, and other cultural norms) play an influential role in children's development of the moral/conventional distinction and how these factors differ across cultural groups. Further, while our model yielded significant findings, our sample was limited to children residing in urban and developed cities,

thus limiting the generalizability of our results. Future work should examine whether other types of methods and measures are better suited for assessing the moral reasoning of children residing in more remote and rural societies, where cultural norms and practices may be vastly different from those assessed in the current study.

The findings from our study reveal some of the issues that can arise when standardized measures alone are used to make cross-cultural comparisons. These methodological issues have been raised before for cross-cultural tests of IQ and aptitude, memory, and personality (Chen, 2008; Irvine & Carroll, 1980; McDowell, 1992; Shuttleworth-Edwards, 2016), but less work has been conducted in the domain of moral reasoning. The benefit of standardized measures is that they allow researchers a uniform method for replicating studies, minimizing bias in the testing process, and producing results that are generalizable to a larger population, beyond the group tested. However, as observed in Study 1, when items on a standardized measure have low cross-cultural equivalence, or carry different meaning for each culture, then the comparison of responses across different cultures can be misleading (Chen, 2008; Hambleton, Merenda, & Spielberger, 2004; Hui & Triandis, 1985; Van de Vijver & Leung, 2000). For instance, our findings from Study 1 alone might lead us to incorrectly conclude that the moral/conventional distinction is nonexistent or weaker in some cultures than others, but our findings from Study 2 suggest that instead the standard measures used may not have been sufficiently valid to capture the moral/conventional distinction in children from non-Western cultures. One reason for this is that when methods are developed primarily for WEIRD groups and then used to test for cross-cultural variability, the results obtained can skew more in favor of the behaviors and cognitions of the majority group. This practice, which is common in psychological research, can be problematic if findings are used

to make claims about groups or the universality of human behaviors (Mason, 2005; McDowell, 1992; Olmedo, 1981).

To help improve the reliability and cultural validity of measures in cross-cultural research, we recommend that researchers establish a standard procedure for collecting and analyzing data across groups and operationalize the concepts being tested (e.g., defining moral and conventional domains), but tailor the specific questions and items in the measures to the local practices and beliefs of the cultural group in question. In the current study we were able to achieve generalizability in our findings by implementing a uniform method of data collection and analysis for all cultural groups. To ensure that items reflect the current cultural and societal climate, it is important that local researchers, experts, and residents be consulted in the process of adapting measures. Therefore, in the second stage of our study we consulted local experts to develop items that were more reflective of the norms and practices of Iranian participants, but still adhered to the conceptual and operational definition of moral and conventional violations. This two-stage model allowed us to achieve both reliability in our methods and better cultural validity in our measures. We also advocate for the use of within-group models over between-group models to make group comparisons, since within-group models allow researchers to first confirm if individuals from the same group interpret measures in a conceptually similar way before comparing the pattern of findings among different groups (Boer, Hanke, & He, 2018; Fischer & Poortinga, 2018; Milfont & Fischer, 2010).

The current study is the first to our knowledge to examine the moral/conventional distinction in WEIRD and non-WEIRD groups of children using both standardized and culturally tailored measures. Though our results do not settle the debate of whether development of the moral/conventional distinction is universal or culture specific, our findings do reveal that a

moral/conventional distinction is more likely to be captured in non-WEIRD groups when measures are tailored to the cultural norms and codes of the group in question. Our results from children residing in non-Western societies also point to the importance of considering the role of cultural, religious, and legal codes when assessing attitudes toward moral and conventional transgressions and comparing moral judgments across cultures. Importantly, our findings highlight how a two-stage approach that we advocate for conducting cross-cultural research can address issues of generalizability and cultural validity within a single study.

Acknowledgments

We are grateful to the parents, children, teachers and school directors who made it possible for us to conduct this research in Canada, India, Iran and Korea. We are particularly grateful to Golshan Shirmohammadali, Fateme Mohammadi, and Samaneh Hamidifar for their help in recruiting families and collecting data in Iran. We also thank Yusuf Kapadia, Priyanka Dixit, and Pallavi Maity Uniyal at Zenith School in Vadodara, India for their help with data collection. Our research in Comox, B.C. would not have been possible without the help of Junyi Chu and the support of the directors and teaching staff at Courtenay Elementary School, École Au-coeur-de-l'île, and École Robb Road Elementary School. We are also thankful to the teaching staff of Myeondong Elementary school in Korea for their assistance with data collection.

Chapter 3, in part is currently being prepared for submission for publication of the material. Yazdi, H. Schneider, R.M., Yoon, E.J., Frank, M.C., Srinivasan, M., Dunham, Y. & Barner, D. The dissertation author was the primary investigator and author of this paper.

References

Apfelbaum, E. P., Phillips, K. W., & Richeson, J. A. (2014). Rethinking the baseline in diversity research: Should we be explaining the effects of homogeneity?. *Perspectives on Psychological Science*, *9*(3), 235-244.

Ardila-Rey, A., & Killen, M. (2001). Middle class Colombian children's evaluations of personal, moral, and social-conventional interactions in the classroom. *International Journal of Behavioral Development*, *25*(3), 246-255.

Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist*, *63*, 602–614.doi: 10.1037/0003-066X.63.7.602

Barrett, C.H., Peterson, C. D., & Frankenhuis, W. E. (2016). Mapping the cultural learnability landscape of danger. *Child Development*, *87*(3), 770-781.

Blake, P. R., McAuliffe, K., Corbit, J., Callaghan, T. C., Barry, O., Bowie, A., Kleutsch, L., Kramer, K.L., Ross, E., Vongsachang, H., Wrangham, R., & Warneken, F. (2015). The ontogeny of fairness in seven societies. *Nature*, *528*(7581), 258-261.

Boer, D., Hanke, K., & He, J. (2018). On detecting systematic measurement error in cross-cultural research: A review and critical reflection on equivalence and invariance tests. *Journal of Cross-Cultural Psychology*, *49*(5), 713-734.

Callaghan, T., Rochat, P., Lillard, A., Claux, M. L., Odden, H., Itakura, S., Tapanya, S., & Singh, S. (2005). Synchrony in the onset of mental-state reasoning: Evidence from five cultures. *Psychological Science*, *16*(5), 378-384.

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology*, *95*(5), 1005.

Chernyak, N., Kang, C., & Kushnir, T. (2019). The cultural roots of free will beliefs: How Singaporean and US Children judge and explain possibilities for action in interpersonal contexts. *Developmental Psychology*, *55*(4), 866.

Chernyak, N., Kushnir, T., Sullivan, K. M., & Wang, Q. (2013). A comparison of American and Nepalese children's concepts of freedom of choice and social constraint. *Cognitive Science*, *37*(7), 1343-1355.

Crane, D. A., & Tisak, M. S. (1995). Mixed-domain events: The influence of moral and conventional components on the development of social reasoning. *Early Education and Development*, *6*(2), 169-180.

Dahl, A. and Killen, M., 2018. Moral reasoning: Theory and research in developmental science. In *The Stevens' handbook of experimental psychology and cognitive neuroscience* (Vol. 4, pp. 323-353). Wiley New York.

Davidson, P., Turiel, E., & Black, A. (1983). The effect of stimulus familiarity on the use of criteria and justifications in children's social reasoning. *British Journal of Developmental Psychology, 1*(1), 49-65.

Dunham, Y., Srinivasan, M., Dotsch, R., & Barner, D. (2014). Religion insulates ingroup evaluations: The development of intergroup attitudes in India. *Developmental Science, 17*(2), 311-319.

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W. A., Pitcairn, T., Ricci-Bitti, P. E., Scherer, K., Tomita, M., & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, *53*(4), 712-717.

Fischer, R., & Poortinga, Y. H. (2018). Addressing methodological challenges in culture-comparative research. *Journal of Cross-Cultural Psychology*, *49*(5), 691-712.

Fu, G., Heyman, G. D., & Lee, K. (2016). Learning to be unsung heroes; Development of reputation management in two cultures. *Child Development*, *87*, 689-699.

Fung, H. (1999). Becoming a moral child: The socialization of shame among young Chinese children. *Ethos*, *27*(2), 180-209.

Goh, J. S., Hebrank, A. C., Sutton, B. P., Chee, M. L., Sim, S. Y., & Park, D. C. (2013). Culture-related differences in default network activity during visuo-spatial judgments. *Social Cognitive and Affective Neuroscience*, *8*, 134-142.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*(1), 98-116.

Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog?. *Journal of Personality and Social Psychology*, *65*(4), 613.

Halberstadt, A. G., & Lozada, F. T. (2011). Emotion development in infancy through the lens of culture. *Emotion Review*, *3*(2), 158-168.

Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (2004). *Adapting educational and psychological tests for cross-cultural assessment*. Psychology Press.

Hardin, E. E., Robitschek, C., Flores, L. Y., Navarro, R. L., & Ashton, M. W. (2014). The cultural lens approach to evaluating cultural validity of psychological theory. *American Psychologist*, 69, 656–668.

Hauser, M., Lee, J., & Huebner, B. (2010). The moral-conventional distinction in mature moral competence. *Journal of Cognition and Culture*, *10*(1-2), 1-26.

Helwig, C. C., Tisak, M. S., & Turiel, E. (1990). Children's social reasoning in context: Reply to Gabennesch. *Child Development*, *61*(6), 2068-2078.

Henrich, J., Heine, S.J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Science*, *33*(2-3), 61-83.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., & Lesorogol, C., (2006). Costly punishment across human societies. *Science, 312*(5781), pp.1767-1770.

Hollos, M., Leis, P. E., & Turiel, E. (1986). Social reasoning in Ijo children and adolescents in Nigerian communities. *Journal of Cross-Cultural Psychology, 17*(3), 352-374.

House, B. R., Silk, J. B., Henrich, J., Barrett, H. C., Scelza, B. A., Boyette, A. H., Hewlett, B.S., McElreath, R., & Laurence, S. (2013). Ontogeny of prosocial behavior across diverse societies. *Proceedings of the National Academy of Sciences*, *110*(36), 14586-14591.

Hui, C. H., & Triandis, H. C. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of Cross-cultural Psychology, 16*(2), 131-152.

Irvine, S. H., & Carroll, W.K. (1980). Testing and assessment across cultures: Issues in methodology and theory. In H. C. Triandis & J. W. Berry (Eds.), Handbook of cross-cultural psychology: Methodology (Vol. 2). Boston: Allyn & Bacon.

Kahn Jr, P. H. (1992). Children's obligatory and discretionary moral judgments. *Child Development, 63*(2), 416-430.

Kanngiesser, P., Rossano, F., Zeidler, H., Haun, D., & Tomasello, M. (2019). Children's respect for ownership across diverse societies. *Developmental Psychology*, *55*(11), 2286.

Kazemipur, A., & Rezaei, A. (2003). Religious life under theocracy: The case of Iran. *Journal for the Scientific Study of Religion*, *42*(3), 347-361.

Kelly, D., Stich, S., Haley, K. J., Eng, S. J., & Fessler, D. M. (2007). Harm, affect, and the moral/conventional distinction. *Mind & Language*, *22*(2), 117-131.

Killen, M., & Smetana, J. G. (1999). Social interactions in preschool classrooms and the development of young children's conceptions of the personal. *Child Development*, *70*(2), 486-501.

Kitayama, S., Duffy, S., Kawamura, T., & Larsen, J. T. (2003). Perceiving an object and its context in different cultures: A cultural look at new look. *Psychological Science*, *14*(3), 201-206.

Kline, M. A., Shamsudheen, R., & Broesch, T. (2018). Variation is the universal: Making cultural evolution work in developmental psychology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1743), 20170059.

Lahat, A., Helwig, C. C., & Zelazo, P. D. (2012). Age-related changes in cognitive processing of moral and social conventional violations. *Cognitive Development, 27*(2), 181-194.

Legare, C. H., & Harris, P. L. (2016). The ontogeny of cultural learning. *Child Development, 87*(3), 633-642.

Mason, T. C. (2005). Cross-cultural instrument translation: Assessment, translation, and statistical applications. *American Annals of the Deaf, 150*(1), 67-72.

McDowell, C. L. (1992). Standardized tests and program evaluation: Inappropriate measures in critical times. *New Directions for Program Evaluation, 1992*(53), 45-54.

Meadon, M., & Spurrett, D. (2010). It's not just the subjects - there are too many WEIRD researchers. *Behavioral and Brain Sciences, 33*(2–3), 104-105.

Mehran, G. (1989). Socialization of schoolchildren in the Islamic Republic of Iran. Iranian Studies, *22*, 35-50.

Mehran, G. (2010). Representations of Arabs in Iranian elementary school textbooks. In A. E. Mazawi & R. G. Sultana (Eds.), World yearbook of education 2010: Education and the Arab 'world': Political projects, struggles, and geometries of power (pp. 361-381). New York, NY: Routledge.

Midgette, A., Noh, J. Y., Lee, I. J., & Nucci, L. (2016). The development of Korean children's and adolescents' concepts of social convention. *Journal of Cross-Cultural Psychology, 47*(7), 918-928.

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research, 3*(1), 111-130.

Mirhadi, T. (1997). Education viii. Nursery schools and kindergartens. In Encyclopedia Iranica. (Vol. VIII, Fasc. 2, pp. 196–199). Retrieved from http://www.iranicaonline.org/articles/education-viii-nursery-schoolsand-kindergartens

Moles, J. A., Chattopadhyay, K., Chilungu, S. W., Cohen, R., Ferreira, J. V., Fleising, U., Henry, F., Kolig, E., Lawless, R., Lefferts, H.L., Momin, A.R. Murray, G.F., Naroll, R., Partridge, W.L., Schneider, H.K., & Studdert-Kennedy, W. G. (1977). Standardization and Measurement in Cultural Anthropology: A Neglected Area [and Comments and Reply]. *Current Anthropology, 18*(2), 235-258.

Nichols, S., & Folds-Bennett, T. (2003). Are children moral objectivists? Children's judgments about moral and response-dependent properties. *Cognition, 90*(2), B23-B32.

Nielsen, M., Mushin, I., Tomaselli, K., & Whiten, A. (2016). Imitation, collaboration, and their interaction among Western and Indigenous Australian preschool children. *Child Development*, *87*(3), 795-806.

Nisan, M. (1987). Moral norms and social conventions: A cross-cultural comparison. *Developmental Psychology*, *23*(5), 719.

Norenzayan, A., & Heine, S. J. (2005). Psychological universals: What are they and how can we know? *Psychological Bulletin*, *131*(5), 763.

Nucci, L. (1981). Conceptions of personal issues: A domain distinct from moral or societal concepts. *Child Development*, *52*(1), 114-121.

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

Nucci, L., Camino, C., & Sapiro, C. M. (1996). Social class effects on northeastern Brazilian children's conceptions of areas of personal choice and social regulation. *Child Development*, *67*(3), 1223-1242.

Nucci, L. P., & Nucci, M. S. (1982). Children's responses to moral and social conventional transgressions in free-play settings. *Child Development*, *53*(5), 1337-1342.

Nucci, L. P., & Turiel, E. (1978). Social interactions and the development of social concepts in preschool children. *Child Development*, *49*(2), 400-407.

Nucci, L., & Turiel, E. (1993). God's word, religious rules, and their relation to Christian and Jewish children's concepts of morality. *Child Development*, *64*(5), 1475-1491.

Nucci, L. P., Turiel, E., & Encarnacion-Gawrych, G. (1983). Children's social interactions and social concepts: Analyses of morality and convention in the Virgin Islands. *Journal of Cross-Cultural Psychology*, *14*(4), 469-487.

Nucci, L., & Weber, E. K. (1995). Social interactions in the home and the development of young children's conceptions of the personal. *Child Development*, *66*(5), 1438-1452.

Olmedo, E. L. (1981). Testing linguistic minorities. *American Psychologist*, *36*(10), 1078.

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, *118*(1), 57.

Robbins, E., & Rochat, P. (2011). Emerging signs of strong reciprocity in human ontogeny. *Frontiers in Psychology*, *2*, 353.

Rochat, P., Dias, M. D., Liping, G., Broesch, T., Passos-Ferreira, C., Winning, A., & Berg, B. (2009). Fairness in distributive justice by 3-and 5-year-olds across seven cultures. *Journal of Cross-Cultural Psychology*, *40*(3), 416-442.

Sanderson, J. A., & Siegal, M. (1988). Conceptions of moral and social rules in rejected and nonrejected preschoolers. *Journal of Clinical Child Psychology*, *17*(1), 66-72.

Schmidt, M. F., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition*, *124*(3), 325-333.

Shuttleworth-Edwards, A. B. (2016). Generally representative is representative of none: commentary on the pitfalls of IQ test standardization in multicultural settings. *The Clinical Neuropsychologist*, *30*(7), 975-998.

Shweder, R. A., Mahapatra, M., & Miller, J. (1987). Culture and moral development. In J. Kagan &: S. Lamb (Eds.), *The emergence of morality in young children* (pp. 1-83). Chicago: University of Chicago Press.

Siegal, M., & Storey, R. M. (1985). Day care and children's conceptions of moral and social rules. *Child Development*, *56*(4), 1001-1008.

Smetana, J. G. (1981). Preschool children's conceptions of moral and social rules. *Child Development*, *52*(4), 1333-1336.

Smetana, J. G. (1985). Preschool children's conceptions of transgressions: Effects of varying moral and conventional domain-related attributes. *Developmental Psychology*, *21*(1), 18.

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

Smetana, J. G., Ball, C. L., Jambon, M., & Yoo, H. N. (2018). Are young children's preferences and evaluations of moral and conventional transgressors associated with domain distinctions in judgments? *Journal of Experimental Child Psychology*, *173*, 284-303.

Smetana, J. G., & Braeges, J. L. (1990). The development of toddlers' moral and conventional judgments. *Merrill-Palmer Quarterly*, *36*(3), 329-346.

Smetana, J. G., Jambon, M., & Ball, C. (2014). The social domain approach to children's moral and social judgments. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (2nd ed., pp. 23-45). New York, NY: Psychology Press.

Smetana, J. G., Rote, W. M., Jambon, M., Tasopoulos-Chan, M., Villalobos, M., & Comer, J. (2012). Developmental changes and individual differences in young children's moral judgments. *Child Development*, *83*(2), 683-696.

Smetana, J. G., Schlagman, N., & Adams, P. W. (1993). Preschool children's judgments about hypothetical and actual transgressions. *Child Development, 64*(1), 202-214.

Song, M. J., Smetana, J. G., & Kim, S. Y. (1987). Korean children's conceptions of moral and conventional transgressions. *Developmental Psychology*, *23*(4), 577.

Srinivasan, M., Kaplan, E., & Dahl, A. (2019). Reasoning about the scope of religious norms: Evidence from Hindu and Muslim children in India. *Child Development, 90*(6), e783-e802.

Tisak, M. S., & Turiel, E. (1988). Variation in seriousness of transgressions and children's moral and conventional concepts. *Developmental Psychology*, *24*(3), 352.

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

Turiel, E. (2006). Thought, emotions, and social interactional processes in moral development. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 7-35). Lawrence Erlbaum Associates Publishers.

Turiel, E. (2008). Thought about actions in social domains: Morality, social conventions, and social interactions. *Cognitive Development*, *23*(1), 136-154.

Turiel, E., Killen, M., & Helwig, C. C. (1987). Morality: Its structure, functions, and vagaries. *The emergence of morality in young children*, 155-243.

Van de Vijver, F., & Leung, K. (2000). Methodological issues in psychological research on culture. *Journal of Cross-Cultural Psychology*, *31*, 33-51.

Van Leeuwen, E. J., Cohen, E., Collier-Baker, E., Rapold, C. J., Schäfer, M., Schütte, S., & Haun, D. (2018). The development of human social learning across seven societies. *Nature Communications*, *9*(1), 1-7.

Wang, Q. (2016). Why should we all be cultural psychologists? Lessons from the study of social cognition. *Perspectives on Psychological Science*, *11*(5), 583-596.

Weston, D. R., & Turiel, E. (1980). Act-rule relations: Children's concepts of social rules. *Developmental Psychology*, *16*(5), 417.

Yau, J., & Smetana, J. G. (2003). Conceptions of moral, social-conventional, and personal events among Chinese preschoolers in Hong Kong. *Child Development, 74*(3), 647-658.

Zimba, R. F. (1994). The understanding of morality, convention, and personal preference in an African setting: Findings from Zambia. *Journal of Cross-Cultural Psychology*, *25*(3), 369-393.

Supplemental Materials: Table of Contents

1. Experimental approaches to testing the moral/conventional distinction

Standardized measures for testing the moral/conventional distinction

Our main goal in Study 1 was to test the universality of the moral/conventional distinction in children from WEIRD and non-WEIRD societies using measures that have been widely used in prior developmental work with participants ranging from 3 years old to adults (Lahat et al., 2012; Nucci, 1981; Nucci & Turiel, 1978; Sanderson & Siegal, 1988; Siegal & Storey, 1985; Smetena 1981, 1985, 1986; Turiel, 1983; 2008). In the child version of the moral/conventional task, an experimenter presents participants with different scenarios of a protagonist committing a moral or conventional violation. Prior research in psychology and philosophy conceptualizes moral violations as acts that negatively impact the welfare of others, whereas conventional violations are acts that violate shared and consensually agreed upon norms (Nucci; 1981; Smetana, 1981, Turiel, 1983, 2002, 2006). For Study 1, three moral items and four conventional items were selected from stimuli used in prior studies with young children (Killen & Smetana, 1999; Smetana, 1981; Smetana & Braeges, 1990; Yau & Smetana, 2003). The moral items were: (a) pushing another child, (b) calling another child a bad name, and (c) ripping another child's drawing. The conventional items were: (a) putting shoes on the lunch table, (b) wearing a bathing suit to school, (c) calling a teacher by their first name and (d) not cleaning up toys in the classroom. After each item was read out loud by an experimenter, children were asked whether what the protagonist did was "good" or "bad", and then rated how good or how bad using a six-point Likert scale represented by smiling to frowning faces (see **Figure 3.8)**. These responses were coded from 1-6 with 1 coded as *very very bad* and 6 coded as *very very good*. Next, children were asked to answer "yes" or "no" to three different questions in a fixed order: Would it be OK in a faraway country? Would it be OK if there was no rule against it? Would it be OK if everyone at school was doing

it? These permissibility judgments have been determined by prior literature to assess the generalizability of the acts wrongness, contingency on rules and contingency on conformity (Hollos, Leis, & Turiel, 1986; Song, Smetana, & Kim, 1987; Yau & Smetana, 2003; Zimba, 1994). Studies of children below the age of five have used other approaches to testing the moral/conventional distinction, such as observations of how children respond to naturally occurring transgressions and the use of individual interviews where children are asked to describe the permissibility of a character committing a moral or conventional violation (Song et al., 1987; Smetana, 1981; Smetana & Braeges, 1990). However, the majority of past research studies on the moral/conventional distinction in children over the age of five employed stimuli and methods similar to those used in the current study.

Figure 3.8. Six-point Likert scale used to measure children's acceptability ratings for each item. Faces from left to right represent: *a little good, very good, very very good, a little bad, very bad, and very very bad*.

Developing a culturally valid measure

In Study 1, we tested for the moral/conventional distinction in four cultural groups using a measure that was initially developed and validated with Western groups and has since been widely used in developmental work to assess children's moral development. We found a clear moral/conventional distinction in Canadian children, but a more complex pattern of findings in Indian, Iranian, and Korean children with Iranian children showing the weakest evidence for the distinction. In Study 2 we tested the possibility of whether the moral/conventional distinction is indeed weaker in non-Western groups, or if the standard items used in Study 1 are less culturally

valid for non-Western groups, such as Iranian children. To test these possibilities, we included all items from Study 1 to see if our results would replicate with a different group of children and added new items in Study 2 that were tailored to the social, cultural, and religious practices of Iran. In Study 1, some of the standard conventional transgressions, such as wearing a bathing suit to school, overlapped with religious and legal concerns in Iranian society, which may have led children to interpret these as having moral implications. Therefore, two of the criteria for Study 2 items were that the conventional transgressions should not (1) overlap with legal, religious, or moral concerns and (2) have legal repercussions in Iran.

We established our culturally tailored items for Study 2 by consulting with nine adults: three researchers in the field of developmental psychology who lived for a minimum of ten years in Iran and six adults currently residing in Iran. All local consultants were well-versed in Iranian society's current cultural norms and religious practices and Iran's religious laws. First, we developed a battery of 16 items that consisted of violations of norms specific to Iranian culture and religious transgressions that are not punishable by Iranian law. For the religious items, we selected religious practices that are widely prevalent in Iran and taught to children in public schools beginning in kindergarten. Our final battery consisted of four conventional items and four religious items that all of our consultants agreed are currently practiced in the majority of Iranian households, taught and enforced in public elementary schools, and are norms that four-year-olds are likely to be familiar with. The conventional items were: (a) turning one's back to an elder, (b) wearing house slippers to a party, (c) wearing shoes inside another person's home, and (d) drinking water in front of guests without offering any to them. The religious items were: (a) not washing up for prayer, (b) eating in front of fasting individuals, (c) wearing red during a holy month of mourning, and (d) telling a joke during prayer.

To see if Iranian children show a more robust moral/conventional distinction in their permissibility ratings of the different transgression types, Study 2 included five additional follow up questions to the permissibility judgment portion of the measure. See Section 5 'Study 2 items' for complete testing materials. The three questions from Study 1 were also included and presented first to children (e.g., would it be OK in a faraway country, if there were no rules against it, and if everyone else was doing it). Next children were asked whether they believed each transgression would be okay if (1) a foreigner committed it and (2) a non-Muslim committed it to see whether children differentiate between moral, conventional, and religious violations and apply them universally to all individuals, or only to Iranians and Muslims. We also asked children whether they believed the transgression (3) is a sin or haram, (4) should be illegal and (5) is "zesht" (e.g., rude/obscene). We asked children about the religious and legal implications of each transgression to see if children distinguished between violations of religious practices that do not have legal repercussions and if they believed that these violations should have legal repercussions. These questions were of interest since some of the conventional items from the standardized measure in Study 1 could have legal repercussions in Iran due to violating religious codes (e.g., not wearing modest clothing in public). Finally, we asked children whether each transgression is "zesht", a word implying that an action is rude, obscene, or socially unacceptable because it is committed in the presence of other individuals. Children responded to each question with a "yes" or "no" which were coded as 1 or 0 respectively.

An additional follow-up question that we asked children after they rated the acceptability and permissibility of each item was to provide a justification for their ratings. The use of open-ended responses is common in moral developmental work to assess how children's moral judgments vary by transgression type, and to identify different categories among children's moral

justifications (e.g., appeal to authority, appeal to welfare of others) (Davidson, Turiel, & Black, 1983; Hollos et al., 1986; Kahn, 1992; Nucci & Weber, 1995; Song et al., 1987; Weston & Turiel, 1980). We asked children why they rated an item as good or bad to see if children give different justifications for moral, conventional, and religious transgressions and if their justifications match the category we anticipated for each item that we added in Study 2.

For instance, we predicted that children would make more religious references (e.g., Allah, the Quran, heaven/hell) for religious items than for moral and culture-specific conventional items. We were also interested in seeing if there was overlap in children's justifications for different transgressions. For instance, children might reason that tearing up another child's drawing (moral item) and calling a teacher by their first name (conventional item) are equally wrong because they are committed with disregard for another person's wellbeing. High overlap in children's reasoning for moral and conventional items from Study 1 could explain why we did not find a significant difference between how children rated the different items. See Section 6 'Study 2: Analysis of justifications' for details on how children's justifications of their transgression ratings were coded and analyzed.

2. Ethical Approval and Consenting Procedure

All study procedures and protocols were approved by the University of California, San Diego's Institutional Review Board. Additional approvals for this project were secured from Yale and UC Berkeley for research teams associated with this project. Licenses and permission were obtained from government agencies in the country of testing.

For Canada and Iran data were collected from one researcher in the team, in India data were collected from four researchers in the team with the assistance of local research assistants, and in Korea data were collected from one researcher in the team.

When testing in elementary or K-12 schools in Canada, India, Iran, and Korea, consent forms were sent home to parents and legal guardians to sign and return prior to testing children. When testing in community centers and libraries in Iran, we received signed consent from parents/legal guardians who accompanied their children to the site. All materials were translated into the local language by bilingual researchers.

3. Participants

Children between the ages of 5 and 10 were recruited for this study in Canada, Iran, and Korea (see **Table 3.1** and **Table 3.2** for demographic details). This age range was selected based on prior research that shows that children undergo age-related changes in their distinctions between morality and social conventions between these years (Killen & Smetana, 1999; Nucci, Camino, & Sapiro, 1996; Smetana & Braeges, 1990; Yau & Smetana, 2003).

In India we tested all children who we had access to and excluded data collected from children over 10 years of age. In India, Iran and Korea we tested as many children as we had access to and permission to test. In Canada, we matched the participants to those tested in Iran by gender and age group. When possible, we tried to match sample sizes. However, we did not have a planned *N* for these studies. We collected data from as many child participants as possible for whom we had consent. Due to the unpredictable nature of collecting data in the field, and because we had a finite period of time in which to collect data, our research team tried to collect data from every child for whom we had consent. When possible, we tried to match children on the basis of age across different sites.

Table 3.1. Number of Subjects by Country, Gender, and Age group in Study 1

		5-6	**7-8**	**9-10**	**Total**
Canada	Female	13	11	14	80
	Male	17	17	8	
India	Female	-	7	14	73
	Male	-	21	19	
Iran	Female	15	12	15	84
	Male	16	18	8	
Korea	Female	-	9	3	31
	Male	-	9	10	

Table 3.2. Number of Subjects by Gender and Age group in Study 2

		5-6	**7-8**	**9-10**	**11-12**	**13-14**	**Total**
Iran	Female	5	8	6	6	5	63
	Male	0	9	10	10	4	

4. Experimental Protocol

Setting

In Canada, we tested children residing in Comox, British Columbia at three elementary schools within an 8-mile radius of each other (Courtenay Elementary School, École Au-coeur-de-l'île, and École Robb Road Elementary School). In India we tested children at two English medium K-12 schools (Zenith and Amrit Vidyalaya School). In Korea, we tested children in Grades 1-3 at an elementary school in Seoul, Korea (Myeondong Elementary school). In Iran, children were tested at elementary schools, community centers, and a public library in Tehran and Karaj (a suburb of Tehran).

Procedure

In India and Korea testing took place in a classroom with teachers present to support the experimenters in data collection. In Canada children were tested in small groups of four or five and in Iran children were tested in pairs or individually in a quiet room. At all testing sites, a teacher or experimenter read each item out loud and children used response sheets to record their answers. When tested in groups, children used privacy screens to prevent them from seeing how other children responded.

Consent from the participant's parent or legal guardian was secured prior to testing each participant. Participants received a small prize for partaking in the study. Participants in Canada were tested in English, participants in Iran were tested in Farsi, participants in Korea were tested in Korean and participants in India were tested in English with a translator fluent in Hindi and Gujarati present. Children were tested either in small groups or individually depending on the testing site. After completing the testing session, children received a small prize for their participation.

5. Materials

In Study 1 we used standardized moral/conventional items from prior research. In Study 2 we used the measures from Study 1 in addition to 8 new culturally tailored items (see below for details on how we determined these items). For Study 1 and 2 we randomized the versions that children received. We tried to balance the versions across gender group and age group when possible. The gender of the protagonist and order of questions were randomized.

Study 1 Items

Standard Moral/Conventional Items:

1. Michael was sitting at the lunch table at school. He then took off his shoes and put them on the table. Was what Michael did good or bad?

2. Nancy and Katie were standing outside. Nancy purposely shoved Katie! Was what Nancy did good or bad?

3. All of the children were playing outside. James decided to wear his bathing suit to school instead of his school clothing. Was what James did good or bad?

4. Sara and Lisa were playing outside. Sara called Lisa a bad name! Was what Sara did good or bad?

5. One day Peter was in class and did not speak politely to his teacher, and he called her by her first name. So instead of saying Mrs. Wilson, he called his teacher by her first name Helen. Was what Peter did good or bad?

6. Anna and Kayla were making drawings. Anna took Kayla's drawing and ripped it up. Was what Anna did good or bad?

7. Carol was playing at school. All of the other children put their toys away in the right places, but Carol put her toys on the floor. Was what Carol did good or bad?

After each item children answered these 5 questions:

1. Was what _____ did good or bad?
2. How good/bad? (a little good/bad, very good/bad, or very very good/bad)?
3. What if _____ did (transgression) somewhere far away, say in another country then is it good/bad?
4. What if there wasn't a rule in _____'s school about (transgression). Would it be okay for _____ to do (transgression) then?
5. What if everyone else at _____'s school did (transgression). Would it be okay for _____to do (transgression) then?

Study 2 Items

Standard Moral/Conventional Items:

1. Nahid and Kimia were standing outside. Nahid purposely shoved Kimia! Was what Nahid did good or bad?

2. Nima was sitting at the lunch table at school. He then took off his shoes and put them on the table. Was what Nima did good or bad?

3. Soraya and Mahsa were playing outside. Soraya called Mahsa a bad name! Was what Soraya did good or bad?

4. All of the children were playing outside. Soroush decided to wear his bathing suit to school instead of his school clothes. Was what Soroush did good or bad?

5. Arezoo and Layla were making drawings. Arezoo took Layla's drawing and ripped it up. Was what Arezoo did good or bad?

6. One day Hessam was in class and called his teacher by her first name. So instead of calling her Mrs. Rahimi he called her Firoozeh. Was what Hessam did good or bad?

7. Tannaz was playing at school. All of the other children put their toys away in the right places, but Tannaz put her toys on the floor. Was what Tannaz did good or bad?

Culture-Specific Conventional Items:

1. One day Sara wanted to draw. She went into the living room, sat on the floor and turned her back to her grandmother who was sitting on the couch. Was what Sara did good or bad?

2. Soheil was invited to have dinner at his friend Pedram's house. When Soheil arrived, he walked into Pedram's living room without taking his shoes off at the front door. Was what Soheil did good or bad?

3. Fariba was sitting at home with her guests. She suddenly got thirsty and poured herself a glass of water and drank it in front of her guests. She did not offer her guests any water. Was what Fariba did good or bad?

4. Kiyarash was invited to his friend Shaheen's birthday party. Kiyarash decided to wear his house slippers to the party instead of dressy/formal shoes. Was what Kiyarash did good or bad?

Religious Items (with no legal punishment for violation):

1. One day during Ramadan Newsha, who was not fasting, got hungry and ate ice cream in front of others in her home who were fasting. Was what Newsha did good or bad?

2. Siyamak was tired at night and didn't have the energy to wash up before his prayer. He said his prayers without washing up. Was what Siyamak did good or bad?

3. During Muharram, Mahsa wore bright, red clothes. Was what Mahsa did good or bad?

4. In the middle of prayer, Daniel told his friend a joke. Was what Daniel did good or bad?

After each item children answered these 10 questions:

1. Was what _____ did good or bad?
2. How good/bad? (a little good/bad, very good/bad, or very very good/bad)
3. What if _____ did (transgression) somewhere far away, say in another country then is it good/bad?
4. What if there wasn't a rule in _____'s school about (transgression). Would it be okay for _____ to do (transgression) then?
5. What if everyone else at _____'s school did (transgression). Would it be okay for _____to do (transgression) then?
6. What if _____ wasn't from Iran and was from another country such as X. Would it be okay for _____to do (transgression) then? Is it okay for a foreigner to do (transgression)?
7. What if _____ isn't Muslim. Then would it be okay? Is it okay for a non-Muslim to do (transgression)?
8. Is (transgression) "haram" (religiously forbidden)?
9. Do you think there should be a law about (transgression)? Should it be illegal?
10. Do you think what _____ did was "zesht" (socially disapproved)?

6. Data Coding & Analysis

Statistical analyses: General approach

Analyses of acceptability ratings

For Studies 1 and 2, within each culture we tested whether children's ratings differed by transgression type and age by running the following linear mixed effects model: Acceptability rating ~ Transgression type (moral; conventional; conventional-Iran; religious)*Age + (Transgression type|subject) + (age|item). This model includes the maximal random effect structure consistent with our design, following the recommendations of Barr, Levy, Sheepers, and Tily, 2013. We began with all design-relevant fixed effects as random slopes and then iteratively removed coefficients until the model converged. We first removed the slope of the item effect, then of the subject.

Because there are only two levels of transgression category in Study 1 (moral and conventional), in these analyses we used beta coefficients and p-values generated by the model as an indication of the difference in ratings between transgression categories.

For Study 2, we investigated differences between transgression types by setting the moral transgression type as the intercept, and then using beta coefficients and p-values attached to conventional, conventional-Iran, and religious transgression types as an indication of the difference in ratings between these three transgression categories and moral transgressions.

Analyses of circumstances ratings

For Studies 1 and 2, within each culture and for each follow-up question type about circumstances (e.g., "Would it be okay if everyone else were doing it?"), we tested if children's decisions about whether a transgression is permissible under certain circumstances varied by transgression type by running the following generalized linear mixed effects model: Permissibility

response (yes/no) ~ Transgression type (moral; conventional; conventional-Iran; religious) *Age + (transgression type|subject) +(age|item). We used a maximal random effects structure, and iteratively removed coefficients until the model converged. We first removed the item slopes, then subject slopes, and finally the item intercept altogether if the model still failed to converge. Because we undertook this model building process separately for each testing site, random effects structures are slightly different across cultures.

For Study 1, we used beta coefficients and p-values generated by the model as an indication of the difference in ratings between transgression categories. For Study 2, we investigated differences between transgression types by setting the moral transgression type as the intercept, and then used beta coefficients and p-values attached to conventional, conventional-Iran, and religious transgression types as an indication of the difference in responses between these three transgression categories and moral transgressions. In Study 2, we also tested whether children made different allowances for foreigners and non-Muslims by exploring their yes/no responses to these questions with the above model.

Statistical analyses: Study 2 items

In Study 2, we measured Iranian children's decisions about whether certain categories of transgression should be illegal; are considered rude; or should be religiously forbidden and if this differs by transgression type. To do this, we used the following generalized linear mixed effects model: Acceptability rating ~ Transgression type (moral; conventional; conventional-Iran; religious)*Age + (transgression type|subject) + (age|item). Once again, we began with a maximal random effects structure, and iteratively removed coefficients until the model converged, starting with item slopes, then subject slopes, and then the item intercept.

For Study 2, we also investigated differences between transgression types by setting the moral transgression type as the intercept, and then used beta coefficients and p-values attached to conventional, conventional-Iran, and religious transgression types as an indication of the difference in responses between these three transgression categories and moral transgressions.

Study 1 Models

Study 1: Analysis of acceptability ratings by age and testing site

Table 3.3. Parameter Estimates for Linear Mixed Effects Models Predicting Acceptability Ratings in Study 1. *$p < .05$; **$p < .01$; ***$p < .001$.

Site Model	**Intercept**	**Transgression (Conventional)**	**Age**	**Transgression:Age**
Canada Acceptability rating ~ Transgression + Age + (1|Participant) + (1|Item)	1.81***	0.50*	0.14*	—
India Acceptability rating ~ Transgression*Age + (1|Participant) + (1|Item)	1.17***	0.28	0.02	0.30**
Korea Acceptability rating ~ Transgression + Age + (1|Participant) + (1|Item)	1.28***	0.34	0.20	—
Iran Acceptability rating ~ Transgression + Age + (Transgression|Participant) + (1|Item)	1.31***	0.09	0.14**	—

Study 1: Analysis of acceptability ratings across cultural groups

Table 3.4. Parameter Estimates for Linear Mixed Effects Models Predicting Cross-cultural Acceptability Ratings in Study 1. *$p < .05$; **$p < .01$; ***$p < .001$.

Model: Acceptability rating ~ Site*Transgression*Age + (1|Participant)	
	Coefficient
Intercept	1.83***
India	-0.66***
Korea	-0.53***
Iran	-0.54***
Transgression (Conventional)	0.46***
Age	0.21***
India:Transgression (Conventional)	-0.18
Korea:Transgression (Conventional)	-0.16
Iran:Transgression (Conventional)	-0.33***
India:Age	-0.19
Korea:Age	-0.07
Iran:Age	-0.11
Transgression:Age	-0.12
India:Transgression:Age	0.42***
Korea:Transgression:Age	0.24
Iran:Transgression:Age	0.22*

Study 1: Analyses of permissibility by circumstance

Would it be okay if it happened in a faraway country?

Table 3.5. Parameter Estimates for Generalized Linear Mixed Effects Models Predicting Acceptability Ratings in Study 1. *$p < .05$; **$p < .01$; ***$p < .001$

Site Model	**Intercept**	**Transgression (Conventional)**	**Age**
Canada Acceptability ~ Transgression + Age + (1|Participant) + (1|Item)	-3.85***	1.50*	-0.24
India Acceptability ~ Transgression*Age + (1|Participant) + (1|Item)	-2.70***	-0.26	-0.59
Korea Acceptability ~ Transgression + Age + (1|Participant) + (1|Item)	-4.67***	1.12	-0.23
Iran Acceptability ~ Transgression + Age + (1|Participant) + (1|Item)	-12.62***	2.74	0.60

Would it be okay if there was no rule against it?

Table 3.6. Parameter Estimates for Generalized Linear Mixed Effects Models Predicting Acceptability Ratings in Study 1. *$p < .05$; **$p < .01$; ***$p < .001$

Site Model	**Intercept**	**Transgression (Conventional)**	**Age**
Canada Acceptability ~ Transgression + Age + (1|Participant) + (1|Item)	-1.98**	1.53*	-1.01*
India Acceptability ~ Transgression*Age + (1|Participant) + (1|Item)	-2.37***	0.40	0.15
Korea Acceptability ~ Transgression + Age + (1|Participant) + (1|Item)	-5.54***	1.50	1.95*
Iran Acceptability ~ Transgression + Age + (1|Participant) + (1|Item)	-8.23***	0.64	0.30

Would it be okay if everyone else was doing it?

Table 3.7. Parameter Estimates for Generalized Linear Mixed Effects Models Predicting Acceptability Ratings in Study 1. *$p < .05$; **$p < .01$; ***$p < .001$

Site Model	**Intercept**	**Transgression (Conventional)**	**Age**
Canada Acceptability ~ Transgression + Age + (1|Participant) + (1|Item)	-4.29***	1.42*	0.80
India Acceptability ~ Transgression*Age + (1|Participant) + (1|Item)	-3.11**	1.12	-0.15
Korea Acceptability ~ Transgression + Age + (1|Participant) + (1|Item)	-3.67***	1.73**	1.77*
Iran Acceptability ~ Transgression + Age + (1|Participant) + (1|Item)	-8.61***	1.03	0.60

Additional Analyses from Study 1

Table 3.8. Mean Permissibility Ratings (1 = *okay*, 0 = *not okay*) by cultural group and circumstance

	In a faraway country		If there's no rule against it		If everyone else did it	
	Moral *M (SD)*	Conventional	Moral	Conventional	Moral	Conventional
Canada	0.10 (0.30)	0.22 (42)	0.34 (0.47)	0.48 (0.50)	0.10 (0.30)	0.19 (0.40)
India	0.09 (0.29)	0.08 (0.27)	0.22 (0.41)	0.25 (0.44)	0.25 (0.43)	0.34 (0.48)
Korea	0.01 (0.10)	0.03 (0.18)	0.07 (0.25)	0.15 (0.36)	0.14 (0.35)	0.33 (0.47)
Iran	0.04 (0.20)	0.09 (0.29)	0.14 (0.35)	0.15 (0.36)	0.04 (0.19)	0.06 (0.25)

Study 2 Models

Study 2: Analysis of acceptability ratings by transgression type and age

Table 3.9. Parameter Estimates for Linear Mixed Effects Models Predicting Acceptability Ratings in Study 2. *$p < .05$; **$p < .01$; ***$p < .001$

Model	Intercept	Conventional	Conventional - Iran	Religious	Age
Acceptability ~ Transgression + Age + (1|Participant)	1.46***	0.06	0.44***	0.06	0.13*

Study 2: Analyses of permissibility ratings of transgression types by age and circumstance

Table 3.10. Parameter Estimates for Generalized Linear Mixed Effects Models Predicting Acceptability Ratings in Study 2. *$p < .05$; **$p < .01$; ***$p < .001$

Circumstance Model	Intercept	Conven -tional	Conv-Iran	Religious	Age	Conv: Age	Conv-Iran: Age	Religious: Age
Faraway country Acceptability ~ Transgression + Age + (1|Participant)	-2.66***	0.69*	1.51***	1.70***	0.69**	—	—	—
No rule Acceptability ~ Transgression + Age + (1|Participant)	-2.04***	0.91***	1.45***	1.00***	0.73**	—	—	—
Everyone else Acceptability ~ Transgression + Age + (1|Participant)	-3.03***	1.26***	1.84***	0.13	0.92***	—	—	—
Foreigner Acceptability ~ Transgression + Age + (1|Participant)	-3.05***	0.54	2.01***	1.89***	0.67	-0.17	0.03	0.78*
Non-Muslim Acceptability ~ Transgression + Age + (1|Participant)	-3.25**	0.39	1.77**	2.38***	-0.08	0.85*	0.77*	1.38***

Study 2: Analyses of acceptability ratings of additional items

Is it (a sin/illegal/rude)

Table 3.11. Parameter Estimates for Generalized Linear Mixed Effects Models Predicting Whether a Transgression is a Sin/Illegal/Rude in Study 2. *$p < .05$; **$p < .01$; ***$p < .001$

Sin/Illegal/Rude *Intercept* Model	**Intercept**	**Moral**	**Conventional**	**Conventional - Iran**	**Religious**	**Age**
Sin *Religious transgressions as intercept* Acceptability ~ Transgression + Age + (1|Participant)	1.99***	-2.02***	-2.43***	-3.03***	—	-1.31***
Illegal *Moral transgressions as intercept* Acceptability ~ Transgression + Age + (1|Participant)	0.08	—	-0.03	-1.50***	0.78**	-1.12***
Rude *Conventional-Iran transgressions as intercept* Acceptability ~ Transgression + Age + (1|Participant)	2.98***	2.85***	0.98*	—	-0.07	-0.54

Study 2: Analysis of justifications

Children's justifications were translated by an Iranian native-speaker. A thematic analysis of responses indicated that children's responses could be categorized into seven justification themes or categories, most of which were similar to those found in prior research (Davidson et al., 1983; Hollos et al., 1986; Nucci, 1981; Smetana, 1985; Song et al., 1987). These themes are described in Table 3 below. Three raters coded all justifications into one of seven themes, and the theme that 2/3 of the raters selected as the best fit for each justification was used. Interrater reliability after selecting the theme that ⅔ of raters agreed on was 97.53%. For the small number of cases where there was no agreement between the three raters, a fourth rater determined the justification theme.

Table 3.12. Description of Themes Identified in Children's Rating Justifications

Justification Themes	
Theme Abbreviation	**Description**
Injury	Action causes physical harm to the victim
Disrespect	Action causes disrespect or offense to the victim, was disregarding of the victim's feelings and wants
Embarrassing	Action reflects poorly on the actor because it is embarrassing or shameful for them or causes embarrassment to those around them
Social order	Action violates social order and rules, is wrong because it goes against what others are doing, is wrong because it's inappropriate in the social context it was committed
Religious	Action is offensive to God or other religious figure, violates what God or religion preaches, violates religious practices and rules
Wrong	Action is wrong without further justification "wrong because it's wrong", causes victim inconvenience, is dirty/bad/gross
OK	Action is acceptable

7. Site Information

Canada

The research was conducted in Comox Valley, a regional district in British Columbia, at two elementary schools. The primary language spoken at both schools is English and only students who were fluent in English were tested. At the time of testing, the population of Comox was 14,208 and the median household income was $69, 254. The majority of residents in Comox speak only English (86.88%) and the second language taught in schools is French. The median age in Comox Valley is 48.3 and 45% of residents have completed a post-secondary education. Most residents identify as being of European or North American origin, and the three largest visible minority groups are Chinese, Black, and South Asian respectively.

India

The research was conducted in the city of Vadodara located in Gujarat Province, India at a K-12 charitable school for low-income children in the district. The majority religious group in Vadodara are Hindus, but the school makes an effort to enroll approximately an equal number of Hindu and Muslim students. The median household income in India in 2013 was $3.168 and 80% of the families of children at this school earned less than $2,000 per year (~$5.50/day). Students of the school are admitted on a first-come, first-serve basis for approximately $10 per month, which is paid for by school trustees in cases of need. Though children receive instruction in English, the most common languages spoken by children at home are Gujarati and Hindi.

Iran

The research was conducted in Tehran, the capital of Iran, and Karaj, a large suburb of Tehran located about 20 miles to the west. Karaj is a major industrial city with more affordable housing available for middle class migrants compared to Tehran. The median household income at the time of testing was equivalent to $28,647 for families in the Tehran region.

Child participants were tested in libraries, after school care programs, and community centers in neighborhoods classified as high, medium and low income. All participants were native speakers of Farsi and all materials and instructions were administered in this language. Though we did not collect income information from families, testing sites were located in neighborhoods with families from predominantly high, middle, or low-income backgrounds. We did not find any effect of neighborhood or testing site on children's comprehension of the testing materials or responses.

All child participants were enrolled in school, where they are exposed to religious teachings through the school curriculum. Children attending K-12 schools in Iran are exposed to religious teachings through the school curriculum. In 1980, after the Iranian revolution, Shi'ite Islamic values became incorporated into the teaching materials of primary grade students. Education in Iran is highly centralized with the Ministry of Education leading the educational planning, financing, curriculum and textbook development for students. The school curriculum includes Quran and Theology/Religion classes where principles of Shia Islam and religious customs are taught. Both private and public schools are required to teach religious practices and use the same uniform set of textbooks for each grade (Mehran, 1997; Mirhadi, 1997).

Though exposure to religious practices and customs in the home varies across households, children are taught Islamic values and religious practices such as praying, fasting, and observing religious holidays in school. Therefore, regardless of the level of exposure children have to religious teachings in the home environment, Iranian children were familiar with the religious items in our testing materials. We did not directly ask children or their families about their own religious practices or knowledge due to the sensitive nature of this question under a religious government, but a in a 2005 survey, 78% of parents indicated that religion is a very important

aspect of their lives and 71% of parents indicated that religious faith should be encouraged upon children at home (World Values Survey, 2005).

After the 1979 Iranian Revolution, Iran became an Islamic republic and Shia Islam was declared as the official state religion. According to Iran's census 99% of the population is Muslim and the remaining 1% are Christian, Zoroastrian, Baha'i and Jewish. However, it is important to note that this statistic may not accurately reflect the religious identification of Iranians due to political and social pressure to identify as Muslim. Other smaller scale surveys suggest a decline in religiosity among Iranians (The Group for Analyzing and Measuring Attitudes in Iran, 2020). Following the revolution, Iran's modern legal system was replaced by an Islamic legal system that is based on Shi'ite sharia law, which acts as a code of living that Muslims should adhere to including family obligations and religious obligations such as prayers and fasting, and financial dealings (Arjomand, 1989).

One of the primary instruments for the government to enforce Islamic codes of behavior was the establishment of a "morality police" known as the Basij Resistance Force (Golkar, 2011). The Basij have been given legal permission to take necessary actions whenever a crime is committed and police are not available. Primarily, the Basij are responsible for confronting and reporting individuals who commit "moral crimes", for instance extramarital sexual relationships, alcohol consumption, owning satellite dishes, not following Islamic dress codes, and gambling (Golkar, 2011). Depending on the severity of the crime, suspects of immoral or un-Islamic behavior are reported to the police, required to pay a fine, sign a commitment letter to not repeat the offense, are given a verbal warning or are detained (Golkar, 2011; Hoodfar & Ghoreishian, 2012). The Basij primarily patrol public areas such as parks and streets but are known to also disrupt private gatherings and parties if they believe Islamic codes are not being observed. It is

unknown to what extent young children in Iran are aware of the Basij and their role in enforcing Islamic norms, but for many Iranian residents their presence instills a heightened sense of intimidation and fear (Golkar, 2011; Khosravi, 2008).

Korea

The research was conducted in Myeondong Elementary school in Seoul Korea located in a low to middle class neighborhood where the average household income is about $3,200 per month. The primary language of instruction at this school is Korean and teachers were involved with data collection. Religious information was not collected from child participants and their families.

Korea is ethnically and culturally homogenous with over 99% of the population identifying as ethnically Korean (Shin, 2006; World Population Review, 2022). Korean society has been characterized as focusing more on traditions, conformity, politeness, and harmonious interdependence than Western cultures (e.g., Ambady, Koo, Lee, & Rosenthal, 1996; Hotltgraves & Yang, 1992; Park & Johnson, 1984). Uniformity in language, race, and cultural background are promoted in Korea's mass media and educational system along with other aspects of ethnic nationalism (Shin, 2006). The Korean Ministry of Education controls school curricula and some aspects of media including television, thereby promoting uniformity in cultural consumption (Song et al., 1987). One of the primary objectives of the Ministry is to foster a sense of morality, responsibility, and community awareness among Korean citizens (Lee, 1996).

From an early age, Korean children are taught to obey their elders, conform to rules, and adhere to cultural conventions (Kim, 1998; Rohner & Pettengill, 1985). For example, in schools they are taught to follow rules about putting items in the right place, dressing appropriately, and using polite language with adults (Song et al., 1987). In line with this, prior work on children's moral reasoning shows that Korean children tend to focus more on social status, courtesy,

obligation, and social roles than American children when evaluating the severity of conventional violations (Song et al., 1987).

Supplemental Materials References

Ambady, N., Koo, J., Lee, F., & Rosenthal, R. (1996). More than words: Linguistic and nonlinguistic politeness in two cultures. *Journal of Personality and Social Psychology, 70*, 996–1011.

Arjomand, S. A. (1989). History, structure, and revolution in the Shi'ite tradition in contemporary Iran. *International Political Science Review*, *10*(2), 111-119.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278.

Davidson, P., Turiel, E., & Black, A. (1983). The effect of stimulus familiarity on the use of criteria and justifications in children's social reasoning. *British Journal of Developmental Psychology*, *1*(1), 49-65.

Golkar, S. (2011). Politics of piety: the Basij and moral control of Iranian society. *The Journal of the Middle East and Africa*, *2*(2), 207-219.

Hollos, M., Leis, P. E., & Turiel, E. (1986). Social reasoning in Ijo children and adolescents in Nigerian communities. *Journal of Cross-Cultural Psychology*, *17*(3), 352-374.

Holtgraves, T., & Yang, J. N. (1992). Interpersonal underpinnings of request strategies: General principles and differences due to culture and gender. Journal of Personality and Social Psychology, *62*(2), 246.

Hoodfar, H., & Ghoreishian, A. (2012). Morality policing and the public sphere: women reclaiming their bodies and their rights. *Sexuality in Muslim contexts: Restrictions and resistance*, 234-268.

Kahn Jr, P. H. (1992). Children's obligatory and discretionary moral judgments. *Child Development*, *63*(2), 416-430.

Khosravi, S. (2008). *Young and defiant in Tehran*. University of Pennsylvania Press.

Killen, M., & Smetana, J. G. (1999). Social interactions in preschool classrooms and the development of young children's conceptions of the personal. *Child Development*, *70*(2), 486-501.

Kim, J. M. (1998). Korean children's concepts of adult and peer authority and moral reasoning. *Developmental Psychology*, *34*(5), 947.

Lahat, A., Helwig, C. C., & Zelazo, P. D. (2012). Age-related changes in cognitive processing of moral and social conventional violations. *Cognitive Development*, *27*(2), 181-194.

Lee, K. S. (1996). Culture and the Korean kindergarten curriculum. *Early Child Development and Care, 123*(1), 167-182.

Mehran, G. (1997). Education xxiv. Education in post-revolutionary Persia, 1979–95. In Encyclopedia Iranica. (Vol. VIII, Fasc. 3, pp. 230–233).

Mirhadi, T. (1997). Education viii. Nursery schools and kindergartens. In Encyclopedia Iranica. (Vol. VIII, Fasc. 2, pp. 196–199).

Nucci, L. (1981). Conceptions of personal issues: A domain distinct from moral or societal concepts. *Child Development, 52*(1), 114-121.

Nucci, L., Camino, C., & Sapiro, C. M. (1996). Social class effects on northeastern Brazilian children's conceptions of areas of personal choice and social regulation. *Child Development, 67*(3), 1223-1242.

Nucci, L. P., & Turiel, E. (1978). Social interactions and the development of social concepts in preschool children. *Child Development, 49*(2), 400-407.

Nucci, L., & Weber, E. K. (1995). Social interactions in the home and the development of young children's conceptions of the personal. *Child Development, 66*(5), 1438-1452.

Park, J. Y., & Johnson, R. C. (1984). Moral development in rural and urban Korea. *Journal of Cross-Cultural Psychology, 15*(1), 35-46.

Rohner, R. P., & Pettengill, S. M. (1985). Perceived parental acceptance-rejection and parental control among Korean adolescents. *Child Development*, 56(2), 524-528.

Sanderson, J. A., & Siegal, M. (1988). Conceptions of moral and social rules in rejected and nonrejected preschoolers. *Journal of Clinical Child Psychology, 17*(1), 66-72.

Shin, G. W. (2006). *Ethnic nationalism in Korea: Genealogy, politics, and legacy*. Stanford University Press.

Siegal, M., & Storey, R. M. (1985). Day care and children's conceptions of moral and social rules. *Child Development, 56*(4), 1001-1008.

Smetana, J. G. (1981). Preschool children's conceptions of moral and social rules. *Child Development, 52*(4), 1333-1336.

Smetana, J. G. (1985). Preschool children's conceptions of transgressions: Effects of varying moral and conventional domain-related attributes. *Developmental Psychology, 21*(1), 18.

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

Smetana, J. G., & Braeges, J. L. (1990). The development of toddlers' moral and conventional judgments. *Merrill-Palmer Quarterly*, *36*(3), 329-346.

Song, M. J., Smetana, J. G., & Kim, S. Y. (1987). Korean children's conceptions of moral and conventional transgressions. *Developmental Psychology*, *23*(4), 577.

The Group for Analyzing and Measuring Attitudes in Iran (GAMAAN). (2020). Iranians' attitudes toward religion: A 2020 survey report. Retrieved from https://gamaan.org/2020/08/25/iranians-attitudes-toward-religion-a-2020-survey-report/

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

Turiel, E. (2002). *The culture of morality: Social development, context, and conflict*. Cambridge University Press.

Turiel, E. (2006). Thought, emotions, and social interactional processes in moral development. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 7–35). Lawrence Erlbaum Associates Publishers.

Turiel, E. (2008). Thought about actions in social domains: Morality, social conventions, and social interactions. *Cognitive Development*, *23*(1), 136-154.

Weston, D. R., & Turiel, E. (1980). Act–rule relations: Children's concepts of social rules. *Developmental Psychology*, *16*(5), 417.

World Population Review. (2022). South Korea population 2022. Retrieved from https://worldpopulationreview.com/countries/south-korea-population

World Values Survey Association. (2005). World values survey 2005. Retrieved from http://www.worldvaluessurvey.org/wvs.jsp

Yau, J., & Smetana, J. G. (2003). Conceptions of moral, social-conventional, and personal events among Chinese preschoolers in Hong Kong. *Child Development*, *74*(3), 647-658.

Zimba, R. F. (1994). The understanding of morality, convention, and personal preference in an African setting: Findings from Zambia. *Journal of Cross-Cultural Psychology*, *25*(3), 369-393.

GENERAL DISCUSSION

Stereotypes, racism, injustice, and wars can all result from intergroup prejudice, the basis of which emerges in early childhood. Many scholars, including psychologists, historians, policy makers, and social activists, want to uncover the origins of prejudices with the shared goal to improve intergroup relations and minimize conflict. This is a complex endeavor since humans differ along many dimensions and it remains unknown whether we all share general cognitive mechanisms that predispose us to prefer our own group over others or if these beliefs are shaped by cultural input. The studies presented in this dissertation help advance our understanding of how and why prejudices develop early in life by examining the extent to which cognitive and cultural factors shape children's attitudes and behaviors toward different groups of people.

Together, my findings revealed that group biases are not entirely contingent on cultural input, since both children and adults are motivated to favor ingroup members over outgroup members even when the groups are artificial and carry no cultural significance (Chapter 1). However, when the groups are meaningful to children and have critical implications for their lives, cultural factors such as status and sociopolitical relations drive children's group preferences (Chapter 2). Finally, I showed that implementing a two-stage model for conducting cross-cultural research on children's moral reasoning allows us to more accurately capture aspects of children's social and cognitive development that are universal in addition to those that are products of their unique cultural environments (Chapter 3).

In Chapter 1, I demonstrated that children and adults are more motivated by the desire to favor their ingroup than the desire to enhance their reputation when sharing resources. Further, participants displayed a strong ingroup bias even though the groups were artificial and arbitrarily determined. I did find evidence that reputation concerns prompted younger and older children to

donate more of their resources in public versus private, but this effect was the same whether the recipient was an ingroup or outgroup member. Interestingly, though adults showed a similar bias to children in their giving behavior, their reasoning for doing so was different: Children judged ingroup giving to be significantly "nicer" than outgroup giving, whereas adults evaluated outgroup giving as more generous, noting that it is less common, requires greater sacrifice, and is driven by less self-serving, ulterior motives. These are the first findings to show that children differ from adults in how they reason about ingroup and outgroup giving.

One reason for the developmental shift in reasoning that we observed could be that children, unlike adults, do not discount self-serving motives when judging other people's giving (Heyman, Barner, Heumann, & Schenk, 2013), and accordingly also do not discount ingroup giving, which adults view as less generous. Another possibility is that children do not perceive ingroup giving as self-serving, but rather as an act of group loyalty. Consistent with this possibility, prior studies show that children expect individuals to prioritize and behave more loyally toward ingroup members than outgroup members (DeJesus, Rhodes, & Kinzler, 2014; Misch, Over, & Carpenter, 2014). Together, the findings of Chapter 1 reveal that children are inclined to favor ingroup members on the basis of category membership alone, and despite evidence of a difference in how adults reason about intergroup giving, both children and adults display a robust ingroup bias in their sharing behavior.

However, minimal groups do not give us the full picture of how group relations work in the real world. By treating all groups as methodologically the same, the minimal group paradigm does not examine how factors such as historical conflict, cultural differences, power dynamics, and intersectionality shape intergroup dynamics. First, even groups that share a common country, language, and ethnic background sometimes engage in violent conflict to gain power and control

over resources (e.g., Hutu and Tutsi in Rwanda, Serbs and Bosniaks in former Yugoslavia). Second, the duration and intensity of conflict between groups can also shape the extent to which ingroup members adopt negative and dehumanizing stereotypes about the outgroup, thereby perpetuating more hostility and mistrust between groups (Schwartz & Struch, 1989). For example, a study of Israeli and Palestinian children shows that the longer that children were exposed to ethno-political violence over a three-year period, the more likely they were to endorse negative and dehumanizing stereotypes of the "enemy" outgroup (Niwa et al., 2014). Third, in the real world, people belong to multiple social groups simultaneously, and how these groups intersect can create distinct experiences for each group member (Lei & Rhodes, 2021). For instance, gender and race intersect such that the experience of a black female in a low status group might be fundamentally different from that of a white male in the same group (Purdie-Vaughns & Eibach, 2008). Together, these examples show that real-world groups have complexities that cannot be captured with minimal groups.

Another important determinant of intergroup bias that was not addressed in the minimal group paradigm used in Chapter 1 is how ingroup cooperation and loyalty are shaped by cultural forces. Ingroup cohesion is higher in collectivist societies than individualistic societies because conformity and obligations to the group are emphasized (Triandis, 1995). Also, group loyalty has been shown to increase in contexts where outgroup members are perceived as a threat to the ingroup's control over resources (Citrin, Green, Muste, & Wong, 1997; Esses, Jackson, & Armstrong, 1998), political power, and nation's distinctive identity (De Figueriredo & Elkins, 2003; Huddy, 2001). In contrast, ingroup loyalty is decreased when groups work together to achieve goals (Sherif, 1966), have flexible group boundaries (e.g., pluralistic societies; de la Garza, Falcon, & Garcia, 1996), or if the outgroup has higher status (Bigler, Spears, Brown, & Markell,

2001; Nesdale & Flesser, 2001). These findings illustrate that the specific nature of ingroups and outgroups matter in shaping their relations.

In Chapter 2, I examined how children form attitudes toward real-world groups and presented evidence that children distinguish between outgroups and view some as favorably as their own group. Surprisingly, Iranian children in our study expressed the most desire to befriend, trust, and affiliate with American children—the outgroup which they indicated as having the highest social status. In contrast, children expressed the least desire to affiliate with Arab children, the outgroup that they perceived as having the lowest status. In contrast to the findings of Chapter 1 and other minimal group studies, Iranian children did not favor the outgroup that was the most similar to their ingroup (e.g., Iranian children from a different school)--challenging a prevalent claim in psychology that prejudices emerge from a general affinity for similar others and dislike of dissimilar others (Aboud, 1988; Allport, 1979; Baron & Banaji, 2006; Brewer, 1979; Mahajan & Wynn, 2012). The findings from Chapters 1 and 2 raise the possibility that, when only limited information about the ingroup and outgroup is available (e.g., minimal groups), children are inclined to base their group preferences on similarity, but if more cultural information about the groups is given, then other dimensions such as sociopolitical relations and group status can be stronger predictors of children's intergroup biases.

Prior research on the role of status in children's intergroup beliefs reveal patterns that are consistent with those observed in Chapter 2. Generally, studies find that children from lower status racial and ethnic groups display less ingroup favoritism than children from racial groups that traditionally have greater status, wealth, and power (Bigler et al., 2001; Dunham, Srinivasan, Dotsch, & Barner, 2013; Horwitz, Shutts, & Olson, 2014; Newheiser, Dunham, Merrill, Hoosain, & Olson, 2014). For example, white American children display robust ingroup favoritism over

other racial and ethnic groups, whereas black American and Hispanic children show less ingroup bias and often express a pro-white bias due to an awareness of their ingroup's low status relative to whites in the U.S. (Baron & Banaji, 2006; Hailey & Olson, 2013; Spencer & Markstrom-Adams, 1990). Similar evidence of children's pro-white biases has been observed in South Africa, where status differences between racial groups are more extreme (Newheiser et al., 2014), and among minority Pacific Islander and Aboriginal children in Australia (Griffiths & Nesdale, 2006). These findings and those presented in Chapter 2 show that outgroup bias is not inevitable and that promoting the view that groups have equal status may help.

Our findings regarding the role of status in intergroup relations builds on prior work demonstrating that the effects of intergroup contact depend on the perceived status of the groups. Increased contact between members of different social groups (e.g., intergroup contact theory; Allport, 1954; Pettigrew & Tropp, 2006) has been found to reduce prejudiced attitudes and promote positive intergroup relations under the condition that the groups are recognized by authority figures, institutions, and laws as having equal status (Brewer & Miller, 1984; Brown & Hewstone, 2005; Savelkoul, Shceepers, Tolsma, & Hagendoorn, 2011). On the other hand, when groups are perceived as unequal, even peaceful intergroup contact between members can have no effect on group attitudes (Bratt, 2008) or, in some cases, can heighten group tensions and negative views of lower status groups by the majority group (Brewer, 1996; Hopkins, 2010; Tropp, 2003; Tropp & Pettigrew, 2005). For example, historical records show that when lower status minority or immigrant groups shift the demographics of local communities, intergroup contact is often met with hostility by members of the majority group (e.g., Horton 1995; Kruse, 2005; Lassiter, 2006; Sniderman, Hagendoorn, & Prior, 2004).

As illustrated in Chapter 2, when children view outgroups as having high status, they are more likely to welcome friendship and trust with outgroup members. Such interactions early in life could have the potential to promote better relations between groups on a global scale. To facilitate peaceful intergroup relations, it is imperative that children not only engage in positive intergroup contact, but also be exposed to cultural messages that emphasize equality among ingroup and outgroup members (see Heyman & Yazdi, 2019, for a review). Overall, Chapter 2 offers some of the only data regarding children's intergroup attitudes in Iran and highlights how cultural messages and affairs shape children's attitudes toward different groups. These results point to the need for more developmental research with non-Western children in real-world settings.

Cross-cultural research and the inclusion of more non-WEIRD groups to psychology studies are essential to understanding how culture shapes cognitive development. Equally important is the question of how to compare groups while remaining sensitive to the social, political, and cultural realities of participants across cultures. In Chapter 3, I raised the problem of validity that arises when researchers use measures that were primarily developed for WEIRD groups to make cross-cultural comparisons of children's social and moral reasoning. I demonstrated how the use of a two-stage model to test for the moral/conventional distinction in children from Canada, India, Iran and Korea tackles the problem of generalizability and cultural validity within a single study. Using a standardized measure that has been widely used in prior developmental work to test for the moral/conventional distinction, I found that Canadian children (the only Western group) showed clear evidence of the moral/conventional distinction, but not children from the non-Western groups. However, when items were tailored to the local practices and beliefs of Iranian children, the non-Western group who showed the weakest evidence of the

moral/conventional distinction on the standardized measure, a robust moral/conventional distinction was detected.

If only standardized measures had been used to test for the moral/conventional distinction in children from WEIRD and non-WEIRD groups, as is commonly practiced in cross-cultural studies, our findings would have led us to believe that the moral/conventional distinction is strongest in children from Western groups and weaker, or nonexistent, in children from some non-Western groups. However, the additional findings obtained from the culturally tailored items in Stage 2 suggest that the group differences observed in Stage 1 can be attributed to a difference in whether conventional items overlap with other concerns for a particular group. For example, Korean children evaluated speaking impolitely to one's teacher as on par with moral violations. Also, the act of wearing a bathing suit to school has been defined by philosophers and developmental psychologists as being a "conventional" violation (Nucci, Turiel, & Encarnacion-Gawrych, 1983; Smetana & Braeges, 1990; Smetana, Rote, Jambon, Tasopoulos-Chan, Villalobos, & Comer, 2012), but in Iran this same act violates strict religious dress codes that are enforced by Iran's 'morality police', a religious police force established by Iran's Islamic government (Golkar, 2011). These cases illustrate the need for research items to reflect the local conventions, beliefs and laws of the participants being tested.

The findings from Chapter 3 highlight the perils of using WEIRD measures for cross-cultural research. One way to ensure that research measures reflect the current cultural climate of participants is to include local researchers, scholars, and residents in the development and implementation of these measures. To establish the culturally tailored items in Stage 2, I consulted with local residents of varying socioeconomic status and educational backgrounds in Iran and ensured that violations of cultural conventions did not overlap with other concerns, such as legal

punishment or violations of religious codes. This collaborative approach to cross-cultural research can help ameliorate the problem of WEIRD measures in developmental psychology and the study of human behavior more broadly.

In Chapter 3, I showed that by combining standardized measures with items tailored to the cultural climate of one non-WEIRD society, Iran, I was able to both replicate findings with a different sample of children and capture the moral/conventional distinction in Iranian children. These results indicate that standardized measures are a valuable tool for maintaining methodological consistency and replicability across groups, but the sole use of a single, untailored method that has been primarily validated with Western groups can lead to inaccurate and misleading conclusions. I argue that to improve both the reliability of findings and cultural validity of measures in cross-cultural research, researchers should follow a standardized procedure for collecting and analyzing data across groups and operationalize the concepts being tested (e.g., defining moral and conventional domains), but the specific items and questions used in the measure should be tailored to the cultural group of focus.

Uncovering the origins of prejudice is one of the biggest and most critical challenges in human research. Understanding why it forms, what factors shape its development, and when it leads to conflict rather than harmony are important questions to investigate. Here, I have demonstrated that both shared cognitive processes and specific cultural factors can shape when, why, and towards whom we exhibit negative biases. Though many questions on the nature of childhood prejudice remain unanswered and more cross-cultural work is needed on this topic, the work presented in this dissertation brings us one critical step closer to understanding the cognitive and cultural mechanisms that shape our social and moral beliefs.

References

Aboud, F. E. (1988). *Children and prejudice*. Blackwell.

Allport, G. W. (1954). *The nature of prejudice.* Addison-Wesley.

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science*,*17*(1), 53-58.

Bigler, R. S., Spears Brown, C., & Markell, M. (2001). When groups are not created equal: Effects of group status on the formation of intergroup attitudes in children. *Child Development*, *72*(4), 1151-1162.

Bratt C. (2008). The Jigsaw classroom under test: No effect on intergroup relations evident. *Journal of Community and Applied Social Psychology*, *18*(5), 403-419.

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, *86*(2), 307.

Brewer, M. B. (1996). When contact is not enough: Social identity and intergroup cooperation. *International journal of Intercultural relations*, *20*(3-4), 291-303.

Brewer M. B., Miller N. (1984). Beyond the contact hypothesis: Theoretical perspectives on desegregation. In Miller N., Brewer M. B. (Eds.), *Groups in contact: The psychology of desegregation* (pp. 281-302). Academic Press

Brown R., Hewstone M. (2005). An integrative theory of intergroup contact. In Zanna M. P. (Ed.), *Advances in experimental social psychology* (Vol. 37, pp. 255-343). Elsevier Academic Press.

Citrin, J., Green, D. P., Muste, C., & Wong, C. (1997). Public opinion toward immigration reform: The role of economic motivations. *The Journal of Politics*, *59*(3), 858-881.

De Figueiredo Jr, R. J., & Elkins, Z. (2003). Are patriots bigots? An inquiry into the vices of in-group pride. *American Journal of Political Science*, *47*(1), 171-188.

De la Garza, R. O., Falcon, A., & Garcia, F. C. (1996). Will the real Americans please stand up: Anglo and Mexican-American support of core American political values. *American Journal of Political Science*, 335-351.

DeJesus, J. M., Rhodes, M., & Kinzler, K. D. (2014). Evaluations versus expectations: Children's divergent beliefs about resource distribution. *Cognitive Science, 38*, 178–193.

Dunham, Y., Srinivasan, M., Dotsch, R., & Barner, D. (2014). Religion insulates ingroup evaluations: The development of intergroup attitudes in India. *Developmental Science, 17*(2), 311-319.

Esses, V. M., Jackson, L. M., & Armstrong, T. L. (1998). Intergroup competition and attitudes toward immigrants and immigration: An instrumental model of group conflict. *Journal of Social Issues, 54*(4), 699-724.

Golkar, S. (2011). Politics of piety: the Basij and moral control of Iranian society. *The Journal of the Middle East and Africa, 2*(2), 207-219.

Griffiths, J. A., & Nesdale, D. (2006). In-group and out-group attitudes of ethnic majority and minority children. *International Journal of Intercultural Relations, 30*(6), 735-749.

Hailey, S. E., & Olson, K. R. (2013). A social psychologist's guide to the development of racial attitudes. *Social and Personality Psychology Compass, 7*(7), 457-469.

Heyman, G. D., & Yazdi, H. (2019). The role of individuation in the development of intergroup relations. *Infant and Child Development, 28*(2), e2121.

Heyman, G. D., Barner, D., Heumann, J., & Schenck, L. (2013). Children's sensitivity to ulterior motives when evaluating prosocial behavior. *Cognitive Science, 38,* 683–700.

Hopkins, D. J. (2010). Politicized places: Explaining where and when immigrants provoke local opposition. *American Political Science Review, 104*(1), 40-60.

Horton, John. 1995. *The politics of diversity: Immigration, resistance, and change in Monterey Park, California*. Temple University Press.

Horwitz, S. R., Shutts, K., & Olson, K. R. (2014). Social class differences produce social group preferences. *Developmental Science, 17*(6), 991-1002.

Huddy, L. (2001). From social to political identity: A critical examination of social identity theory. *Political Psychology, 22*(1), 127-156.

Kruse,K. M. (2005) *White flight: Atlanta and the making of modern conservatism*. Princeton University Press.

Lassiter, Matthew D. 2006. *The silent majority: suburban politics in the Sunbelt South*. Princeton University Press.

Lei, R. F., & Rhodes, M. (2021). Why developmental research on social categorization needs intersectionality. *Child Development Perspectives, 15*(3), 143-147.

Mahajan, N., & Wynn, K. (2012). Origins of "us" versus "them": Prelinguistic infants prefer similar others. *Cognition, 124*(2), 227-233.

Misch, A., Over, H., & Carpenter, M. (2014). Stick with your group: Young children's attitudes about group loyalty. *Journal of Experimental Child Psychology, 126,* 19-36.

Nesdale, D., & Flesser, D. (2001). Social identity and the development of children's group attitudes. *Child Development, 72*(2), 506-517.

Newheiser, A. K., Dunham, Y., Merrill, A., Hoosain, L., & Olson, K. R. (2014). Preference for high status predicts implicit outgroup bias among children from low-status groups. *Developmental Psychology*, *50*(4), 1081.

Niwa, E. Y., Boxer, P., Dubow, E. F., Huesmann, L. R., Landau, S., Shikaki, K., & Dvir Gvirsman, S. (2016). Negative Stereotypes of Ethnic Outgroups: A Longitudinal Examination Among Palestinian, Israeli Jewish, and Israeli Arab Youth. *Journal of Research on Adolescence*, *26*(1), 166-179.

Nucci, L. P., Turiel, E., & Encarnacion-Gawrych, G. (1983). Children's social interactions and social concepts: Analyses of morality and convention in the Virgin Islands. *Journal of Cross-Cultural Psychology*, *14*(4), 469-487.

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751.

Purdie-Vaughns, V., & Eibach, R. P. (2008). Intersectional invisibility: The distinctive advantages and disadvantages of multiple subordinate-group identities. *Sex Roles*, *59*(5), 377-391.

Savelkoul, M., Scheepers, P., Tolsma, J., & Hagendoorn, L. (2011). Anti-Muslim attitudes in the Netherlands: Tests of contradictory hypotheses derived from ethnic competition theory and intergroup contact theory. *European Sociological Review*, *27*(6), 741-758.

Schwartz, S. H., & Struch, N. (1989). Values, stereotypes, and intergroup antagonism. In *Stereotyping and prejudice* (pp. 151-167). Springer, New York, NY.

Sherif, M. (1966). *In common predicament: Social psychology of intergroup conflict and cooperation*. Houghton Mifflin.

Smetana, J. G., & Braeges, J. L. (1990). The development of toddlers' moral and conventional judgments. *Merrill-Palmer Quarterly (1982)*, 329-346.

Smetana, J. G., Rote, W. M., Jambon, M., Tasopoulos-Chan, M., Villalobos, M., & Comer, J. (2012). Developmental changes and individual differences in young children's moral judgments. *Child Development*, *83*(2), 683-696.

Sniderman, P. M., Hagendoorn, L., & Prior, M. (2004). Predisposing factors and situational triggers: Exclusionary reactions to immigrant minorities. *American Political Science Review*, *98*(1), 35-49.

Spencer, M. B., & Markstrom-Adams, C. (1990). Identity processes among racial and ethnic minority children in America. *Child Development*, *61*(2), 290-310.

Struch, N., & Schwartz, S. H. (1989). Intergroup aggression: Its predictors and distinctness from in-group bias. *Journal of Personality and Social Psychology*, *56*(3), 364.

Triandis, H. C. (1995). *Individualism and collectivism*. Westview Press.

Tropp L. R. (2003). The psychological impact of prejudice: Implications for intergroup contact. *Group Processes & Intergroup Relations*, 6, 131-149.

Tropp, L. R., Pettigrew T. F. (2005). Relationships between intergroup contact and prejudice among minority and majority status groups. *Psychological Science*, 16, 951-956.