THE DEVELOPMENT OF MORAL FOUNDATIONS IN 2- AND 4-YEAR-OLDS

By

Sarah Elizabeth Peverill

Submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

at

Dalhousie University

Halifax, Nova Scotia

August 2020

Table of Contents

List of Tables

List of Figures

Abstract

This dissertation examined moral development in children using Moral Foundations Theory (MFT) as a guiding framework. MFT states that there are several moral 'foundations', or areas in which we may hold moral values and beliefs. MFT also asserts that moral foundations have both innate and environmental influences. The moral foundations examined in this dissertation were care, fairness, loyalty, authority, and sanctity. Therefore, relative to existing literature, this dissertation examined moral development from a broader perspective and at a younger age (for specific foundations). One hundred and seventy-eight children (102 2-year-olds, 76 4-year-olds) were included in this dissertation, as well as their parents. All children completed the Moral Foundations Puppet Task (MFPT), where they watched puppets doing good or bad acts relevant to each foundation. They subsequently indicated which puppet they preferred, and which puppet was bad. Only 4-year-olds completed the Moral Foundations Questionnaire for Kids (MFQK), a verbal measure of their sensitivity to each foundation and social norm violations. Parents completed the Moral Foundations Questionnaire (MFQ; Graham et al., 2011), an adult measure of moral foundations. Analyses conducted in Chapter 3 revealed that 2-year-olds' puppet choices did not differ significantly from chance on the MFPT. Four-year-olds were sensitive to the care, fairness, authority, and sanctity foundations, while their puppet choices for the loyalty foundation did not differ significantly from chance. Chapter 4 examined 4-year-olds' responses on the MFQK. The pattern of responses was consistent with the MFPT; 4-year-olds were more sensitive to care, fairness, authority, and sanctity, and less sensitive to loyalty (and social norm violations) on the MFQK. Finally, Chapter 5 examined relationships between parents' and children's moral concerns. In general, few relationships were found, and those found were unexpected and difficult to predict. However, using parents' responses on the MFQ, previous relationships between moral foundations and political orientation were replicated. Overall, the current dissertation supports the existence of moral beliefs outside of care and fairness in young children. It highlights that greater diversity of approaches and methods is warranted in moral development research. It also provides useful future directions in the study of MFT and moral development.

List of Abbreviations and Symbols Used

ANOVA	Analysis of variance
BIC	Bayesian information criterion
β	Beta parameter
X^2	Chi-square
CI	Confidence interval
r	Correlation coefficient
α	Cronbach's alpha
DF	Degrees of freedom
DV	Dependent variable
F	F statistic
H	Hypothesis
IQR	Inter-quartile range
Max	Maximum value
R^2	McFadden's pseudo r-squared
M	Mean
Min	Minimum value
MFPT	Moral foundations puppet task
MFQ	Moral foundations questionnaire
MFQK	Moral foundations questionnaire for kids
MFT	Moral foundations theory
OR	Odds ratio
p	P-value

N	Sample size
SDT	Social domain theory
SD	Standard deviation
SE	Standard error
t	T-statistic
USA	United States of America
WEIRD	Western, educated, industrialized, rich, and democratic

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Chris Moore. His support and unparalleled confidence in my abilities provided the encouragement I needed to complete my dissertation. He has consistently trusted my judgment, valued my opinions, and supported my plans. His thoughtful guidance always brought me back to the theory and purpose of my research, and kept me focused on what was important.

I would also like to thank the remaining members of my dissertation committee, Dr. Sophie Jacques and Dr. Sean Mackinnon. I could not have asked for a kinder, more supportive committee. They provided constructive feedback and input, while also celebrating my work and accomplishments. Their direction over the past few years has been invaluable. I would further like to thank Dr. Mackinnon for his assistance with statistical analyses and data visualization.

The completion of my dissertation would not have been possible without the relentless support and effort of our Early Social Development Lab (ESDL) coordinator, Stef Hartlin. Stef is the glue that holds our lab together and keeps it running. She has always been there if I needed help with a project or just someone to talk to. Her expertise in ethics proposals, participant recruitment, and conducting experiments is beyond compare, and I am tremendously grateful for her assistance.

I also want to recognize and acknowledge other members of the ESDL. I was very fortunate to have an army of volunteers and research assistants help me with this project: Ty Pratt, Kyle Cameron, Emma Viner, Ami Nishimori, Tasnim Ali, Olivia MacDonald, Ainsley Loyd, Mya Dockrill, Hayley MacDougall, Natalie Njoku, and Abigail Green. These amazing individuals were there for early morning drives to daycares, for the trials

and tribulations of testing with 2-year-olds, and for last-minute participant bookings. They worked tirelessly on data entry, data checking, and video coding. ESDL undergraduate student Anastasia Gvozdeva also made a significant contribution to this project. Furthermore, my fellow graduate students at the ESDL, including Jenn Richards, Julie Longard, Robin Curtis, and Erin Dempsey, comprehensive students Nicole Ali and Lydia Muyingo, and international students Vivian Li, Fatemah Keshvari, and Rita Xiao, as well as post-doctoral fellow John Corbit, have been an irreplaceable source of encouragement, advice, and aid over the past five years. I would further like to thank Robin Curtis for his work on the MFQK and assistance with implementing it in my study.

Next, I want to thank my friends. Christiane Whitehouse and Noelle Strickland have been there to calm my worries and escape into the woods with me on a moment's notice. Jason Isaacs and Emre Fertan also helped to keep me sane and stay afloat. Taylor Losier, Sam Moore, and DJ Livermore continued to support me from across the country and the world. So many other friends from the Psychology and Neuroscience department have provided motivation, respite, and support. My cohort helped me survive the years of classes and beyond. I was also supported by the community and friendships I found through Camp Rotary and Seven Bays Bouldering.

Finally, my family. To Mum, Dad, and Meghan, thank you for your never-ending love, encouragement, and support. You have always been my biggest cheerleaders. To Pascal, there is so much to thank you for – the food, the laundry, the dishes. However, above all, I am grateful for your love, your steadfast reassurance, and for believing in me when I did not think I could make it through.

The funding agencies who made my research and this dissertation possible are as follows: Killam Trusts, the Social Sciences and Human Research Council of Canada, and the Nova Scotia Graduate Scholarship program.

Chapter 1: Introduction

A quick glance at the daily news headlines will show the wide number of issues on which modern society is divided. From abortion to family composition to end-of-life care, people have different opinions and give importance to different factors. At their roots, these conflicts seem linked to a wide diversity of moral values or beliefs. Some people value independence and new ideas, while others put more emphasis on hierarchy and tradition. Some value national pride and loyalty above all else, while others treasure individuality and personal freedom. It remains unclear exactly how and why these beliefs develop and change over time. It is reasonable to assert that environment plays a crucial role, given the similar ideologies often present within members of a family or community. However, there are similar moral codes across cultures as well (e.g., regarding murder, incest, etc.), which may suggest a more universal, innate moral code.

Although influenced by the earlier work of Jean Piaget, Lawrence Kohlberg is commonly referred to as the pioneer of research on moral development. Kohlberg believed that morality is self-constructed by children through their experiences. His research focused almost exclusively on issues of justice, while subsequent researchers added concerns of care and fairness to the table (e.g., Gilligan, 1982; Smetana, 2013; Turiel, 1983). However, upon examination of existing cultural or societal conflicts, these do not seem to encompass the full range of moral values present across groups. Issues such as respect for authority, religion, or faithfulness are also frequently cited as reasons for opinions or beliefs. In recognition of this, Jonathan Haidt (2012) outlined Moral Foundations Theory (MFT). He proposed several moral 'foundations,' or areas in which we may hold moral values and beliefs, that extended beyond issues of care, fairness, and

justice to include respect for authority, ingroup loyalty, and more. He also asserted that moral development is influenced by both environmental and nativist factors, not simply rationalist self-construction. While MFT has gained traction in adult research (e.g., Frimer et al., 2013; van Leeuwen et al., 2012), it has yet to be examined in children. Therefore, the purpose of this dissertation was to examine MFT in young children, and in doing so broaden the study of moral development beyond issues of care, fairness, and justice. This research will be important for full consideration and understanding of existing societal debates and divisions.

A Brief History of Research on Moral Development

The earliest known study of moral development dates back to 1894 (Osborn, 1894; Wendorf, 2001), when F.W. Osborn asked children what constitutes good and bad behaviour. Based on his research, Osborn stated that morality develops as children become less egoistic, and more altruistic. Since the late 1800s, there have been two primary factors identified in the development of morality: the role of society (i.e., environmental influences), and the role of the individual themselves (i.e., nature; Wendorf, 2001). For example, John Dewey (1892) believed that morality comes from sympathy, and that sympathy develops from the interaction of individual personality with society. In the early 1900s, the study of morality was heavily influenced by the American testing movement (Wendorf, 2001). Researchers attempted to quantify how much morality an individual possessed. This is an arguably impossible and certainly controversial task and may be why this research is largely ignored today.

In the 1920s, Jean Piaget was beginning his career as a psychologist. He is perhaps best known for his work on cognitive development, but he also studied morality.

Rather than attempt to quantify it, he instead approached morality from a cognitive perspective and examined children's moral reasoning, with a particular focus on children's understanding of rules and justice (Piaget, 1948; Wendorf, 2001). While playing marbles, he asked children about the origins of rules and what makes rules fair. He also asked for children's opinions of the actors, authority figures, and punishments described in moral stories. Based on his findings, Piaget (1948) developed a stage theory of moral development. In the first stage, children have little consciousness of rules (up to age 5). In the second stage (ages 5-10), the heteronomous stage, children believe rules are sacred and unchangeable. Punishment is based on observable consequences and should be equal to the misdeed. In the final stage, the autonomous stage (ages 10/11+), children believe rules can be changed if enough people agree. They also take intent or attenuating circumstances into account when serving justice. According to Piaget (1948), progression through these stages depends on cognitive development and experiences with other children. As children mature cognitively, they use reason to interpret different experiences or viewpoints and form their own understanding of moral concepts.

While Piaget was one of the first to study morality in a systematic fashion, many of his findings and assertions about children's abilities have since been disproved. For example, Nelson (1980) found that children much younger than 10 years old (as young as 3) can judge the intentions of others. Similarly, there is evidence that even young children do not treat all rules as the same or absolute, and can distinguish between more serious transgressions (e.g., hitting) and more social-conventional violations (e.g., snacking in class; Smetana et al., 2012). In general, Piaget largely underestimated children's abilities, in part due to flawed methodologies (Gelman & Baillargeon, 1983; Hopkins, 2011).

However, his work is still considered valuable for highlighting the importance of cognitive development in moral reasoning, and for stimulating subsequent research – most notably, the work of Lawrence Kohlberg.

Lawrence Kohlberg developed one of the most well-known theories on moral development, in part by refining Piaget's previous work (Kohlberg & Hersh, 1977). He conducted two-hour moral dilemma interviews with children aged 6-16 years old, and examined their moral decisions, reasoning, and rationales (Kohlberg, 1963). In doing so, he realized that moral development is not complete at ages 10-11, as Piaget had suggested, but instead becomes more complex into adolescence and adulthood. To account for his participants' responses, he developed a theory of moral development that includes 3 levels, and 2 stages in each level (see Table 1). He believed that all children pass through these stages sequentially, with higher stages of moral reasoning replacing (rather than adding to) lower stages (Kohlberg, 1963). Each stage represents a way of thinking about moral dilemmas or forming moral judgments. The stages focus on children's understanding of rights, justice, punishment, and intentions versus consequences of actions (Kohlberg, 1963). Like Piaget, Kohlberg (1963) believed that both social experience and cognitive development are key for moral development. He indicated that progression through these stages occurs as children actively reason about and make sense of the social world around them.

Kohlberg's theory is considered a more accurate description of children's moral reasoning than that of Piaget, but it is not without criticisms. One of the biggest critiques is that his theory is culturally biased towards societies that emphasize and value individualism (Dien, 1982). Shweder et al. (1987) studied Kohlberg's stages in India and

the United States with children aged 5-13 years old. They found that with age, children from the US saw fewer issues as universal moral principles, and more issues as social convention – while the opposite was true for children from India. Overall, they concluded that culture changes how people think about moral issues, and that not all cultures follow Kohlberg's stages.

Table 1

Lawrence Kohlberg's levels and stages of morality

Level	Description	Stages	Orientation	Beliefs
Pre-conventional	Rules are external, not internalized. Child conforms to rules to avoid punishment or get rewards.	1	Punishment and obedience	Judgments of acts are based on consequences and authority.
		2	Naïve hedonism	Reciprocity informs judgments; right actions yield benefits and rewards.
Conventional	Children follow rules and norms for social praise, approval, maintenance of social order, and avoidance of blame.	3	"Good boy" or "good girl"	Behaviour is moral if it is approved by others.
		4	Social-order-maintaining morality	Respecting authority and maintaining social order is highly important.
Post-conventional	Right/wrong is determined by broad principles of justice, rather than rules or authority.	5	Social contract	Good acts or laws are those that have been agreed upon by a society
		6	Morality of individual principles of conscience	What is right is defined by one's own ethical principles, which are abstract and not concrete laws.

Note. Information summarized from Kohlberg (1963) and Kohlberg and Hersh (1977).

Another critique of Kohlberg's theory pertains to gender, and the fact that it was developed from research done predominantly on boys. Gilligan (e.g., Gilligan & Attanucci, 1988) highlighted that the moral judgments of females do not fit neatly within Kohlberg's theory and stages. Based on her own research, she highlighted two different moral orientations: a justice perspective (similar to Kohlberg's stages), used predominantly by males, and a care perspective (focusing on compassion, protection, and attachment), used predominantly by females (Gilligan, 1982; Gilligan & Attanucci,

1988). Although some of Gilligan's claims have also been criticized (e.g., Garmon et al., 1996 found no differences between the moral stages or judgments of men and women), she highlighted the importance of considering all genders in moral development research, and broadened the study of morality to include issues of caring and compassion.

The theories of Piaget and Kohlberg were similar in that they both discussed rules and how children's views of rules develop and change over time. However, Elliot Turiel (1983) noted that there are different types of rules that children encounter: moral rules, which focus on welfare and basic human rights (e.g., no hitting, cheating, causing harm, violating rights), and social-conventional rules, which pertain to social etiquette or social norms (e.g., no snacking in class). Following research by himself and his former PhD student, Judith Smetana, this was later expanded into Social Domain Theory (SDT). In brief, SDT states that morality (encompassing issues of fairness, justice, care, and welfare) is just one piece of children's developing social knowledge (Smetana, 2013). The other domains of social development include social conventions, and personal issues. Social conventions are context-specific, mutually agreed upon, and changeable matters of authority, tradition, and social norms. Personal issues are preferences and decisions that have consequences only for the private domains of life (e.g., bodily control, personal health, privacy, choice of friends). In SDT, each domain (moral, social, personal) is viewed as a separate system with its own developmental trajectory. Like the theories of Kohlberg and Piaget, SDT suggests these domains are self-constructed and developed through interactions with the social environment – again focusing primarily on peer relationships. However, unlike Kohlberg and Piaget, SDT does acknowledge that parents are important. Specifically, parents help children make sense of their social interactions,

explain things in a developmentally appropriate way, and provide emotional support throughout childhood (Smetana, 1999).

To study the developmental trajectories of the different social domains, Smetana et al. (2012) presented 2- to 4-year-olds with social rules interviews over a 1 year timespan (with 93% retention). During the interviews, children were presented with both moral (physical or emotional harm) and conventional (e.g., wearing your bathing suit to daycare) transgressions. Smetana et al. (2012) asked children about the acceptability of the act and appropriate punishments (e.g., "is __ OK? What if the teacher didn't see them? Should they get in trouble?"). All children viewed moral transgressions as more wrong and more deserving of punishment relative to conventional transgressions. Older preschoolers (approximately 3.5 and older) were less likely to punish both types of acts than younger preschoolers, but still punished moral transgressions more than conventional ones. During the year-long study, children grew in their understanding that moral transgressions were wrong regardless of rules or authority. Turiel (2008) conducted a similar study with children in first, third, fifth, and seventh grade. He found children viewed moral transgressions as wrong even in the absence of rules or authority figures. However, conventional transgressions were not considered wrong if no rule existed or a teacher said it was okay. No age differences were observed; these findings were consistent across ages. Therefore, Turiel (2008) concluded that distinctions between moral and conventional domains form at an early age and are maintained as children grow older, with transgressions in the moral domain being seen as more severe.

In summary, the research conducted in the early to mid 20th century sparked decades of research on moral development. The prominent theories presented here – i.e.,

those of Piaget, Kohlberg, Turiel, and Smetana – each build upon the next and therefore have notable similarities. All fall into the realm of rationalism, as reasoning is considered paramount to moral development (Haidt, 2012). They assert that morality is self-constructed by children as they reason about and interpret their social experiences. While Smetana (2013) discussed children's emotional responses to transgressions, rationalist theories typically ignore emotions and certainly do not consider them to influence moral reasoning. These theories are also similar in the criticisms that they face. Most notably, all were created from a Western, Liberal, and individualistic standpoint. As such, they primarily emphasize issues of fairness, justice, and care – all largely focused on treating *individuals* well (Haidt, 2012). Concerns about hierarchy, loyalty, or respect for authority, rules, and tradition, are not viewed as moral, or are seen as 'lower' forms of moral reasoning. Piaget (1948) even said that moral development only occurs as children become increasingly free from their parents' authority. However, when you look at the range of moral concerns across the world, or even within North America, issues of fairness, justice, and care simply do not seem to account for everything.

Cross-Cultural Considerations

In recent decades, research on morality has expanded to include people from different groups and cultures. Shweder et al. (1987), mentioned briefly above, presented 39 short stories to both adults and children in India and the USA. Some acts described in the stories were seen as wrong to both cultures (e.g., kicking a sleeping dog), some were wrong in the USA but acceptable in India (e.g., son claiming most property after father's death, leaving little for daughter), and some were wrong in India but acceptable in America (e.g., 25-year-old son addressing father by first name). Shweder et al. (1987)

found that morality in India was broader than in the USA, such that violations of authority, purity, and tradition, in addition to care, fairness, and justice, were viewed as universally and unalterably wrong. Based on this research, Shweder et al. (1997) asserted that morality extends beyond issues of care, fairness, and justice. They developed a 'big three' of morality: autonomy, community, and divinity. Autonomy pertains to rights violations, and encompasses care, fairness, and justice. The other codes extend beyond these issues; community pertains to violations of hierarchy and community values, and divinity pertains to issues of sanctity and purity.

While Shweder et al. (1997) was one of the first to consider broader, cross-cultural representations of morality, there were some methodological flaws in the research upon which their theory was based. Most notably, they did not consider if people in India viewed more acts as immoral because they ultimately caused harm (e.g., by offending spirits, preventing reincarnation; Haidt, 2012). Wanting to address this confound, Haidt et al. (1993) created 'harmless' taboo stories where no person or spirit was harmed. These fell primarily into categories of disrespect (e.g., a woman cuts up her old American flag for which she has no use into rags) and disgust (e.g., a family eats their pet dog that was killed by a passing car, and nobody sees). They administered these, as well as stories similar to those used by Kohlberg (1963) and Turiel (2008), to adults from different socioeconomic backgrounds in Philadelphia and in different regions of Brazil (characterized by both cultural and wealth differences). In addition to other questions, they explicitly asked their participants whether someone was harmed in each story.

Amongst their participants, Haidt et al. (1993) found considerable variation in the issues and stories that people considered moral. Interestingly, the biggest differences

were between socioeconomic classes, rather than cities or cultures. In general, Americans and upper-class Brazilians viewed a narrower range of concerns as moral, making distinctions between moral and social conventions. In contrast, Brazilians from less wealthy or less industrialized cities, as well as lower-class Americans, moralized more of the stories and included issues of disrespect and disgust in their moral evaluations. This was true even when controlling for perceptions of harm – in fact, controlling for harm made the differences even bigger. Haidt et al. (1993) concluded that the moral domain indeed extends beyond issues of care, fairness, and justice, and varies between groups. Haidt (2012) further highlighted that, based on this early research, existing rationalist theories could not account for the range of moral responses they received – especially pertaining to disgust and disrespect. He suggested that other influences must be present as well, such as environmental, innate, or evolutionary influences.

Moral Foundations Theory (MFT)

Together with Craig Joseph, Jonathan Haidt (Haidt & Joseph, 2008) developed Moral Foundations Theory (MFT). This theory was designed to incorporate findings from Shweder (Shweder et al., 1987, 1997), as well as Haidt's early research described above (Haidt et al., 1993). It encompasses Haidt's views that there are more moral domains than just care, fairness, and justice. Influenced by the growing field of evolutionary psychology, it also outlines influences on moral development other than rationalist self-construction. Finally, when conducting their moral interviews, Haidt et al. (1993) observed that their participants often struggled to come up with a reason for their opinion or belief. Therefore, unlike rationalist and cognitive theories, MFT does not ignore the impact of emotions and intuitions on moral thinking. Altogether, the theory is best

understood by reviewing its four primary tenets: nativism, cultural learning, pluralism, and intuitionism.

Nativism

Borrowing an analogy from Marcus (2004), Haidt (2012) stated that nature provides the "first draft" of morality (p. 152). Innate in this sense does not mean it cannot be changed; it means that morality is "organized in advance of experience" (Graham et al., 2013, p. 61). Haidt (2013) added that innate does not mean present at birth; children may not be able to access certain moral concerns or issues until other development has occurred (e.g., perspective taking, cognitive development, etc.). Haidt (2013) also noted that innate does not mean that the same values and beliefs will be expressed in all people and cultures. Evolution and genetics guide the organization of the brain, but that draft can still be revised by experience. Haidt (2012) specifically outlined that the brain is organized into different areas, which he referred to as 'moral foundations.'

Haidt and Joseph (2008) explained that their concept of moral foundations is similar to the learning modules described by Sperber (1994). These foundations were created through challenges faced by our ancestors over millions of years. Evolution favoured individuals who could solve social problems effectively by detecting patterns and responding appropriately (Haidt & Joseph, 2008). Foundations represent the kinds of concerns that are likely to become moralized, and it is more difficult to get people to care about issues outside of those areas. In response to criticisms from Suhler and Churchland (2011), who described this use of modularity as "murky" (p. 2105), Haidt and Joseph, (2011) acknowledged that it lacks supporting computational and neurobiological data. They were hopeful this would occur in the future, but also asserted that incompleteness is

not a reason to abandon a theory.

To further explain their notion of modularity and innateness, Haidt (2013) borrowed the notion of psychological preparedness from Seligman (1971). A common example of psychological preparedness is that humans are prepared to and easily develop a fear of snakes, while it is rare to develop a fear of flowers. Similarly, moral foundations are areas in which we are more prepared to learn moral ideas and beliefs. For example, it is easy to teach children that hitting is bad, but perhaps harder to teach that it is wrong to dance (see Footloose; Ross, 1984). Haidt and Joseph (2011) further highlighted the emergence of cross-cultural moral phenomena as evidence that morality is psychologically prepared. They gave the example of cooties (where people of the opposite sex or low in popularity are 'contagious'), a childhood game tied strongly to concepts of purity and pollution, which emerges at similar developmental timepoints across the world (Samuelson, 1980). Overall, MFT's version of innateness is that morality is organized through evolution and genetics into moral foundations, or areas where moral concerns are more likely to occur. They are called foundations because they are not the final product. Although foundations limit what morality will encompass, the environment and experience can suppress, alter, or magnify a particular foundation (Haidt, 2013).

Cultural Learning

Haidt and Joseph (2008) continued with Sperber's (1994) notion of modularity to explain how and why morality is impacted by environmental influences. The first draft of morality gets filled in and revised by our experiences and culture, so that we can navigate the culture or society we live in. The capacities of the moral foundations (i.e., in detecting

particular social patterns) are expanded or refined, and develop into sophisticated and varied perceptions, beliefs, virtues, and judgments. This revision process can lead to both the reduction and expansion of a particular moral foundation (Haidt, 2012). For example, in a Hindu community, a moral foundation pertaining to authority may be activated by actions and taught beliefs (e.g., respect for gods or politicians), and expanded to include culture-specific beliefs, facts, and movements (e.g., bowing). In a secular American community, the same foundation may not be activated or expanded as much, so someone's beliefs and practices about authority would be less involved – or they may even hold anti-authoritarian beliefs (Graham et al., 2013).

In addition to modifying and expanding the moral foundations, cultural learning can cause the same action to be linked to foundations in different ways (Haidt, 2012). For example, for some people spanking triggers judgments of harm, while for others it may be linked to respect for rules and authority. This depends on lessons that are taught, previous experience, and the strength and development of a particular foundation. It means that conflicting adult moralities can be built upon the same innate moral foundations (Haidt, 2012). Some cultures use all foundations and build on them in specific ways, while others just build on a few. Therefore, the 'first draft' of morality is revised by culture and experience, producing finished 'books' that look very different across cultures. If you look simply at the finished books, it would be difficult to decipher what comprises the universal first draft of morality – i.e., how many moral foundations there are, and what moral concerns they encompass. However, by examining commonalities in evolutionary theory and anthropological observations, hypotheses can be formed (Haidt, 2012).

Pluralism

Determining the number of basic elements that comprise and describe morality is no easy task. In most circumstances, scientists value simplicity and parsimony. For example, Lawrence Kohlberg (1971) created a parsimonious, monist theory of morality: he believed there was only one moral virtue (justice), which takes the same form across cultures and settings. However, Graham et al. (2013) argued that when you reduce explanations of human behaviour to single constructs, you risk losing "descriptive completeness" (p. 56). They noted that evolutionary thinking supports pluralism (i.e., multiple moral foundations); evolution does not occur in the most parsimonious fashion, but instead creates solutions to problems out of whatever is available and however it can. Evolutionary thinking also supports functionalism, such that evolved structures have a purpose (Graham et al., 2013). Therefore, moral foundations are structures or modules that responded to problems by our ancestors; the best foundations are the ones that could determine a particular social pattern and respond effectively.

Based on evolutionary, cross-cultural, cross-species, and other evidence, Haidt and his colleagues have proposed five moral foundations: care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, and sanctity/degradation (Graham et al., 2013; Haidt, 2012; Haidt & Joseph, 2008). Continuing with Sperber's (1994) theory of modularity, Haidt and Joseph (2008) discussed the 'proper' versus the 'actual' domain for each moral foundation. The proper domain refers to the concerns the foundation evolved to detect. The actual domain refers to the concerns that are actually detected by the foundation, which are broader (Haidt & Joseph, 2008). As humans have experienced rapid societal changes, the many different things that now activate our moral foundations

likely do not contribute to survival in the same way.

For the care/harm foundation, Haidt and Joseph (2008) stated that this evolved from attachment systems designed to protect offspring; parents who were better at detecting the suffering and distress of offspring were more likely to keep them alive. However, the actual domain for most individuals includes distress from people and animals of all ages, cute images, kind acts, and more. For fairness/cheating, the proper domain evolved to facilitate and obtain the benefits of reciprocal altruism from strangers or distant kin (Trivers, 1971). Those who were better able to assess fairness or detect signs of cheating were more likely to get or keep resources that contributed to survival. The actual domain can now detect a wide range of concerns, from cheating and cooperation to rights or institutions of justice. Loyalty/betrayal pertains to the human partiality for organizing into groups and tribes, and the resulting benefit to survival. While the original triggers were likely disloyalty from family or tribe members, current triggers can include displays of loyalty towards nations, sports teams, businesses, occupations – groups that influence individuals' identities and their attitudes towards outgroup and ingroup members (see Social Identity Theory; Tajfel & Turner, 1979). For authority/subversion, both humans and primates form hierarchical societies, so those that can navigate those social relationships appropriately have an increased chance of survival. This foundation is now commonly activated by interactions with bosses, police, leaders, and other social hierarchies. Finally, the sanctity/degradation foundation evolved to detect and avoid disease, rotten food, and other pathogens. However, Haidt and Joseph (2008) argued that this process of evaluation, disgust, and rejection was adapted for social and moral purposes. Concerns about keeping sacred objects safe from destruction

or pollutants, treating the body as a temple, chastity, sexual deviance, or even immigration can be tied to feelings of impurity or disgust, and represent some of the current sanctity triggers.

To provide justification for the moral foundations they proposed, Haidt and Joseph (2008) highlighted that there are similarities in how different cultures deal with these foundations, such as initiation rituals to strengthen ingroups, different markings or clothes to denote hierarchy, or purity and pollution rules about toileting, birth, etc. Furthermore, they noted that their foundations line up quite well with other cross-cultural theories of morality (e.g., Rai & Fiske, 2011; Shweder et al., 1997). Certainly, issues of care/harm and fairness/cheating are frequently included in theories of morality (e.g., (Kohlberg & Hersh, 1977; Smetana, 2013), and so those moral foundations are unlikely to be challenged. However, in anticipation that issues of loyalty, authority, and sanctity were less likely to be viewed as moral by people from liberal, educated, or scientific backgrounds, Haidt and Graham (2009) provided additional justification. They highlighted that loyalty, authority, and sanctity serve many positive functions, and that it is understandable that people or groups are distressed by actions that diminish or disrespect these constructs. For example, Durkheim (1951) found that increased social integration or ingroups (e.g., religion, war, large family) decreased rates of suicide. Durkheim (1973) also emphasized that loss of both traditional authority and strong community leads not to egalitarian liberty, but to anomic anarchy.

As anticipated, the foundations selected by Haidt and his colleagues were not without criticism. For example, Suhler and Churchland (2011) argued that their taxonomy of moral foundations that their taxonomy was contrived, and ignored other

good candidates as well as group differences in foundation content. Haidt and Joseph (2011) acknowledged that the five foundations they proposed are a starting point, and that reorganization could occur with more research. However, they also argued that more principled approaches to building theories (e.g., Kohlberg & Hersh, 1977) will be simple and elegant, but crumble under empirical data. Haidt and his colleagues have openly stated that the existence of other moral foundations is quite plausible, and even went so far as to offer a prize to anyone who could "demonstrate the existence of an additional foundation, or show that any of the current five foundations should be merged or eliminated" (Graham et al., 2013, p. 99). They offered five criteria of foundationhood: the foundation must be a common concern in third-party normative judgments, it must involve automatic, emotional evaluations, it should be culturally widespread, there must be evidence of innate preparedness (e.g., cross-cultural, anthropological data), and an evolutionary model should clearly demonstrate the adaptive challenge it solved. While liberty has been presented as a possible candidate (Graham et al., 2013), the current study will focus on the original five foundations proposed by Haidt and Joseph (2008).

Intuitionism

The final tenet of MFT is intuitionism. Although this tenet is important to the theory, it is less relevant for the current study, and will be described only briefly here. This tenet was the product of Haidt et al.'s (1993) interviews with participants about harmless taboos and moral vignettes. When conducting these interviews, they noticed that people would declare a certain act (e.g., a family eats their pet dog that was killed by a passing car, and nobody sees) to be unacceptable or wrong, but would be unable to provide a reason. Haidt (2012) referred to this as being "morally dumbfounded" (p. 29).

It led him to assert that moral decisions are made based upon intuitive, emotional, or gut reactions. Therefore, a moral foundation is something that elicits a strong, automatic reaction. While Haidt (2012) did state that humans engage in moral reasoning, he argued that this is not done to formulate decisions, but is rather to support the intuitive judgments that were already made.

Empirical Evidence for MFT

In order to empirically examine Moral Foundations Theory (MFT), Graham et al. (2011) developed the Moral Foundations Questionnaire (MFQ, see Appendix A). They sought to determine individuals' level of concern for each moral foundation. They generated many items representing each foundation, and administered these items to a large, heterogenous sample ($N = 3825$). They then narrowed down these items by looking at the underlying factor structure, item loadings, scale internal consistency, and correspondence with existing scales. The final version of the MFQ is a 32-item questionnaire, with 6 questions for each of the five moral foundations (care, fairness, loyalty, authority, and sanctity), and 2 'catch' items to determine inattentive responses. The MFQ is divided into two parts: moral relevance items, where participants rate how much different concerns influence their judgments of right and wrong (e.g., whether or not someone suffered emotionally), and moral judgment items, where participants rate their agreement with different statements or actions (e.g., it can never be right to kill a human being).

With the final version of the MFQ complete, Graham et al. (2011) sought to test its reliability and validity with a large online sample of adults ($N = 34,476$). They found that the separate scales for each moral foundation showed good internal consistency,

especially given the small number of items (Cronbach's α range from .65-.84). Test-retest reliability (ranging from 28-43 days between tests) was also good; Pearson correlations for each foundation were .71 for care, .68 for fairness, .69 for loyalty, .71 for authority, and .82 for purity (all *ps* <.001). To examine the underlying factor structure of the MFQ, Graham et al. (2011) conducted both an exploratory and a confirmatory factor analysis, and found that a five-factor model was indeed the best fit for the data. Finally, they examined external validity by looking at correlations with and predictive validity for existing measures. They found that scores on each moral foundation showed positive correlations (average $r = .51$) with relevant measures (e.g., a measure of empathy for the care foundation, religious attendance for the sanctity foundation). They also found that scores predicted attitudes towards relevant social groups (e.g., labour unions for the fairness foundation, military for the authority foundation), and that the MFQ was a better predictor for relevant scales, opinions, and behaviours than an existing comprehensive measure of moral values. Overall, Graham et al. (2011) found good evidence for the internal and external validity of the MFQ, as well as evidence to support MFT itself – for example, support for the existence of five separate moral foundations.

Armed with a good measure of their five proposed moral foundations, Haidt and his colleagues sought to establish the pragmatic validity of MFT through novel, useful, empirical advances (Graham et al., 2013). One of their first areas of study was to look at cross-cultural differences in moral foundations. Using their large online sample, Graham et al. (2011) compared participants from Eastern cultures ($n = 2258$) and Western cultures ($n = 104, 983$). Those from Eastern cultures scored slightly but significantly higher on the Loyalty and Sanctity foundations relative to those from Western cultures, in line with

previously found cultural differences in collectivism and concerns about purity (Graham et al., 2013; Shweder et al., 1997; Triandis et al., 1988). These differences have also been replicated on non-English speaking samples (Kim et al., 2012). Interestingly, cross-cultural differences have been linked to historical threats faced by countries. For example, van Leeuwen et al., 2012) found that historical country levels of pathogen prevalence predicted scores on the sanctity, loyalty, and authority foundations, while van Leeuwen and Park (2009) found that increased perceptions of social danger correlated with higher scores on the 'binding' foundations (i.e., loyalty, authority, and sanctity). This also provides support for some of the original triggers underlying moral foundations proposed by Haidt (2012).

Although evidence supports cross-cultural differences in moral foundations, the small size of these differences suggests that within-culture variation may be greater than between-culture variation (Graham et al., 2013; Vauclair & Fischer, 2011). Indeed, some of the biggest differences in moral foundations have been found between people from the same country who have different political orientations. In an online sample of 1613 adults, Graham et al. (2009) found that the care and fairness foundations were highly endorsed by all individuals, with self-identified liberals scoring slightly but significantly higher than self-identified conservatives. In contrast, the loyalty, authority, and sanctity foundations were rated significantly higher by conservatives than by liberals, who on average gave little value to those foundations. This effect is very robust; it has been found in the USA, United Kingdom, Canada, Australia, the Netherlands, and across Asia, Europe, and South America (Graham et al., 2009, 2011; van Leeuwen & Park, 2009). It suggests that the most complex political debates are likely to highlight different opinions

on respect for tradition, authority, and physical or spiritual purity, whereas issues of care and fairness are more likely to be common ground. Overall, these differences highlight that morality is indeed broader than issues of care and fairness.

In addition to group differences in moral foundations, empirical evidence supports the claims of MFT that biology is an important factor in morality (Haidt, 2012). Researchers have examined the relationship of dispositional and personality traits with moral and political values. For example, Inbar et al. (2012) found that heightened disgust sensitivity was associated with increased political conservatism. Similarly, Lewis and Bates (2011) found significant correlations between the Big Five personality traits and moral foundations. They found that higher scores on the care and fairness foundations were associated with more openness, neuroticism, and agreeableness, while higher scores on the loyalty, authority, and sanctity foundations were associated with greater conscientiousness and extraversion, and less neuroticism. People with dissimilar dispositions may approach or interpret events differently, and therefore develop unique moralities even with similar life experiences (Graham et al., 2013).

Other evidence supporting the role of biology has been psychophysiological. Lewis et al. (2012) found that higher scores on particular foundations (e.g., care) were associated with increased gray-matter volume in specific brain areas or structures (e.g., the dorsomedial prefrontal cortex, an area associated with empathy). Similarly, using fMRI, Parkinson et al. (2011) found that reading stories about physical harm, dishonesty, and sexual disgust activated different neural systems in participants' brains. For example, stories about sexual deviance caused greater activation the anterior insula (associated with moral emotions, including disgust), relative to stories about harm and dishonesty.

This research not only supports the influence of biology on morality, but also favours pluralism and the theory that morality is not based upon one single virtue, construct, or mental faculty. Overall, there is compelling evidence that morality is influenced by more than rationalist self-construction.

The final area of adult empirical evidence that will be reviewed here pertains to the relationship between morality and emotions. Of the different emotions, anger and disgust have been the most studied. Disgust is thought to be integral for judgments within the sanctity foundation. Helzer and Pizarro (2011) manipulated feelings of disgust using reminders of physical cleansing and found that participants given this reminder were more unforgiving of sexual purity violations. In a similar study, Horberg et al. (2009) found that this effect was unique to sanctity violations; disgust manipulations did not impact judgments of justice or care/harm. However, it is worth noting that a replication study conducted using this type of priming (i.e., reminders of cleanliness) suggested that such effects are likely smaller than originally detected (Johnson et al., 2014). Using different methods, Russell and Giner-Sorolla (2011) studied both disgust and anger. They gave participants vignettes which varied on sanctity (high vs. low taboo), harm (self vs. other), and intent to commit a violation (intent vs. no intent). They found that anger and disgust each contributed to moral judgments. Disgust specifically responded to bodily violations, not harm, and was not influenced by intentionality. Anger responded to disgust and harm, but was impacted by intentionality (i.e., more intent = more anger). In general, this research shows that anger and disgust are both tied to moral judgments. They appear to respond differentially to different types of moral violations, and anger seems more open to revision based on new information (Graham et al., 2013). Most

importantly, the pluralism and intuitionism of MFT have enabled the study of different moral emotions with diverse moral concerns, which were previously unstudied in monist or rationalist theories (Graham et al., 2013).

Developmental Considerations of MFT

A noteworthy limitation of the research presented above is that it was conducted exclusively using adult populations. However, MFT is in part a developmental theory and makes assertions regarding how morality develops over time. MFT suggests that it is easy for children to learn or develop virtues tied to moral foundations, but hard to develop virtues outside of those areas (Jonathan Haidt & Graham, 2009). Haidt and Bjorklund (2007) highlighted that the innate moral foundations emerge at different developmental timepoints, so children will start showing concerns (e.g., harm, purity) at different ages. Their reactions may at first be inappropriate or ineffective, and become more refined with learning and experience. Finally, because MFT suggests that morality is revised through experience, the virtues and concerns children develop should be related to their culture, environment, and experiences (Haidt & Bjorklund, 2007). Graham et al. (2011) highlighted the need for descriptive research on children's moral judgments to support the claims and tenets of MFT, especially regarding loyalty, authority, and sanctity. The paragraphs that follow represent a summary of the existing developmental literature pertinent to each foundation. Of note, the focus is on early childhood, as it is most relevant to the current study.

Care

There is a vast amount of research on children's concern for issues of care and harm, starting from a very early age. In a well-known study, 6- and 10-month-old infants

watched as a puppet unsuccessfully tried to climb a hill (Hamlin et al., 2007). They then saw a helper push the puppet up the hill, or a hinderer push it down. There was also a 'neutral' puppet who pushed an inanimate object up/down a hill. Hamlin et al. (2007) found that infants preferred the helper over the hinderer, as well as the helper over the neutral puppet and the neutral puppet over the hinderer. This, as well as subsequent studies, suggested that children at a young age prefer others who help, and dislike those who hinder or harm. This has been found in children as young as 3 months old, and later in the preschool period (Hamlin & Wynn, 2011; Vaish et al., 2010). It has generally been studied in the context of completing a goal (e.g., climbing a hill, opening a box, retrieving a toy), and is apparent in children's choices, prosocial behaviour, and people they trust (Hamlin et al., 2011; Hamlin et al., 2007; Hamlin & Wynn, 2011, 2012; Vaish et al., 2010). Furthermore, even infants will take intent to help vs. hinder into account (Hamlin, 2013).

Studies conducted with toddlers and preschoolers shed even more light on the beliefs children hold regarding issues of care and harm. Children as young as 18 months show helping behaviour that appears to be intrinsically motivated, and will help others even under conditions of anonymity (Engelmann et al., 2016; Hepach et al., 2012, 2017). They also act prosocially towards victims of harm, and protest actions that harm others (Vaish et al., 2011). In general, children genuinely want to see other's needs met, and show concern for the welfare of others (Hepach et al., 2012, 2016, 2017). Furthermore, when asked explicitly, children as young as 2 years old view both physical and emotional harm as wrong and deserving of punishment, regardless of authority figures or rules (Helwig et al., 2001; Jambon & Smetana, 2014; Smetana et al., 2012). This continues into

adolescence and adulthood, although children are more able to take the situation into account (e.g., necessary harm) and less likely to punish others as they grow older (Helwig et al., 2001; Jambon & Smetana, 2014). Overall, there is good evidence to show that infants, toddlers, and preschoolers are sensitive to issues of care and harm.

Fairness

Fairness, like harm, has also been extensively studied from a young age. In a number of studies, toddlers and preschoolers have watched actors distributing resources equally or unequally between two third-party recipients. As young as 15 months old, children prefer and expect others to distribute resources fairly (Geraci & Surian, 2011; Schmidt & Sommerville, 2011; Sloane et al., 2012; Sommerville et al., 2013). This has been shown through children's looking time, prosocial behaviour, and associations of praise versus reprimands (DesChamps et al., 2016; Olson & Spelke, 2008; Sommerville et al., 2013; Surian & Franchin, 2017). At 3 years old, children even show emotional reactions to unfair distributions, and will also protest unfair distributions to themselves and others (LoBue et al., 2011; Rakoczy et al., 2016). In terms of their own distributive behaviour, children will divide resources equally between two friends as young as 2 years old (Li et al., 2016). Children's distributive fairness increases with age over the preschool period, a correlation that is mediated by their numerical understanding (Chernyak et al., 2016).

While children share fairly between two third-party recipients from a young age, the story is different when it comes to sharing resources between themselves and someone else. Children show evidence of a judgment-behaviour gap for sharing; they believe that both themselves and others should share fairly as young as age 3, but do not

share fairly themselves until the ages of 7 or 8 (Fehr et al., 2008; Li et al., 2016; Smith et al., 2013). However, there are a few exceptions to this. For example, 3-year-olds will share equally with someone they collaborated with to complete a task, and 18-24-month-olds are more generous when there is no possession of resources indicated (Melis et al., 2013; Ulber et al., 2015; Warneken et al., 2011). Indeed, there is some evidence that children show the underpinnings of equity from an early age. Toddlers (21-month-olds) expect someone who worked harder to receive more resources (Sloane et al., 2012), and 3-year-olds share more with a collaborator than a free rider (Melis et al., 2013). In a forced choice scenario, 3- to 5-year-olds will distribute more resources to recipients who worked harder, but still prefer to distribute resources equally if possible (Baumard et al., 2012). Overall, although children tend to share selfishly until later in childhood, the research shows that toddlers and preschoolers understand and care about concepts of equality, generosity, and, to some extent, equity.

Loyalty

In comparison to care and fairness, loyalty has not been studied as much in infants or toddlers. One study did find that 9- and 14-month-olds preferred others with similar food preferences (Hamlin et al., 2013); this could be seen as early signs of ingroup bias, but could also just be due to similarity. Most studies start at the age of 3. At this age, there is evidence that children have a bias, and will act more prosocially (e.g., share more), towards members of their ingroup (Benozio & Diesendruck, 2015; Buttelmann & Böhm, 2014; Enesco et al., 2011; Plötner et al., 2016). This has typically been studied using the minimal group paradigm, where children are assigned a group based on a superficial characteristic (e.g., shape on a coin; Atkin & Gummerum, 2012). Ingroup bias

has also been found towards friends, family members, and ethnic group in children aged 3-4 years old (Enesco et al., 2011; Lu & Chang, 2016; Olson & Spelke, 2008). Ingroup bias appears to strengthen with age (Benozio & Diesendruck, 2015; Buttelmann & Böhm, 2014; Enesco et al., 2011), although Mulvey (2016) found that it decreased with age. In contrast to ingroup bias, outgroup hate (i.e., negative actions against an outgroup) develops later, and is not seen until age 8 (Benozio & Diesendruck, 2017; Buttelmann & Böhm, 2014). There is some evidence that both ingroup bias and outgroup hate are stronger in boys (Benozio & Diesendruck, 2015; Buttelmann & Böhm, 2014).

Ingroup bias and outgroup hate both represent the behaviours that children show, and are likely related to opinions of loyalty and disloyalty. However, it is also important to assess how children assess and judge these concerns; morality is inherently tied to evaluating the actions of others. This has been primarily studied in older children, with a few studies looking at children aged 4 or 5 years old. Children at this age view teasing or excluding ingroup members as wrong regardless of rules or authority, and reject disloyal ingroup members (Abrams et al., 2003; Rhodes & Chalik, 2013). They also view disloyalty against their ingroup as more negative than disloyalty against an outgroup (Abrams et al., 2003, 2008; Atkin & Gummerum, 2012), and generally hold ingroup members to higher standards than outgroup members (Schmidt et al., 2012). Children's concerns about group loyalty grow stronger with age and into adolescence (Abrams et al., 2003; Fu et al., 2016). Interestingly, there is some evidence that children see harm violations as worse than loyalty violations (Atkin & Gummerum, 2012). However, it is worth noting that this research was conducted from a Western perspective with a Western sample. In summary, there is evidence of ingroup bias at age 3, and concern for group

loyalty at age 4. However, studies have not been conducted below these ages, so it is difficult to determine when these concerns first emerge.

Authority

Like loyalty, concern for issues of authority has not been studied in infants or toddlers; the earliest studies start at age 4. Furthermore, it has almost exclusively been studied from a liberal, Western perspective, which does not view respect for authority as moral, and views unquestioningly accepting rules as a more immature stage of moral reasoning. However, it is still worth summarizing this literature. There is some evidence that children see disrespect for authority as wrong; elementary school children will reject playmates that are disobedient and have low respect for authority (Martín-Antón et al., 2016; Young & Avdzej, 1979). Four- to 10-year-olds will also endorse compliance even if they disagree with a rule (Smetana et al., 2014). Across different cultures, respect for authority has been found to be important in mother-daughter relationships; Dixon et al. (2008) found that less respect for authority predicted greater mother-daughter conflict. However, respect for parental authority varies by culture, social, and political contexts, with increases seen in more conservative Christians, in Latina and African American families, and following acts of terrorism (Dixon et al., 2008; Fang et al., 2003; Starks & Robinson, 2005).

The most frequent line of study when it comes to children's beliefs about authority has been when and in what contexts children respect or believe in authority. Children in grades kindergarten to six do not see authority figures as having authority in all situations or events (Kim, 1998; Laupa & Turiel, 1993). For example, children reject the authority of their school principal outside of the school context (Laupa & Turiel,

1993). Similarly, children's acceptance of authority varies by type of event. Smetana et al. (2014) found that children are more accepting of rules about moral (e.g., stealing) than personal (e.g., what toy to play with) actions. In general, studies have found that children's acceptance of authority decreases with age. Lagattuta et al. (2010) found that between the ages of 4 and 7, children develop the knowledge that people can sometimes resist authority and feel good about it. Similarly, Kuhn and Laird (2011) found that between grades 5 and 7, children's belief in the legitimacy of parental authority declines. Overall, children's concern for authority has not been studied at ages younger than 4, and valuing authority has typically been viewed by researchers as a more primitive form of moral development, which could bias existing research findings.

Sanctity

Finally, the research on the development of sanctity and related concerns is extremely limited. We found only 3 articles that dealt explicitly with the moralization of sanctity-related concerns. Rottman and Kelemen (2012) found that 7-year-olds will moralize novel behaviours (e.g., aliens putting sticks on their head) relatively easily, and that both norms of unnaturalness (i.e., being told an action violated natural order) and feelings of disgust contributed to the development of new purity morals. Hahn and Garrett (2017) found that children as young as 3 years old view harming the environment as morally wrong, which could be considered a sanctity-related concern (i.e., environment and nature are sacred). In a study with 6- to 10-year-olds, Helwig and Prencipe (1999) found that the older children viewed flag burning as morally wrong, but the younger children did not understand the symbolic nature of a flag.

In our search for sanctity-related literature, we also found a number of articles that

dealt with the emotion and label of disgust (which is commonly associated with this foundation). Stevenson et al. (2010) found that parents influenced children's (2-16 years of age) disgust reactions to different elicitors (e.g., sneezing on food), and that facial expression is important to this influence. Stevenson et al. (2014) further found that both fear and disgust increase preschooler's attention to environmental threats (e.g., rotten food). Furthermore, multiple studies (e.g., Askew et al., 2014; Muris et al., 2012) have established a relationship between disgust and fear in school-aged children. Danovitch and Bloom (2009) found that children in grades kindergarten, 2, and 4 will use the term 'disgusting' to describe non-physical moral violations (e.g., being mean), but use it more consistently for physically disgusting events.

Lastly, other literature examined in our literature search for the sanctity foundation included studies examining children's perception of handwashing, disease transmission, or related concepts. However, the literature we found dealt predominantly with children's understanding of such concepts, rather than their moral beliefs about them (e.g., Dingman et al., 2020; Raman & Gelman, 2005, 2008). Overall, the research on children's concern for issues of sanctity and purity is very limited, especially in toddlers and preschoolers. It is difficult to draw any concrete conclusions regarding the beliefs young children hold under this foundation.

Dissertation Aims

As previously described, Moral Foundations Theory (MFT) makes specific claims regarding how morality develops. These claims predicate the necessity of developmental research to support and establish the theory. Descriptive research throughout childhood on children's concern for a wide range of moral issues is necessary to determine when the

different moral foundations emerge and how they develop over time (Graham et al., 2011). There is also a need for research examining different aspects of children's environments and experiences, and their relationships with children's moral concerns. Such environmental relationships will determine if and how cultural learning impacts moral foundation development. Furthermore, as outlined in the previous section, there is much research examining children's concern for issues of care and fairness from a very early age. However, the research on authority and loyalty is more limited, especially in early childhood, and these concerns have not typically been viewed as issues of morality. The research on sanctity is even more inadequate, with only a few preliminary investigations into children's feelings of disgust and sanctity-related concerns. Therefore, there is certainly a need to broaden the study of morality in developmental literature, particularly with regards to early development. To address these concerns, the aims of the current dissertation were as follows:

1. Examine 2- and 4-year-olds' level of concern for the five moral foundations proposed by Haidt and colleagues: care, fairness, loyalty, authority, and sanctity (e.g., Graham et al., 2013). Specifically, we examined the moral sensitivities of 2- and 4-year-olds using a behavioural puppet measure, similar to the work of Hamlin et al. (2007).
2. Examine cross-sectional changes in moral concerns by comparing responses of 2- and 4-year-olds to the same behavioural measure. This allowed for a preliminary investigation into how children's moral concerns may change with development and cultural learning.
3. While behavioural puppet tasks are most suitable for research with 2-year-olds,

they limit the amount and type of data that can be obtained. Therefore, the third aim of the current study was to conduct a more in-depth study of 4-year-olds' moral concerns using a previously developed and validated measure, the Moral Foundations Questionnaire for Kids (MFQK). This aim included a comparison of the two measures (puppet task and MFQK) administered to 4-year-olds.

4. In addition to examining concerns within each foundation, another goal of this dissertation was to investigate patterns of correlation between each of the foundations. For example, studies have found that care- and fairness-related behaviours develop at similar ages (Schmidt & Sommerville, 2011; Sommerville et al., 2013). Furthermore, Haidt (2012) distinguished between the 'individualizing' foundations of care and harm, versus the 'binding' (i.e., group-related) foundations of loyalty, authority, and sanctity. Therefore, we thought it would be important to investigate associations within and outside of the different 'categories' of moral foundations.
5. Finally, the last aim of the current study was to examine relationships between parents' concern for the moral foundations (assessed using the Moral Foundations Questionnaire; MFQ) and children's responses on either of the tasks administered. During the toddler and preschool years, parents are one of, if not the largest influence on the beliefs and opinions children are taught and the experiences or environments they encounter. Therefore, examining patterns of correlation represents a preliminary investigation into MFT's tenet of cultural learning (Graham et al., 2013).

This dissertation took the form of one large data collection project. Over the

course of a year, we administered our measures to a large group of 2-year-olds (behavioural puppet task), 4-year-olds (behavioural puppet task and MFQK), and their parents (MFQ). The methods for each of these measures are outlined in the following chapter. The results pertaining to each study objective are spread across the three subsequent chapters; Chapter 3 covers Objectives 1, 2, and 4, Chapter 4 covers Objectives 3 and 4, and Chapter 5 deals primarily with Objective 5. Of note, more specific aims and hypotheses are outlined in each relevant chapter.

Chapter 2: Methods

Participants

Following study approval from the Dalhousie University Research Ethics Board, participants were recruited from the Halifax Regional Municipality and surrounding areas. Parent/guardians provided informed consent prior to participation, and children provided assent. Some participants ($N = 81$) were tested at our university lab following responses to study advertisements on Facebook or sent out through our lab database. Others ($N = 97$) were tested at their daycare facilities if parents returned consent forms that had been distributed by daycare teachers. Children had to be typically developing to participate, and both parent/guardians and children had to speak English fluently.

In total, 211 children, and their parent/guardians, were recruited. There were 135 two-year-olds (76 male, 59 female). However, 33 two-year-olds (21 male, 12 female) were excluded due to poor attention, refusal to participate, lack of comprehension, and/or language delays or other developmental delays. This left a final sample of 102 two-year-olds (55 male, 47 female) and 76 four-year-olds (39 female, 37 male; no 4-year-olds were excluded). A-prior power analyses were conducted for the main analyses anticipated prior to data collection; the parameters and results of these analyses are shown in Table 2. Of note, when it came time for data analysis, some tests changed upon examination of the data or other considerations. For example, we conducted generalized linear models rather than correlations to determine relationships between parents and children's morals (see Chapter 5). However, the majority of the analyses stayed the same. Results of power analyses indicated that 49-128 were required. Therefore, we initially planned to recruit approximately 75 two-year-olds and 75 four-year-olds (150 total). However, based on the

response patterns of 2-year-olds and the frequency of missing data, additional 2-year-olds were recruited.

Table 2

Results of a-priori power analyses

Test	Anticipated analysis	α	Power	Effect size	Effect size justification	# of tails	*N*
Binomial probability	Puppet preferences	.05	0.8	Detect if proportion of 0.7 is different from chance	Large effects (e.g., proportions of 0.8 or greater; Hamlin et al., 2007) seen in studies examining toddlers' moral puppet preferences	Two	49
Fisher's exact test	Differences in puppet preferences between age group	.05	0.8	Detect a difference of 0.25 between proportions from 2 groups	Previous research showing small to medium differences between 2- and 4-year-olds, e.g., Smetana et al. (2012)	Two	128
Bivariate correlation	Correlations between foundations	.05	0.8	Medium correlation, $r = 0.3$	Medium to large correlations found between individualizing and binding foundations (e.g., Graham et al., 2011)	Two	84
Bivariate correlation	Correlations between parents' and children's morals	.05	0.8	Small to medium correlation, $r = 0.25$	Small to medium effect sizes for parent influences on children's moral behaviours (e.g., (Karmakar, 2015; Sengsavang & Krettenauer, 2015)	Two	123

Full descriptive data for the final, useable sample are presented in Table 3. Two-year-olds were on average 30.77 months old (*SD* = 3.55 months), while 4-year-olds were on average 52.66 months old (*SD* = 3.37 months). One parent/guardian of each participant completed a demographics questionnaire. Parent/guardian respondents were most frequently female (86.5%), identified as white (87.6%), and had completed some form of postsecondary education (90.7%). Annual household incomes varied, with 23.5% less than $75000 (Canadian), 23.6% ranged from $75000 to less than $100000, and 50.6% were $100000 or more (2.2% did not respond). We also assessed the religious and political orientations of parent/guardian respondents. The most common religious

Table 3

Demographic information for the final, included sample

Category	Options endorsed*	Percentage
Child gender	Female	48.3%
	Male	51.7%
Parent/guardian gender	Female	86.5%
	Male	13.5%
Gross family income	Less than $25000	4.5%
	$25000-less than $40000	5.6%
	$40000-les than $75000	13.5%
	$75000-less than $100000	23.6%
	$100000 or more	50.6%
Parent/guardian education	Some high school	0.6%
	High school diploma or equivalent	0.6%
	Some college, trade school, or equivalent	3.9%
	Some university	3.9%
	Diploma from college, trade school, or equivalent	25.3%
	Bachelor or undergraduate degree	44.9%
	Some graduate school	1.7%
	Master's	23.0%
	Degree in medicine, dentistry, veterinary medicine	0.6%
	Earned doctorate	4.5%
Parent/guardian cultural background, race, or ethnicity	White	87.6%
	Black or African Nova Scotian	1.7%
	Hispanic, Latino, or Spanish	0.6%
	Asian	2.2%
	Asian Indian	1.1%
	Middle Eastern	0.6%
	First Nations, Inuit, or Metis	0.6%
	Bi-racial	5.1%
	Other	0.6%
Parent/guardian religious orientation	Christian	41.6%
	Muslim	1.1%
	Buddhist	1.1%
	Hindu	1.1%
	Jewish	1.7%
	Atheist	4.5%
	Agnostic	10.7%
	Non-religious	34.3%
	Other	3.9%
Parent/guardian political orientation	Liberal	49.4%
	Conservative	8.4%
	Moderate	14.6%
	Socialist	3.9%
	Libertarian	0.6%
	Environmentalist	7.3%
	Other	12.9%

**Note*. Response options that were not endorsed are not shown here; full survey with response options is shown in Appendix B. Percentages that do not add to 100 mean the remaining percentage chose not to provide a response for that question (there were no missing surveys).

orientations were Christian (41.6%) and atheist, agnostic, or non-religious (49.5% combined). The most common political orientations were liberal (49.4%), moderate (14.6%), other (12.9%), and conservative (8.4%).

To examine differences between 2-year-olds that were included versus excluded, we compared age, income, and parent/guardian education using three Independent Samples Mann-Whitney U tests. There was no difference in the level of parent/guardian education attained (p = .178) or family income (p = .057). Included 2-year-olds (M = 30.77 months, SD = 3.55) were significantly older than excluded 2-year-olds (M = 28.48 months, SD = 3.17; p = .001).

Measures and Procedure

Moral Foundations Puppet Task

The Moral Foundations Puppet Task (MFPT) was completed by both 2- and 4-year-olds. It is a behavioural task used to examine children's concern for the five moral foundations proposed by Haidt (2012). The stimuli for this task were videos of puppets engaging in different moral acts. There were ten videos, two for each foundation. In each video, one 'bad' puppet committed a moral transgression and another 'good' puppet was morally conforming (according to the designated moral foundation). Each video showed both puppets engaging in these acts twice to allow for better comprehension and habituation to the acts. The same two puppets were used for each video. However, which puppet was 'good' or 'bad' was counterbalanced. The act the child saw first (good vs. bad) was also counterbalanced. The 'good' and 'bad' acts for each video and moral foundation are shown in Table 4.

Each video was presented to participants according to the same procedure.

Following introductions and assent, participants sat on a mat on the floor. In front of them was a cardboard 'house' with a window. There were two experimenters sitting on either side of the house; one experimenter knew which puppet was good or bad, and the other experimenter was blind to this information (and could not see the video). The participant was introduced (e.g., "This is cow, and this is tiger.") to the two hand puppets for the relevant video. They were also informed that the puppets "lived in the house." Subsequently, the puppets went into the house (i.e., behind the cardboard house). Through the window of the house, there was a computer screen on which the puppet videos played. This set-up was based on research showing that 2-year-olds will take a scene presented on a computer screen as real if they are made to believe that they are looking through a window (Troseth & DeLoache, 1998). Showing videos rather than live shows also ensured greater consistency across participants.

After watching the video of the two puppets engaging in moral acts, the blind experimenter brought the puppets back out of the house. They asked the children two questions: "Which puppet do you like best?" and "Which puppet did something bad?". Children could point or respond verbally. Two manipulation checks were also asked to examine children's comprehension of the videos. Participants were asked: "Which puppet [description of good act]?" and "Which puppet [description of bad act]?". For example, for the care foundation, Scenario 1, participants were asked: "Which puppet hugged?" and "Which puppet hit?". Children's responses were recorded as 'good puppet,' 'bad puppet', 'neither' or 'both', and 'unclear'. Each task was also videotaped for reliability coding. After participants responded to each scenario, they were given stickers as a reward. All 10 videos were typically presented in one testing session. However, if

necessary (due to attention, daycare schedules, etc.), the scenarios were split over two testing sessions (approximately 10% of participants).

Table 4

'Good' and 'bad' acts for each scenario and moral foundation

	Scenario 1		Scenario 2	
	Good puppet	Bad puppet	Good puppet	Bad puppet
Care	Hugs another puppet	Hits another puppet	Pets a dog	Kicks a dog
Fairness	Distributes 6 cookies equally between 2 puppets	Distributes 6 cookies unequally (5 vs. 1) between 2 puppets	Keeps 3 stickers and shares 3 with another puppet	Keeps 5 stickers and shares 1 with another puppet
Loyalty	Chooses (i.e., stands beside) own team vs. opposite team*	Chooses (i.e., stands beside) opposite team vs. own team	Helps own team lift something heavy	Helps opposite team lift something heavy
Authority	Goes to bed when asked by parent	Says 'no' and does not go to bed when asked by parent	Gets off table when asked by parent	Says 'no' and does not get off table when asked by parent
Sanctity	Smells flower	Stomps on and destroys flower	Smells pile of feces, shakes head and keeps walking	Smells pile of feces, rolls in it, and walks away

*Note: Teams were identified by shirt colour (green vs. blue).

Moral Foundations Questionnaire for Kids

The Moral Foundations Questionnaire for Kids (MFQK; Curtis & Moore, 2019, unpublished study) was completed by 4-year-olds only. This is a verbal questionnaire where children answer questions based on short vignettes. It measures concern for the five moral foundations proposed by Haidt (2012), as well as odd behaviours or social norm violations (e.g., eating soup with a fork). There were four vignettes for each of the five foundations, as well as the odd behaviours. Each vignette showed a child (same gender as the participant) committing a moral transgression relevant to a particular foundation, or an act that violated social norms. All vignettes are shown in Table 5. The majority of 4-year-olds completed the MFQK after the MFPT; however, a small subset of participants (approximately 15%) completed it beforehand due to the availability of

daycares, parents, or lab volunteers.

Table 5

Vignettes for the Moral Foundations Questionnaire for Kids

Care	Fairness	Loyalty	Authority	Sanctity	Odd Behaviours
You see a girl punch another girl in the stomach	You see a girl cheating in a race by taking a shortcut	You see a girl telling secrets about her sister to people her sister doesn't like	You see a girl ignore her parents when they tell her to stop watching TV	You see a girl rubbing poop on herself in the shower	You see a girl sleeping in her raincoat, instead of sleeping in pajamas
You see a girl calling a girl stupid	You see a girl cheating in a board game	You see a girl reading her sister's diary	You see a girl wearing a hat at school, even after the teacher asks her not to	You see a girl using a dirty diaper as a pillow	You see a girl wearing her pajamas to school
You see a girl stomp on the tail of her pet cat	You see a girl cut to the front of the line	You see a girl teach a secret password to people who are not in her club	You see a girl calling her parents bad words	You see a girl drinking pee with her dinner	You see a girl eating her soup with a fork
You see a girl kick a stray dog	You see a girl taking all of the cookies, leaving none for others	You see a girl score a goal against her own team to help the other team win	You see a girl calling her teacher bad words	You see a girl loudly burping and farting while eating	You see a girl eat dessert before dinner is served, instead of eating it afterwards

**Note*. The gender of the vignette was matched to participant gender; only female items are shown here.

The MFQK was administered on a touch-screen computer using SuperLab 5 software. Each vignette was presented in the same way. A recording of the vignette (e.g., 'you see a boy calling another boy stupid') was played out loud, while a picture (see Appendix C) was displayed on the computer screen. Following the vignette, the child was asked (via recording) "Is [action] bad?", and two large buttons ('yes' and 'no') appeared on the screen. If the child chose 'yes,' they were then asked, "How bad is it?". Participants responded to this question using a 5-point pictoral Likert scale ranging from not very bad (smallest and lightest coloured circle) to very very bad (largest and darkest coloured circle; see Appendix D). They were also asked "Should he/she be punished?", again with yes and no response options. Finally, regardless of whether they answered

'yes' or 'no' to "Is this bad?", all participants were asked "How does him/her doing this make you feel?". They responded using a pictorial emotion scale (see Appendix E), and response options included happy, sad, angry, yucky, and neutral. Prior to starting the MFQK, all participants completed a short training session to ensure comprehension of the yes and no buttons, as well as the severity and emotion scales. An experimenter was also close by in case any questions or problems arose.

Summary scores were created for each foundation, as well as the odd behaviours scale. If a child answered 'no' to "Is this bad?", their response was coded as 0. If a child answered 'yes', the code for their response corresponded to their selection on the severity scale, ranging from 1 (not very bad) to 5 (very very bad). We then created an average of the 4 vignettes for each foundation and odd behaviours scale. This was only calculated if the participant had recorded a response for all 4 items, otherwise it was considered missing data. Summary scores ranged from 0 to 5, with higher numbers reflecting greater concern for the relevant foundation. Responses to the punishment and emotion questions were recorded as categorical variables. For the moral foundation summary scores, we examined the internal consistency of each subscale within our sample of 4-year-olds. The internal consistencies were adequate for the care (Cronbach's $\alpha = .73$), authority ($\alpha = .68$), and sanctity ($\alpha = .72$) subscales, but lower for the fairness ($\alpha = .41$), loyalty ($\alpha = .43$), and odd behaviours ($\alpha = .57$) subscales. Therefore, analyses with the latter subscales should be interpreted with caution.

Moral Foundations Questionnaire

The Moral Foundations Questionnaire (MFQ; see Appendix A) was administered to one parent/guardian of each child participant. Other than being fluent in English, we

did not enforce inclusion or exclusion criteria for the parent/guardian, to better enable equal access to research participation and to minimize missing data. Those who participated at the lab completed this questionnaire while their child was completing the tasks described above; those who participated through daycares completed and returned it with their signed consent form. The MFQ (Graham et al., 2011) is a 32-item questionnaire measuring adults' concern for the five moral foundations of care, fairness, loyalty, authority, and sanctity. The 32 items are divided into two parts: moral relevance items, and moral judgment items. For moral relevance items, participants rated the extent to which different concerns factor into their judgments of what is right or wrong (e.g., whether or not someone suffered emotionally). They responded using a 6-point Likert scale ranging from 0 (not at all relevant) to 5 (extremely relevant). For the moral judgment items, participants decided how much they agreed or disagreed with particular actions, virtues, or statements (e.g., it can never be right to kill a human being). They responded using a 6-point Likert scale ranging from 0 (strongly disagree) to 5 (strongly agree). Across the two sections there were six questions for each of the five foundations. There were also two 'catch' items designed to indicate inattentive or insincere response patterns (e.g., it is better to do good than bad).

The MFQ yields scores that indicate concern for each foundation. For this study, we created an average score for each foundation, which was calculated as long as participants had answered 80% (or 5/6) of questions for that foundation. Total scores ranged from 0 to 5, with greater scores indicating greater concern for a particular foundation. In our sample of parents, the internal consistencies (α) were .49, .43, .73, .55, and .74 for the care, fairness, loyalty, authority, and sanctity foundations, respectively.

When Graham et al. (2011) examined the reliability and validity of the MFQ, they found internal consistencies ranging from .65 to.84. They also found that test-retest reliability (after 28-43 days) ranged from .68 to .84, while exploratory factor and confirmatory factor analyses indicated that a five-factor model was the best fit for a large sample (N = 34,476). Graham et al. (2011) also found evidence of external validity; the MFQ predicted relevant personality features and social group attitudes.

Demographics Questionnaire

The demographics questionnaire was also completed by one parent/guardian of each child participant and was completed at the same time as the MFQ. The survey assessed child gender identity, parent gender identity, gross family income, number of siblings, highest level of education obtained by the parent/guardian, religious orientation, political orientation, and parent/guardian cultural or ethnic background. The full survey and response options are shown in Appendix B.

Chapter 3: Children's Responses to the Moral Foundations Puppet Task

To properly examine the key tenets of Moral Foundations Theory (MFT), developmental research is required (see Chapter 1). Therefore, in the current chapter, we specifically assessed children's concern for moral foundations (care, fairness, loyalty, authority, sanctity) from a very early age (2-years-old). MFT makes developmental claims that have yet to be properly examined with descriptive, developmental research (Graham et al., 2011). It asserts that morality is both partly innate and influenced by the environment. In Haidt's (2012) version of innateness, innate does not mean present at birth; moral foundations can emerge at different ages, likely relating to development in other areas. However, it is still important to examine the moral foundations from an early age – it is key for determining at what stage each foundation develops, and to examine what each looks like when children have experienced less cultural influence. Furthermore, especially when it comes to loyalty, authority, and sanctity, there has been very little research conducted below the ages of 4 or 5. Therefore, we felt it was important to extend the existing literature to younger ages.

Conducting research with children as young as 2 years old meant that we first needed to develop a task that measured moral concerns in this age group. In our approach, we took inspiration from the tasks developed and administered by Hamlin and colleagues (Hamlin, 2013; Hamlin et al., 2007; Hamlin & Wynn, 2011). We developed puppet shows and examined children's subsequent reactions and puppet preferences. However, given that Hamlin et al. (2007) conducted their research with preverbal infants, we had slightly more freedom regarding the content of our measure. That is, while we limited language when possible, we did use some age-appropriate language (e.g., go to

bed, get down). Other differences included that we used videotaped rather than live puppet shows to increase standardization between participants. We developed two different puppet shows or scenarios for each moral foundation, each one depicting both a 'good' and a 'bad' puppet. Using only two scenarios did not cover the full content of each foundation, especially when language limitations are considered. However, many of these foundations were previously unexplored in children this young, and it was our hope that this research would stimulate further developmental research on moral foundations.

The creation of our scenarios was based both on concerns within each foundation, and on existing gaps in the literature. For the care foundation, there is relatively less research on children's perception of physical harm than children's perception of someone who hinders another from completing a goal (e.g., Hamlin et al., 2007). Furthermore, there were no studies that looked at harm to non-human animals. Therefore, our scenarios involved physical harm (and physical caring) to 'humans' (other puppets) and non-human animals (emotional harm would rely too much on language). For the fairness foundation, studies have examined toddler's reaction to unequal distributions of resources between two third parties (e.g., Geraci & Surian, 2011). However, there is little to no research examining toddler's perceptions of unequal or ungenerous sharing between an actor and a third party. Therefore, one of our scenarios depicted sharing (equally or unequally), and the other depicted distribution of resources to third parties (equally or unequally).

For the remaining foundations, there is no research on toddlers' concerns for relevant issues. Therefore, we simply developed scenarios that involved triggers for each foundation. For loyalty, both scenarios depicted betraying an ingroup, versus helping or joining an ingroup. Groups were teams based on puppet shirt color; this was felt to be

easier to understand than depictions of families, organizations, etc. For the authority foundation, both scenarios involved acts of simple disobedience (versus obedience); disrespect (e.g., name calling) was felt to rely too heavily on verbal comprehension. Finally, for the sanctity foundation, one scenario involved physical disgust and contamination. For the other, we wanted to depict the destruction (or preservation) of something that could be considered sacred to a small child, but avoid confounding this with care or harm. Therefore, we chose the destruction (vs. conservation) of a flower. For all scenarios and foundations, we used puppets with similar facial expressions to avoid preferences based on friendliness. We also used similar movements between the 'good' and 'bad' puppets (e.g., both jumping after hugging or hitting).

In summary, using a behavioural puppet task, we aimed to examine moral foundations in young children. During the puppet task, participants watched puppet shows depicting 'good' and 'bad' acts according to each moral foundation. They were then asked to point or name 1) the puppet they liked best (preference), and 2) which puppet they thought was bad (judgement). We recruited 2-year-olds to study MFT in toddlers, as well as to extend the existing moral development literature on loyalty, authority, and sanctity. We also recruited 4-year-olds to examine cross-sectional changes in moral concerns. Our hypotheses (and justifications) were as follows:

H1: We expected that 2-year-olds would be sensitive to our care and fairness scenarios. That is, they would prefer the 'good' puppet and judge the 'bad' puppet as bad at rates significantly higher than chance for each of the care and fairness scenarios.

Justification: There is an abundance of literature showing that from early infancy or toddlerhood, children are sensitive to concerns of care or fairness (e.g., Hamlin &

Wynn, 2011; Sommerville et al., 2013). Of note, given the dearth of existing literature for other foundations (loyalty, authority, sanctity), no predictions were made for 2-year-olds' responses to these foundations (although exploratory analyses were still conducted).

H2: We expected that 4-year-olds would be sensitive to our care, fairness, loyalty, and authority scenarios. That is, they would prefer the 'good' puppet and judge the 'bad' puppet as bad at rates significantly higher than chance for each of these foundations.

Justification: Similar to 2-year-olds, there is ample research showing that 4-year-olds are sensitive to care and fairness (e.g., Li et al., 2016; Smetana et al., 2012). While there is relatively less research examining their concern for authority and loyalty, there is still some evidence that 4-year-olds are sensitive to disloyalty and disobedience (Abrams et al., 2003; Smetana et al., 2014). We again made no predictions for 4-year-olds' responses to the sanctity foundation.

H3: We expected that 4-year-olds would show more concern than 2-year-olds for the loyalty and authority foundations. That is, for these two foundations, they would prefer the 'good' puppet and judge the 'bad' puppet as bad more frequently than 2-year-olds.

Justification: Although we made no specific predictions regarding 2-year-olds' concern for loyalty and authority, given the developmental context, we feel it was reasonable to expect that 4-year-olds would be more sensitive to these foundations. Both involve concepts (e.g., hierarchy, team membership) that children might not understand at very early stages in development. Furthermore, during the preschool period, issues of loyalty and authority are frequently taught and discussed (e.g., obedience, starting to play team sports). Of note, no predictions for age differences were made for the care, fairness,

and sanctity foundations, although for different reasons. For care and fairness, studies have shown that these remain strong concerns across the lifespan, especially in samples from Western cultures (Helwig et al., 2001; LoBue et al., 2011). Therefore, there was no reason to expect age differences, although we cannot specifically test 'no difference' using hypothesis testing. For sanctity, no predictions were made simply given the lack of research and prior knowledge in this area.

H4: Using summary scores created for each foundation on the MFPT (see below), we expected to see significant, positive correlations between care and fairness, loyalty and authority, loyalty and sanctity, and sanctity and authority.

Justification: Haidt (2012) proposed that the care and fairness foundations were related, as they both deal with concerns related to individuals. In contrast, he suggested that the loyalty, authority, and sanctity foundations deal with group-related concerns, and referred to them as 'binding' foundations. Indeed, Graham et al. (2011) found that care and fairness were correlated more strongly with each other than with the other three foundations, and the same applied to loyalty, authority, and sanctity. Therefore, it is reasonable to expect that moral foundations will show similar developmental trajectories within these categories (individualizing vs. binding). Indeed, care- and fairness-related behaviours have been shown to develop at similar ages (Schmidt & Sommerville, 2011; Sommerville et al., 2013).

Methods

The measure of interest for this chapter and analysis was the Moral Foundations Puppet Task, described in detail in Chapter 2. As a brief reminder, 102 two-year-olds and 76 four-year-olds completed this task. They were presented with two puppet shows

within each moral foundation (see Chapter 2, Table 3), each one depicting a 'good' and 'bad' puppet. They were then asked which puppet they liked best (preference), and which puppet did something bad (judgment), as well as two questions to check their comprehension. Children's responses were recorded as 'good puppet,' 'bad puppet', 'neither' or 'both', and 'unclear'. Of note, if children's responses were unclear, this was considered to be missing data. Furthermore, for the 'neither' or 'both' responses, our original intention had been to divide these responses equally amongst the 'good puppet' and 'bad puppet' responses. However, given the low frequency of such responses (2 or less for each question), these data were also treated as missing.

To examine correlations between foundations, we also created a summary variable for each foundation. The preference and judgment questions were first coded so that a score of 1 represented a more sensitive answer or more concern for the relevant foundation (i.e., 'good puppet' for the preference question, 'bad puppet' for the judgment question). Within each foundation, we then summed each of the four relevant variables (2 scenarios per foundation, 2 questions per scenario). This was only done if participants had data for all four variables. This created an ordinal level variable for each foundation, ranging from 0-4, where higher scores indicated greater concern for the relevant moral foundation. We examined the internal consistency for these summary variables using Cronbach's alpha (equivalent to Kuder-Richardson 20 with binary variables). As shown in Table 6, α ranged from .62 to .72, demonstrating adequate internal consistency for small scales of binary items.

If parents consented, children's responses to the MFPT scenarios were videotaped for reliability coding (i.e., 10 scenarios/videos per child). Reliability coding was

completed by a third party who was blind to which puppet was good or bad in each scenario. They coded the puppet selected by participants in response to each of the four questions (preference, judgment, and two manipulation checks) per scenario; their coding was then compared to the original coding and percent agreement was calculated. Of note, in addition to lack of parental consent, some scenarios were not videotaped due to equipment malfunction, lack of space in daycare testing settings, or child variables (e.g., hyperactivity, distractibility). In total, 1309 scenarios from 178 participants (74.21% of all scenarios administered) were videotaped. Given four questions to be coded per scenario, this produced 5236 responses to be coded. However, only 97.08% were codeable due to video angle, child moving out of the picture, etc. Therefore, 5083 individual responses were coded for interrater reliability. The overall percent agreement between the original and third-party coding was 95.75%, indicating excellent interrater reliability. Percent agreement per question and average percent agreement per scenario is shown in Appendix F. Given the high percent agreement, the original coding was kept for all analyses.

Table 6

Internal consistency for summary scores on the MFPT

Foundation	Alpha
Care	.67
Fairness	.72
Loyalty	.62
Authority	.65
Sanctity	.66

Data Analysis

Prior to examining our hypotheses, the first analysis conducted in this chapter was to examine participants' responses to manipulation checks. As a reminder, during the

MFPT, we asked participants two manipulation checks following each video to assess their comprehension of the videos. For example, for the care foundation, Scenario 1 (hitting versus hugging another puppet), we asked participants: “Which puppet hit? Which puppet hugged?”. Given that there were two checks, there were four possible ways in which participants could respond. They could answer both questions correctly, both questions incorrectly, the first question correctly and the second incorrectly, or the second question correctly and the first incorrectly. Therefore, if children were responding at chance level, we would expect to see each combination appear approximately 25% of the time. We conducted one-tailed binomial tests to see if the proportion of correct responses (answering both questions correctly, expected proportion = .25) versus incorrect responses (answering one or both questions incorrectly, expected proportion = .75) differed significantly from chance. These were conducted separately in each age group, considering probable differences in comprehension between 2- and 4-year-olds.

Hypotheses 1 through 3, and associated exploratory analyses, were tested together through a series of probability and non-parametric tests (given binary data). Firstly, we conducted two-tailed binomial probability tests for all participants (i.e., 2- and 4-year-olds combined) to determine if the puppet choices (good vs. bad puppet) of the entire sample differed significantly from chance (expected proportion = .50). An individual binomial test was conducted for each question (preference and judgment) of each scenario (2 scenarios per foundation), as it was possible that children’s responses differed depending on the scenario or question asked. Therefore, this step was essentially testing to see if there was a ‘main effect’ of scenario on children’s puppet preferences and judgments in the overall sample. For the next step, we examined the effect of age.

Specifically, we conducted two-tailed Fisher's exact tests to examine how age impacted puppet choice for each question of each scenario. Fisher's exact test was chosen over Chi-Square as it does not assume 5 counts per cell. Odds ratios were also calculated. Finally, our last step was, for questions and scenarios that showed age differences, to conduct separate binomial tests within each age group. These were, in effect, 'post-hoc' tests; if there was an age difference, these tests showed if each age group's puppet choices differed significantly from chance for that particular question or scenario (expected proportion = .50). Thus, altogether, for Hypotheses 1 and 2 (regarding which foundations 2- and 4-year-olds would be sensitive to), if there was no age difference, the first binomial tests (conducted in the overall sample) told us how 2- and 4-year-olds' responded to given scenario or question. If there was an age difference, we relied instead on the second binomial tests to tell us how they responded. Finally, Hypothesis 3 was tested solely via Fisher's exact test (examining the impact of age on puppet choices). Of note, given that we only made predictions for certain foundations, not all of these analyses tested hypotheses – however, this series of analyses was conducted on all foundations and scenarios for exploratory purposes and for the sake of completeness.

As shown below, 2-year-olds' responses to the manipulation checks were perhaps less than desirable (approximately 50% correct or less). Therefore, although this was not originally planned, we conducted more binomial tests solely with 2-year-olds, specifically those 2-year-olds that answered both manipulation checks correctly. That is, for a given scenario, we selected the 2-year-olds who answered the manipulation checks correctly for that specific scenario, and tested to see if their puppet choices differed significantly from chance for the preference and judgment questions (expected proportion

= .50). This analysis was purely exploratory and meant to help with interpretation of Hypotheses 1-3.

Finally, to test the final hypothesis regarding correlations between foundations, we conducted Spearman's correlations (and partial Spearman's correlations controlling for age in months) between each foundation. These correlations were conducted between the summary variables for each foundation described above. Spearman's correlations were used (i.e., rather than Pearson's) given the ordinal nature of these variables. Of note, although we only made predictions regarding correlations between some foundations, all possible inter-foundation correlations were examined for exploratory purposes and the sake of completeness.

All of the tests completed in this section involved a large number of tests, thereby greatly inflating the family-wise error rate. To control for this, we employed the Benjamini and Hochberg (1995) procedure, which controls for false discovery rate across multiple comparisons. It was conducted separately for each group or family of analyses that were run. This procedure is considered a good alternative to the Bonferroni correction, which has been criticized as being too conservative. It is a sequential procedure, that first rank-orders p-values for a family of tests. Subsequently, the corrected threshold for significance is calculated for each individual p-value. This threshold is based on the rank of the p-value, the total number of tests, and the false discovery rate selected (in comparison to the Bonferroni procedure, which takes into account only the desired alpha level and number of tests). In this paper, for ease of both reporting and interpretation, we have reported only one corrected significance level per group of tests (rather than the threshold for each individual p-value, even though the

procedure was applied to each p-value). This significance threshold is the largest threshold that was met; for example, if the largest p-value found to be significant was .02, and the significance threshold corresponding to said p-value was .025, it is this value (.025) that was reported as the corrected significance threshold (and anything over that threshold was not significant). In keeping with common standards in the literature, the false discovery rate selected for all analyses was .05. Overall, where corrected significance threshold is reported in this paper, it was always based on the Benjamini and Hochberg (1995) procedure.

Results

Manipulation Checks

Table 7 shows the percentage of 2- and 4-year-olds who answered both manipulation checks correctly for each scenario and moral foundation. As described

Table 7

Percentage of 2- and 4-year-olds that answered both manipulation checks correctly on the MFPT, by scenario and moral foundation

		% That answered both manipulation checks correctly	
Foundation	Scenario	2-year-olds	4-year-olds
Care	1	58.8	77.6
	2	49.0	89.5
Fairness	1	48.0	77.6
	2	38.2	77.6
Loyalty	1	36.3	59.2
	2	36.3	59.2
Authority	1	51.0	90.8
	2	43.1	88.2
Sanctity	1	57.8	94.7
	2	55.9	96.1

Note. Care Scenario 1 depicted hitting vs. hugging, Scenario 2 depicted kicking vs. petting. Fairness Scenario 1 depicted equal vs. unequal distribution of cookies, Scenario 2 depicted equal vs. unequal sticker sharing. Loyalty Scenario 1 depicted joining own vs. opposite team, Scenario 2 depicted helping own vs. opposite team. Authority Scenario 1 depicted going to bed vs. not going to bed, Scenario 2 depicted getting off table vs. not getting off table. Sanctity Scenario 1 depicted smelling vs. destroying flower, Scenario 2 depicted rolling vs. not rolling in feces.

in the previous section, we conducted one-tailed binomial tests to examine the proportion of correct responses (answering both questions correctly, expected proportion = .25) versus incorrect responses (answering one or both questions incorrectly, expected proportion = .75). All tests produced significant results (all *p*'s <.001), indicating that both age groups answered the manipulation checks correctly at a rate significantly higher than chance. However, as shown in Table 7, the rates of correct responding, especially for 2-year-olds, were not necessarily desirable.

Examining 2- and 4-year-olds' concern for moral foundations (H1-H3)

As described above, our first step in testing Hypotheses 1-3 was to conduct two-tailed binomial probability tests in our overall sample for each question (preference and judgment) of each moral foundation scenario (corrected significance threshold = .035). Results are shown in Table 8. Asterisks show significant p-values, which indicate that children chose one puppet over another at a rate significantly higher than chance. To summarize, the binomial tests indicated that our participants (2- and 4-year-olds combined) chose one puppet over the other in the expected direction (i.e., good puppet for the preference question, bad puppet for the judgment question) at rates significantly higher than chance for the majority of the care, fairness, authority, and sanctity scenarios and questions. No significant preferences were shown for the loyalty foundation.

Our second step in testing Hypotheses 1-3 (and especially Hypothesis 3) was to examine the impact of age on puppet choice. We conducted a two-tailed Fisher's exact test examining age group (2 vs. 4) and puppet choice (good vs. bad) for each question of each scenario (corrected significance threshold = .035). Descriptive data showing the percentage of 2- and 4-year-olds that chose the good and bad puppets for each question

Table 8

Results of binomial probability tests for the preference and judgment questions of each scenario in 2- and 4-year-olds combined on the MFPT

Foundation	Scenario	Question	N	% Good puppet	% Bad puppet	p
Care	1	Preference	174	55.7	44.3	.150
		Judgment	173	33.5	66.5	<.001*
	2	Preference	177	64.4	35.6	<.001*
		Judgment	176	27.3	72.7	<.001*
Fairness	1	Preference	178	65.2	34.8	<.001*
		Judgment	178	30.9	69.1	<.001*
	2	Preference	175	56.9	44.0	.130
		Judgment	170	37.1	62.9	.001*
Loyalty	1	Preference	171	48.0	52.0	.646
		Judgment	168	47.0	53.0	.488
	2	Preference	170	49.4	50.6	.939
		Judgment	165	44.8	55.2	.213
Authority	1	Preference	174	60.9	39.1	.005*
		Judgment	173	31.8	68.2	<.001*
	2	Preference	175	61.7	38.3	.002*
		Judgment	175	34.3	65.7	<.001*
Sanctity	1	Preference	176	61.9	38.9	.002*
		Judgment	173	30.6	69.4	<.001*
	2	Preference	176	60.2	39.8	.008*
		Judgment	175	34.9	65.1	<.001*

Note. * indicates significant p-values (after application of the Benjamini and Hochberg 1995 procedure). Care Scenario 1 depicted hitting vs. hugging, Scenario 2 depicted kicking vs. petting. Fairness Scenario 1 depicted equal vs. unequal distribution of cookies, Scenario 2 depicted equal vs. unequal sticker sharing. Loyalty Scenario 1 depicted joining own vs. opposite team, Scenario 2 depicted helping own vs. opposite team. Authority Scenario 1 depicted going to bed vs. not going to bed, Scenario 2 depicted getting off table vs. not getting off table. Sanctity Scenario 1 depicted smelling vs. destroying flower, Scenario 2 depicted rolling vs. not rolling in feces.

and scenario are shown in Figure 1, while results of each Fisher's exact test are shown in Table 9. Of note, odds ratios show the probability of 4-year-olds picking the good puppet relative to 2-year-olds. Ratios greater than 1 indicate that 4-year-olds were more likely to pick the good puppet, while ratios less than 1 indicate that they were less likely to pick the good puppet.

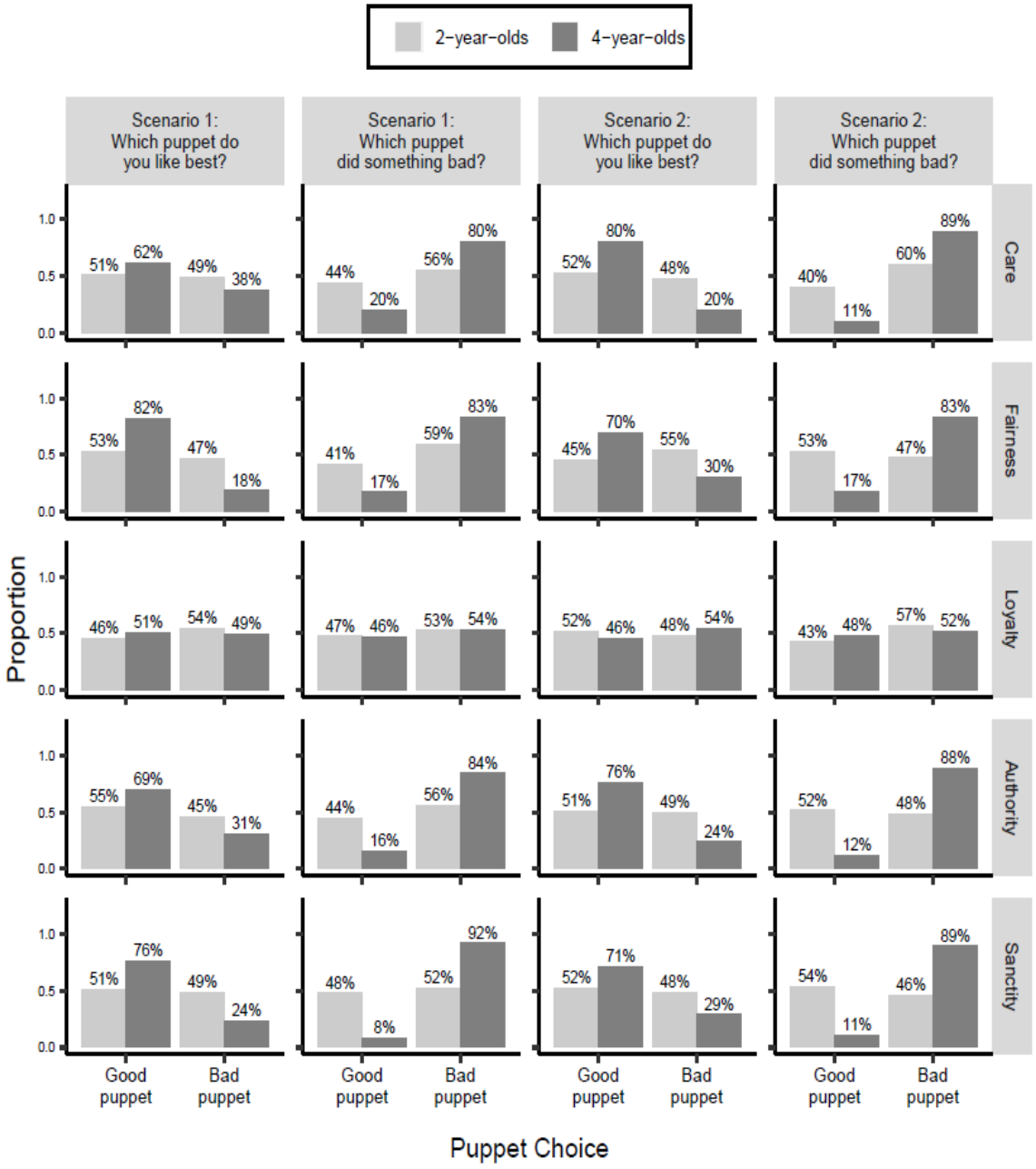


Figure 1. Percentage of 2- and 4-year-olds that selected the good and bad puppets for the preference and judgment questions of each scenario and moral foundation on the MFPT. Care Scenario 1 depicted hitting vs. hugging, Scenario 2 depicted kicking vs. petting. Fairness Scenario 1 depicted equal vs. unequal distribution of cookies, Scenario 2 depicted equal vs. unequal sticker sharing. Loyalty Scenario 1 depicted joining own vs. opposite team, Scenario 2 depicted helping own vs. opposite team. Authority Scenario 1 depicted going to bed vs. not going to bed, Scenario 2 depicted getting off table vs. not getting off table. Sanctity Scenario 1 depicted smelling vs. destroying flower, Scenario 2 depicted rolling vs. not rolling in feces. *N*'s for 2-year-olds ranged from 94-102, and for 4-year-olds ranged from 71-76).

Table 9

Examining the relationship between age group (2- vs. 4-year-olds) and puppet choice on the MFPT using Fisher's exact test

Foundation	Scenario	Question	p	OR	95% CI OR
Care	1	Preference	.169	1.56	[0.85, 2.86]
		Judgment	.001*	0.31	[0.15, 0.62]
	2	Preference	<.001*	3.68	[1.85, 7.32]
		Judgment	<.001*	0.18	[0.08, 0.41]
Fairness	1	Preference	<.001*	3.94	[1.96, 7.91]
		Judgment	.001*	0.29	[0.14, 0.60]
	2	Preference	.002*	2.77	[1.47, 5.19]
		Judgment	<.001*	0.19	[0.09, 0.39]
Loyalty	1	Preference	.542	1.21	[0.66, 2.22]
		Judgment	1.00	0.96	[0.52, 1.78]
	2	Preference	.445	0.78	[0.43, 1.43]
		Judgment	.530	1.24	[0.67, 2.31]
Authority	1	Preference	.060	1.88	[1.00, 3.54]
		Judgment	<.001*	0.24	[0.11, 0.49]
	2	Preference	.001*	3.16	[1.63, 6.11]
		Judgment	<.001*	0.13	[0.06, 0.28]
Sanctity	1	Preference	.001*	3.10	[1.60, 5.98]
		Judgment	<.001*	0.09	[0.04, 0.24]
	2	Preference	.013*	2.27	[1.20, 4.26]
		Judgment	<.001*	0.10	[0.04, 0.23]

Note. * indicates significant p-values (after application of the Benjamini and Hochberg 1995 procedure). Care Scenario 1 depicted hitting vs. hugging, Scenario 2 depicted kicking vs. petting. Fairness Scenario 1 depicted equal vs. unequal distribution of cookies, Scenario 2 depicted equal vs. unequal sticker sharing. Loyalty Scenario 1 depicted joining own vs. opposite team, Scenario 2 depicted helping own vs. opposite team. Authority Scenario 1 depicted going to bed vs. not going to bed, Scenario 2 depicted getting off table vs. not getting off table. Sanctity Scenario 1 depicted smelling vs. destroying flower, Scenario 2 depicted rolling vs. not rolling in feces.

Finally, our last step in testing Hypotheses 1-3 was to conduct follow-up or 'post-hoc' binomial tests for the questions and scenarios that indicated significant age differences. That is, we conducted separate binomial tests in each age group for the questions and scenarios that showed a significant age difference (corrected significance threshold = .025). As the results were very consistent, they will just be described here, rather than shown in a table. For every Fisher's exact test that revealed a significant relationship between age and puppet choice (i.e., all except for the loyalty scenarios), the follow-up binomial tests showed the same pattern. Two-year-olds' puppet choices did not

differ significantly from chance (*p*'s ranging from .057-1.00), while 4-year-olds showed significant puppet preferences in the expected direction (all *p*'s .001 or less) – i.e., they chose the good puppet for the preference question, and the bad puppet for the judgment question..

Re-visiting Hypotheses 1 through 3

To summarize the above results and how they relate to our hypotheses, in Hypothesis 1, we predicted that 2-year-olds would prefer the 'good' puppet and judge the 'bad' puppet as bad at rates significantly higher than chance for each of the care and fairness scenarios. However, 2-year-olds did not show significant puppet preferences in response to the care or fairness scenarios. They also did not show significant puppet preferences in response to loyalty, authority, or sanctity foundations (no predictions were made for 2-year-olds on these foundations). The one exception was the preference question of authority Scenario 1 (goes to bed vs. does not go to bed). For this question, the overall sample picked the good puppet at a rate significantly higher than chance, and there was no significant effect of age.

In Hypothesis 2, we predicted that 4-year-olds would prefer the 'good' puppet and judge the 'bad' puppet as bad at rates significantly higher than chance for the care, fairness, loyalty, and authority foundations (no predictions were made for sanctity). We found that 4-year-olds' puppet choices for both the preference and judgment questions of each scenario differed significantly (in the expected direction) from chance for the care, fairness, and authority, with the exception of the preference question for care Scenario 1 (hit vs. hug). This was also true for the sanctity foundation (this was exploratory). For the loyalty foundation, 4-year-olds did not show a significant preference for either question

of Scenario 1 (joins own vs. opposite team) or Scenario 2 (helps own vs. opposite team).

Overall, 2-year-olds did not appear to be sensitive to any of the moral foundations, while 4-year-olds were sensitive to the care, fairness, authority, and sanctity foundations, but not loyalty.

Exploratory analysis: Examining 2-year-olds who answered manipulation checks correctly

As described above, 2-year-olds' responses to our manipulation checks, although above chance level, were not optimal. Therefore, for a given scenario, we selected the 2-year-olds who answered the manipulation checks correctly for that specific scenario, and tested to see if their puppet choices differed significantly from chance for the preference and judgment questions (using two-tailed binomial probability tests for each question/scenario, corrected significance threshold = .025). This analysis was exploratory and meant to aid with interpretation of Hypotheses 1-3. Significant results are indicated with an asterisks in Table 10. Interestingly, 2-year-olds who answered manipulation checks correctly picked the bad puppet over the good puppet at rates significantly higher than chance for seven out of ten of the judgment questions (exceptions were the care 1, loyalty 1, and sanctity 2 scenarios). In contrast, they only picked the good puppet over the bad puppet for three of the preference questions (the fairness 1, loyalty 2, and authority 2 scenarios). Therefore, even when examining only the 2-year-olds that had good comprehension of the videos, the moral foundations did not seem to strongly impact their puppet preferences, although they were better able to say which puppet had done something bad.

Table 10

Results of binomial probability tests for the preference and judgment questions of each scenario on the MFPT, using only 2-year-olds that answered both manipulation checks correctly

Foundation	Scenario	Question	N	% Good puppet	% Bad puppet	p	Consistent with previous tests?
Care	1	Preference	59	55.9	44.1	.435	Consistent
		Judgment	59	35.6	64.4	.036	Consistent
	2	Preference	50	58.0	42.0	.322	Consistent
		Judgment	50	22.0	78.0	<.001*	Inconsistent
Fairness	1	Preference	49	69.4	30.6	.009*	Inconsistent
		Judgment	49	24.5	75.5	<.001*	Inconsistent
	2	Preference	39	64.1	35.9	.108	Consistent
		Judgment	36	27.8	72.2	.011*	Inconsistent
Loyalty	1	Preference	37	62.2	37.8	.188	Consistent
		Judgment	37	40.5	59.5	.324	Consistent
	2	Preference	37	70.3	29.7	.020*	Inconsistent
		Judgment	37	24.3	75.7	.003*	Inconsistent
Authority	1	Preference	52	61.5	38.5	.126	Inconsistent
		Judgment	51	29.4	70.6	.005*	Inconsistent
	2	Preference	44	70.5	29.5	.010*	Inconsistent
		Judgment	44	27.3	72.7	.004*	Inconsistent
Sanctity	1	Preference	59	64.4	35.6	.036	Consistent
		Judgment	58	31.0	69.0	.005*	Inconsistent
	2	Preference	57	63.2	36.8	.063	Consistent
		Judgment	56	37.5	62.5	.081	Consistent

Note. *indicates significant *p*-values (after application of the Benjamini and Hochberg 1995 procedure). Care Scenario 1 depicted hitting vs. hugging, Scenario 2 depicted kicking vs. petting. Fairness Scenario 1 depicted equal vs. unequal distribution of cookies, Scenario 2 depicted equal vs. unequal sticker sharing. Loyalty Scenario 1 depicted joining own vs. opposite team, Scenario 2 depicted helping own vs. opposite team. Authority Scenario 1 depicted going to bed vs. not going to bed, Scenario 2 depicted getting off table vs. not getting off table. Sanctity Scenario 1 depicted smelling vs. destroying flower, Scenario 2 depicted rolling vs. not rolling in feces.

Correlations between moral foundations (H4)

As previously described, we created a summary variable (ordinal level, ranging from 0-4) for each moral foundation on the MFPT. Higher scores on this variable indicated greater concern for the relevant moral foundation; descriptive statistics are shown in Table 11. In Hypothesis 4, we predicted significant, positive correlations between the care and fairness, loyalty and authority, loyalty and sanctity, and sanctity and

authority foundations. To test this hypothesis, we conducted Spearman correlations to examine the relationship between foundations. Table 12 shows bivariate Spearman correlations between foundations (corrected significance threshold = .02) with partial Spearman correlations (controlling for age in months) in parentheses (corrected significance threshold = .005). On the MFPT, the sanctity foundation showed a significant positive correlation with the care, fairness, and authority foundations. The authority foundation also showed a significant correlation with the fairness foundation. However, when controlling for age in months, these correlations shrunk in size by approximately 33-50% and were no longer significant.

Table 11

Descriptive statistics for moral foundation summary scores on the MFPT

Foundation	Mean (*SD*)	Median	Mode	IQR (Q1-Q3)	Min	Max
Care	2.60 (1.34)	3.00	4.00	2.00-4.00	0.00	4.00
Fairness	2.54 (1.39)	2.00	4.00	2.00-4.00	0.00	4.00
Loyalty	2.07 (1.35)	2.00	2.00	1.00-3.00	0.00	4.00
Authority	2.52 (1.32)	2.00	4.00	2.00-4.00	0.00	4.00
Sanctity	2.56 (1.34)	3.00	4.00	2.00-4.00	0.00	4.00

Note. IQR = interquartile range, Q1 = 25th percentile, Q3 = 75th percentile.

Table 12

Bivariate Spearman correlations between moral foundation summary scores on the MFPT [with partial correlations controlling for age in months][1]

	Care	Fairness	Loyalty	Authority	Sanctity
Care	1				
Fairness	.13 [-.01]	1			
Loyalty	.15 [.17]	-.02 [-.01]	1		
Authority	.06 [-.10]	.21*[.08]	.03 [.04]	1	
Sanctity	.26*[.13]	.33***[.21]	-.01 [-.00]	.33*** [.21]	1

Note. ***p<.001, **p<.01, *p<.02. Following application of the Benjamini and Hochberg (1995) procedure, the corrected significance threshold for the bivariate correlations was .02, and for the partial correlations was .005.

[1]The analyses run in Table 12 were also conducted using only the children who passed the relevant manipulation checks, to reduce error and noise. Of note, this led to a significant reduction in the sample size for each correlation. No new relationships were found; those that were found were the same as in Table 12, and many shown in Table 12 become non-significant (likely due to the reduction in N).

Discussion

The purpose of this chapter was to examine moral foundations (care, fairness, loyalty, authority, and sanctity) in toddlers and preschoolers. We developed a measure to assess moral foundations in 2-year-olds; this allowed us to investigate Moral Foundation Theory (MFT) from an early age and extend existing research on moral development to concerns of loyalty, authority, and sanctity. We also recruited 4-year-olds to examine cross-sectional age differences in moral foundations. Four-year-olds showed concern, through significant puppet preferences, for the care, fairness, authority, and sanctity foundations, although were unexpectedly not sensitive to our loyalty scenarios. In contrast, 2-year-olds did not show consistent concern for any of the scenarios presented. This was unexpected for the care and fairness scenarios, and also meant that we saw significant age differences (with 4-year-olds showing greater concern) for all foundations but loyalty. Finally, the pattern of correlations between foundations was unexpected; sanctity was positively correlated with care, fairness, and authority, while fairness was also positively correlated with authority. That said, these correlations were non-significant when controlling for age in months.

Two-year-olds

In interpreting the results of this chapter's analysis, we felt it would be sensible to start with the most unexpected – that 2-year-olds were not sensitive to any of our scenarios depicting care and fairness violations. As described in detail in Chapter 1, there is extensive research showing that children are sensitive to care and fairness from infancy or early toddlerhood. Therefore, proper explanation of our results necessitated re-visiting this literature for a thorough comparison of study methods, styles, and interpretations.

Starting with the care foundation, both scenarios in the current study depicted acts of physical harm versus physical caring (of humans or non-human animals). In contrast, the vast majority of previous research examining care in infants or toddlers has presented participants with scenarios that involved helping or hindering the completion of a goal (e.g., going up a hill, opening a box, picking up a toy; Hamlin et al., 2011; Hamlin, 2013; Hamlin et al., 2007; Hamlin & Wynn, 2011; Hepach et al., 2012, 2016, 2017). Research showing that children are sensitive to physical harm (e.g., hitting, destroying belongings) has not been conducted younger than age 3 (Engelmann et al., 2016; Jambon & Smetana, 2014; Vaish et al., 2010, 2011). The one exception would be Smetana et al. (2012), who did find that, as young as 2.5 years old, children believed acts of physical and emotional harm were deserving of punishment. However, they also found that older children (i.e., 4-year-olds) were more likely to say that such acts were universally wrong.

In general, although there are many studies suggesting that young children are sensitive to issues of caring and prosocial behaviour, upon further examination, very few look at acts of physical harm, and even fewer examine physical harm younger than age 3. Therefore, while 2-year-olds may be sensitive to helping and hindering the completion of goals, as shown in previous studies, our results suggest that they are not sensitive to physical violence or physical caring. One possible interpretation of this difference could be that in previous studies, infants and toddlers simply wanted to see the action or goal completed and did not care about the intended harm of an act. That is, they are sensitive to goal completion, rather than acts of helping vs. harming. However, some studies (e.g., Hepach et al., 2017) did control for this possibility. Furthermore, we think it is important to reiterate that moral foundations are believed to encompass a wide variety of actions

and beliefs (Haidt, 2012), and become more refined over time. Therefore, it is simply possible that for toddlers, physical harm is not yet strongly linked to the care foundation; this connection may develop as children develop more understanding and have more exposure to teachings, culture, and more.

As with the care foundation, a re-examination of the literature highlighted some key methodological differences that could explain 2-year-olds' apparent lack of concern for our fairness scenarios. In the current study, our fairness scenarios depicted a) puppets distributing resources equally or unequally between two third parties, and b) puppets sharing equally or unequally between themselves and another puppet. The former scenario, children's evaluations of distributive fairness, has frequently been studied in toddlers and infants. However, the measures used to assess opinions of these situations in children younger than 3 have been almost exclusively looking time measures (DesChamps et al., 2016; Geraci & Surian, 2011; Schmidt & Sommerville, 2011; Sloane et al., 2012; Sommerville et al., 2013). Therefore, it is possible that toddlers care about distributive fairness, but this only appears in more 'unconscious' measures (i.e., looking time). Behavioural or preference measures may add more noise and error to the data or be more strongly impacted by other influencing and competing factors (e.g., animal preference, impulsivity or inattention, poor comprehension). It is also worth noting that the reliability of looking time research has been questioned. Oakes (2017) highlighted that looking time studies are often done using small sample sizes with low statistical power, which can cause both false positive and false negative results. Combined with the bias to publish significant results, this means that existing research may overestimate the strength or existence of certain phenomena (e.g., Erdmann et al., 2018).

While some studies have used measures other than looking time to assess children's concern for distributive fairness, these studies have not been conducted in children younger than 3 (e.g., Baumard et al., 2012; Melis et al., 2013; Olson & Spelke, 2008; Rakoczy et al., 2016; Warneken et al., 2011; Wittig et al., 2013). Likewise, children's opinions of how others should share (between themselves and a third party, rather than two third parties) has not been examined younger than age 3 (Fehr et al., 2008; Smith et al., 2013). Therefore, it is difficult to compare our 2-year-olds' results with such research. Of note, previous research has examined toddlers' (as young as 18 months old) own distributive and sharing behaviour, and has generally found that they make equal distributions but share selfishly (Chernyak et al., 2016; Li et al., 2016; Melis et al., 2013; Ulber et al., 2015; Warneken et al., 2011). However, as research has often shown, there can be differences in how children act and how they judge others (e.g., Smith et al., 2013). Furthermore, given that morality inherently involves judgment of others, research on prosocial behaviour does not easily extend to or compare with the current study. In general, there are both methodological and age differences that could explain the difference in our results and results of previous studies. This highlights the necessity of thorough research across all ages that employs large samples and a variety of different methods.

When interpreting the lack of concern for care and fairness shown by our sample of 2-year-olds, it is worth considering lack of comprehension as a possible explanation. This is especially true seeing that 2-year-olds in our study were not sensitive to any of the foundations – care, fairness, loyalty, authority, or sanctity. Of note, with a few exceptions (e.g., Hepach et al., 2016, 2017), most research in 2-year-olds has been conducted using

real actors or stimuli, while our study used videos. In toddlers, there is a phenomenon referred to as the video deficit (Anderson & Pempek, 2005). As late as 30 months old, studies have shown that toddlers learn less from videos than they do from live action or real life (Hayne et al., 2003). This has been found for imitation of actions and words, as well as object retrieval (Barr & Hayne, 1999; Linebarger & Walker, 2016; Troseth & DeLoache, 1998). However, the video deficit does not mean toddlers retain no information from video, and it is generally small or will disappear, depending on task, by 24 months old (Krcmar, 2010; Suddendorf, 2003). Furthermore, some factors have been shown to reduce the deficit, including repetition, looking at the monitor through a window, and social meaningfulness (e.g., having interacted with the people in the video; Krcmar, 2010; Troseth & DeLoache, 1998), and many of these factors were employed in the current study. Therefore, steps were taken to minimize video deficit, and our study also varied greatly from video deficit research (e.g., we did not ask children to imitate actions or retrieve objects). However, considering variation in children's development, it is possible that deficits in toddlers' ability to process videos impacted the results that were found.

In anticipation that children's comprehension of the moral foundation videos could vary, we did employ manipulation checks to examine children's comprehension. As shown in Table 8, 2-year-olds answered these checks correctly at rates significantly higher than chance (36-58% of the time), although certainly less frequently than 4-year-olds (who were correct 59-96% of the time). Furthermore, even when we examined solely those 2-year-olds who answered the manipulation checks correctly, they did not show the same overwhelming response or pattern to the preference and judgment

questions that the 4-year-olds showed for most foundations. Indeed, they seemed to be most sensitive to the judgment questions, with only a few significant puppet preferences. Interestingly, this represents something of a judgment-behaviour, or judgment-action, gap. Judgment-actions gaps are well-known moral phenomena (Stephens, 2018; Williams & Gantt, 2012). There exist numerous theories to explain them, which will not be described here – rather, it is sufficient to say that there is commonly a gap between moral opinions and behaviour across all ages and including children (e.g., Smith et al., 2013). Overall, we think it is reasonable to say, based on our analysis of 2-year-olds that understood the videos, that comprehension was not the only factor explaining why 2-year-olds did not show consistent, significant preferences and judgments to moral foundations. Numerous other factors – such as the methodological differences described above, judgment-action gaps, or even simply less concern for the presented foundations – were likely also at play.

Four-year-olds

In contrast to 2-year-olds, 4-year-olds showed consistent concern for all foundations except for loyalty, as shown by their responses to both the preference (preferring the good puppet) and judgment (judging the bad puppet) questions. Regarding the care foundation, it was expected that 4-year-olds would be sensitive to issues of care and harm. Previous research has shown that as young as 3 years old, children will protest physical harm (Vaish et al., 2011). Four-year-olds believe that both physical and emotional harm is wrong regardless of rules or authority, and preschool children are even less forgiving of necessary harm relative to older children (Jambon & Smetana, 2014; Smetana et al., 2012). Therefore, the current study is consistent with existing research

demonstrating that 4-year-olds are sensitive to acts of harm. To our knowledge, this is the first study showing that they are also sensitive to harm directed towards non-human animals. Of note, there was a small anomaly in our results; care Scenario 1 (hit vs. hug another puppet) did not impact 4-year-olds' puppet preference, only their judgment. However, given the amount of research showing that 4-year-olds are sensitive to harm, as well as significant preferences or judgments for 3 out of 4 of the care foundation questions, we believe that little attention should be given to this anomaly; it could simply be sampling variation or measurement error. Instead, we think that the overall pattern in the current study, and in past research, shows that 4-year-olds are indeed sensitive to issues of care and harm.

Like care, it was expected that 4-year-olds would be sensitive to our scenarios depicting distributive fairness and sharing. From an early age, children will distribute resources fairly between third parties (Chernyak et al., 2016; Li et al., 2016; Ulber et al., 2015). Furthermore, as young as 3 years old, children show a broad range of fairness behaviours, will protest unfair distributions, and the fairness of others impacts their prosocial behaviour (Baumard et al., 2012; Melis et al., 2013; Olson & Spelke, 2008; Rakoczy et al., 2016; Warneken et al., 2011; Wittig et al., 2013). Although children do not share fairly themselves until at least the age of 7, they believe that others should (Smith et al., 2013). Overall, our findings that 4-year-olds were sensitive to distributive fairness and sharing, in both preferences and judgments of others, is consistent with previous literature. Like care, it also provides a small addition to the existing body of literature; although previous research had shown that children believed others should share fairly, how the sharing of others impacts children's moral judgments and personal

preferences had not been studied. Therefore, our study further adds that 4-year-old children prefer actors who share fairly and believe that selfish sharing is morally bad.

Although the previous literature on children's concern for authority is much less extensive than the literature on care or fairness, there was still enough to suggest that 4-year-olds would be sensitive to acts of disobedience or disrespect (Martín-Antón et al., 2016; Smetana et al., 2014; Young & Avdzej, 1979). Therefore, 4-year-old's responses to authority scenarios in the current study were both expected and consistent with previous research. Furthermore, it fits with the developmental context. The parent/guardians of our sample were predominantly white, liberal, and wealthy, a population that typically puts less emphasis on the importance of authority (Graham et al., 2011). However, during the preschool period especially, even politically liberal parents frequently teach rules, ask for obedience, and engage in discipline. Therefore, it is understandable that their children would be sensitive to disobedience in others. Although we had made no specific predictions about 4-year-olds' concern for sanctity given the scarcity of literature, a similar interpretation could be made for the results that we found. That is, although the parent/guardians in our study were those that typically value sanctity concerns very little, toileting concerns, cleanliness, and what to protect versus destroy are principles heavily instilled in the preschool period. Therefore, it makes sense that the 4-year-olds in our sample were consistently sensitive to issues of sanctity and disgust.

Unlike the other foundations, 4-year-olds' responses to loyalty were unexpected. In the current study, 4-year-olds' preferences and judgments of loyal versus disloyal puppets were not significantly different from chance. Therefore, we again re-visited the literature to search for potential explanations, methodological differences, and

explanations. As outlined in Chapter 1, in addition to preferring ingroup members over outgroup members, 4- and 5-year-olds view acts of disloyalty as wrong (e.g., Abrams et al., 2003, 2008; Atkin & Gummerum, 2012). However, studies examining children's perceptions of loyalty have typically done so by assigning participants to groups, or using existing groups (e.g., family), and then examining their opinions of real or imagined ingroup members (Abrams et al., 2003, 2008; Enesco et al., 2011; Lu & Chang, 2016). To our knowledge, only one study has examined preschool children's perceptions of third party (i.e., where they were not part of a team) disloyalty. However, the loyalty violations used were confounded with violations of care (Rhodes & Chalik, 2013). Therefore, it is possible that that it is harder for young children to understand or be sensitive to loyalty violations when they are not a member of the team themselves. Similarly, for third-party scenarios, the group memberships may need to be stronger than simple colour teams (e.g., friendship, family). It is also worth noting that our depictions of loyalty were actually kind; that is, puppets who acted disloyally did so by helping the opposite team, rather than harming their own team (e.g., hurting their chances of winning). This represents another divergence from previous studies. Overall, our results suggest that future research examining children's opinions of different acts disloyalty in ingroup members and third parties is warranted.

Relationships between foundations

After examining 2- and 4-year-olds reactions to each individual foundation, the final goal of this chapter was to look at correlations between foundations (in the entire sample). Using a summary variable for each foundation, we found that sanctity was positively correlated with care, fairness, and authority. Fairness was also positively

correlated with authority; all other correlations were non-significant. This pattern of correlations was not what we expected. Based on research and adults, we had expected to see strong, positive correlations between the two individualizing foundations (care and fairness), and the three binding foundations (loyalty, authority, and sanctity). However, given what children were and were not sensitive to, this result seems appropriate. That is, 2-year-old's responses were no different from chance, and 4-year-olds showed concern for issues their parents likely value less given our wealthy, liberal sample (i.e., authority, sanctity). Therefore, it seems unlikely that we would have found the same correlations previously established in adults. The correlations were likely also influenced by measurement error, given the lower internal reliability values for the MFPT summary scores (see Table 6).

While unexpected, the pattern of correlations found was interesting, nonetheless. We found it particularly interesting that the sanctity foundation was related to all other foundations, with the exception of loyalty. This further highlights the difference in children's reactions to our loyalty scenarios relative to the other foundations. Moreover, it suggests that concern for sanctity may be important for the development of other foundations; it is possible that greater understanding of and concern for what is pure or sacred facilitates moral understanding and judgment in other domains as well. The authority foundation showed the next highest amount of correlations with other foundations, and a similar interpretation could be employed – that children who are more sensitive to rules and disobedience are more likely to develop or learn concern for other moral violations. Of note, the correlations we found decreased and became non-significant when controlling for age in months. However, it does not mean that they are

less important – it simply means that age or development influences the relationships that were found. Furthermore, these analyses were purely correlational and certainly not predictive. Therefore, future longitudinal research to examine how early moral concerns predict later ones, and to test our hypotheses (e.g., that concern for sanctity may facilitate concern for other areas), would be interesting and warranted.

Limitations and future directions

Some limitations of the methods and analyses presented in this chapter have been addressed in previous paragraphs, most notably those discussing the results found for 2-year-olds (e.g., it was difficult to determine if 2-year-old's responses were due to a lack of concern for the issues presented , comprehension, etc.). However, there are a few more to present here. The Moral Foundations Puppet Task (MFPT) was developed to be appropriate for use with 2-year-olds. However, this meant that we were limited in the scenarios and concerns we could show; moral foundations are broad and encompass more complexity and diversity than two simple scenarios (Haidt, 2012). Furthermore, the dependent variables of interest were dichotomous (which puppet was preferred or judged to be bad). Therefore, while we interpreted the results of the MFPT to indicate *if* children were sensitive to the moral issues presented, it does not indicate *how* sensitive children were, or degrees of concern. Although we did create summary variables for correlational analyses, these variables were still based on dichotomous responses, and were ordinal level with 5 possible values (0-4). In general, this meant that we could not effectively compare foundations; for example, we know 4-year-olds were sensitive to both care and authority, but we do not know if they show more concern for one or another. Thus, designing tasks to assess degree of concern for moral issues in toddlers and preschoolers

would be an important contribution to the literature.

In addition to resolving methodological issues, there are several other interesting avenues of future research. It was often interesting to observe, anecdotally, how children's outward reactions corresponded with their preferences and judgments. A subset of participants laughed at the 'bad' actions; most of these participants still judged the bad puppet, but some (not all) also preferred the bad puppet. There were other participants (largely 2-year-olds) who showed strong emotional reactions (e.g., being sad or angry over the puppet hitting), which did not appear to be reflected in their puppet choices. The current study was not designed to analyze emotional reactions; video recordings prioritized children's pointing over their faces, and many video recordings were missing due to lack of consent or other reasons. However, in future research, it would be interesting to examine children's emotional reactions to moral violations, and how that corresponds with their preferences or moral judgments, as well as other potential influencing factors (e.g., temperament, self-regulation, etc.). Overall, as highlighted a few times during this discussion, we believe our study demonstrates the importance of continued research on children's moral development, across time and ages, and particularly using a variety of methods and scenarios to study all five moral foundations.

The analyses and results presented in the current chapter represent a preliminary investigation into the moral concerns of 2- and 4-year-olds from a non-prescriptive, non-judgmental point of view. We made no decisions regarding what types of moral preferences or judgments were 'better;' instead we simply examined how children responded to the moral foundations proposed by Haidt (2012). We found that 2-year-

olds' responses did not differ significantly from chance, which particularly highlighted that they might not be as consistently sensitive to care and fairness as previously thought. Our sample of 4-year-olds were sensitive to issues of care, fairness, authority, and sanctity, while their responses to loyalty scenarios did not differ significantly from chance. The results we found were largely interpreted individually; the meaning of these analyses and those in subsequent chapters for Moral Foundations Theory (MFT) will be discussed in detail in Chapter 6.

Chapter 4: Measuring Concern for Moral Foundations with a Verbal Questionnaire

The Moral Foundations Puppet Task (MFPT) represented an initial investigation into the moral concerns of 2- and 4-year-old children. However, the MFPT is limited in what it can assess. The dichotomous responses tell us whether or not children were sensitive to the moral foundations, but do not indicate how sensitive, or degrees of concern. Furthermore, the MFPT was designed to rely very little on language to be appropriate for use with 2-year-olds, but this also limited the scenarios and moral concerns that could be displayed. Similarly, the MFPT depicts only two scenarios per foundation, some of which were very similar – however, each moral foundation encompasses a larger range of concerns. Therefore, in the current chapter, we sought to broaden our investigation of young children's concern for the five moral foundations (care, fairness, loyalty, authority, and sanctity). We focused exclusively on 4-year-olds, as they can participate in more sophisticated measures of moral development.

To broaden our assessment of moral foundations in 4-year-olds, we required a longer, and likely more verbal, measure – similar to previously developed adult measures. The Moral Foundations Questionnaire (MFQ; Graham et al., 2011) is both verbal and measures degree of concern for moral foundations. However, it is perhaps not the best model for children, as it requires abstract moral judgments. Furthermore, Clifford et al. (2015) outlined some limitations of the MFQ. In addition to using abstract judgments, rather than concrete scenarios, asking people to rate moral relevance may not assess how they actually make moral judgments. Frimer et al. (2013) also noted that the MFQ may focus too much on issues that divide political parties (e.g., chastity), rather than on the full moral spectrum. To address these limitations, Clifford et al. (2015)

developed a series of vignettes to assess adults' concern for the five moral foundations, in addition to social norm (i.e., non-moral) violations. They sought to develop stimuli that would measure participants' judgments of concrete behaviours and be suitable for use in behavioural and neuroimaging paradigms. They controlled the length and readability of all vignettes and began each vignette with "You see…," to encourage participant visualization of the behaviour.

When developing their vignettes, Clifford et al. (2015) covered particular concerns or behaviours with each moral foundation. For the care foundation, they focused on behaviours that depicted emotional harm to a human, physical harm to a human, and physical harm to a non-human animal. For the fairness foundation, they focused on instances of cheating or free riding. The loyalty vignettes each involved individuals putting themselves ahead of their ingroup (family, country, sports team, company, or school). All behaviours in loyalty vignettes had to publicly threaten the reputation of the ingroup, and there had to be a clear outgroup in competition with the ingroup. For the authority foundation, vignettes were comprised of disobedience or disrespect towards a traditional authority figure (e.g., boss, parent) or an institution or symbol of authority (e.g., police). All vignettes for the sanctity foundation involved elements of physical disgust, including sexual deviance, degrading behaviours (e.g., drunkenly making out with strangers on a bus), and contamination concerns (e.g., using a stranger's toothbrush). Finally, Clifford et al. (2015) also developed vignettes depicting social norm violations that were meant to be unusual but not morally wrong (e.g., lifting weights in business clothes). These vignettes were meant to serve as controls for neuroimaging studies, but also prevent participants from expecting something morally wrong for every vignette

(Clifford et al., 2015). Across all foundations and social norm vignettes, Clifford et al. (2015) tried to avoid cross-contamination (e.g., care/harm behaviours that involve social hierarchy) and topics that could invoke political rather than moral concerns (e.g., race or gender).

To examine the validity of their vignettes, Clifford et al. (2015) administered them to large adult samples across two studies. After reading each vignette, participants were asked a) how bad the action in the vignette was, and b) why the action was morally wrong (e.g., it violated norms of care, fairness, authority, etc.). They were also asked to rate the comprehensibility, imageability (how easily they could imagine it), and frequency (how often they see it in real life) of each vignette, as well as the strength of their emotional response (1-5). Clifford et al. (2015) only kept vignettes that were classified as the intended foundation by at least 60% of participants and excluded vignettes that were classified as an unintended foundation by 20% or more of participants. After narrowing down the vignettes, Clifford et al. (2015) further examined the internal and external validity of their vignettes. Using both exploratory and confirmatory factor analysis, they determined underlying factors associated with each of the five moral foundations. They also found evidence of a division for the care foundation (into emotional, physical human, and physical animal harm). Subscales created from the moral foundation vignettes were predicted by the extensively validated and corresponding MFQ subscales (βs = .24-.47; all ps < .001) and showed expected relationships with political ideology. Overall, Clifford et al. (2015) produced a set of 90 vignettes for future research on moral foundations and extended existing measures by assessing judgments of concrete moral violations.

Given the reliance on concrete situations, rather than abstract moral judgments, the vignettes developed by Clifford et al. (2015) are a good model for a child measure of moral foundations. Therefore, prior to the current study, Curtis and Moore (2019) developed a similarly designed verbal questionnaire for children aged 4 years and older. Like Clifford et al. (2015), they created vignettes that started with "You see…" to increase the likelihood that children would imagine the situations. To aid children's comprehension, they also paired each vignette with a stick figure picture of the scenario. Simple language was used, although because scenarios were read out loud to children, reading level was not accounted for. In their initial development of the vignettes, Curtis and Moore (2019) attempted to cover specific concerns for each foundation. For the care foundation, this included physical harm to a human, emotional harm to a human, and physical harm to a non-human animal. For the fairness foundation, vignettes depicted cheating, inequality, and free riding (e.g., not helping clean up a mess they made). The authority foundation included both disobedience and disrespect, while the loyalty foundation involved betraying ingroups such as family (e.g., telling siblings' secrets), teams (e.g., celebrating with opposite team), and friends (e.g., not helping friend who is bullied). For the sanctity foundation, vignettes initially included scenarios depicting physical disgust or contamination (e.g., rubbing poop on self), violations of privacy (e.g., spying on someone in the shower), violations of sacred objects (e.g., ripping up photo of a deceased family member), and environmental degradation. Finally, like Clifford et al. (2015), Curtis and Moore (2019) developed vignettes showing social norm violations, to serve as a comparison for how bad children rate things in general, and so they did not expect a moral violation in every vignette.

After developing their initial vignettes, Curtis and Moore (2019) examined the suitability, validity, and comprehensibility of their vignettes. They administered the vignettes to 50 parents of children aged 4 to 11 years old. Items that received low approval from parents (less than 60% approval) were excluded from the final measure. In addition to parents, Curtis and Moore (2019) administered the vignettes to 81 undergraduate students. Following Clifford et al. (2015), they asked these students to rate why they felt each item was bad (e.g., it violated norms of care, fairness, etc.). Items with less than 70% agreement with the intended moral foundation were excluded from the final measure. Curtis and Moore (2019) also administered the previously validated Moral Foundations Questionnaire (MFQ, Graham et al., 2011) to the same sample of undergraduate students. They found modest to moderate positive correlations between relevant foundations on the MFQ and moral foundation vignettes for children (ranging from $r = .22$, $p = .126$ to $r = .55$, $p<.001$). Finally, Curtis and Moore (2019) piloted the remaining vignettes with a sample of 28 children aged 4 to 10 to ensure items were appropriate and comprehensible.

The revision of the moral foundation vignettes for children according to parents' approval, foundation endorsed, and age appropriateness led to a reduction in the breadth of areas covered under each foundation. For the care foundation, at least one vignette remained depicting each physical harm to a human, emotional harm to a human, and physical harm to a non-human animal. For fairness, items covered cheating and inequality, but free riding proved too difficult to represent with age appropriate items. For the loyalty foundation, the best vignettes depicted disloyalty to a group club, and disloyalty to a sibling. For authority, both disobedience and disrespect were represented

in the vignettes. Finally, for the sanctity foundation, numerous items were deemed inappropriate by parents, while others were not endorsed as the intended foundation. Therefore, the remaining items all involved elements of physical disgust or contamination. In their final measure, the Moral Foundations Questionnaire for Kids (MFQK), Curtis and Moore (2019) included four vignettes for each of the five moral foundations, as well as four vignettes depicting social norms violations. Children are asked to rate the badness of each violation, if the act should be punished, and how it makes them feel. The full details of the MFQK are described in Chapter 2, and the full scale is shown in Appendix C.

Overall, the primary goal of this chapter was to get a broader picture of 4-year-olds' concern for moral foundations using a verbal measure (MFQK). We sought to examine and compare their degree of concern for each moral foundation, as well as correlations between the different foundations. We also compared 4-year-olds responses to moral violations on the MFQK and Moral Foundations Puppet Task (MFPT, used in Chapter 3). Our hypotheses and justifications are listed below:

H1: We expected that 4-year-olds would show a high degree of concern for care and fairness. That is, on average, they would rate violations in these foundations as 'very bad' or 'very very bad.'

Justification: There is extensive literature showing that 4-year-olds are sensitive to violations of care and fairness.

H2: For the loyalty, authority, and sanctity foundations, we expected that children would rate violations as at least a little or quite bad.

Justification: As previously outlined, although there is relatively less literature,

there is still evidence that 4-year-olds are sensitive to disloyalty and disobedience (Abrams et al., 2003; Smetana et al., 2014). For the sanctity foundation, although the literature is scarce, our prediction is based on the premise that all of the MFQK sanctity vignettes involve physical disgust (e.g., feces), and such issues are commonly discussed during the preschool years (e.g., toileting, cleanliness).

H3: We expected that 4-year-olds would show significantly less concern for the authority and loyalty foundations, and rate violations as significantly less bad, relative to the care and fairness foundations.

Justification: Some researchers have found that children rate disobedience and group violations as less severe than physical or relational acts of harm or injustice (Atkin & Gummerum, 2012; Smetana et al., 2012).

H4: Similar to Chapter 3, we expected to see significant, positive correlations between children's scores for the care and fairness, loyalty and authority, loyalty and sanctity, and sanctity and authority foundations on the MFQK.

Justification: Previous research has shown strong positive relationships between the individualizing (care and fairness) and binding (loyalty, authority, sanctity) foundations (Graham et al., 2011).

H5: We expected to see significant positive correlations between 4-year-olds' responses to the same foundations (e.g., care with care, fairness with fairness) on the MFPT and MFQK.

Justification: Although these measures were slightly different – the MFPT involved preference and judgment, while the MFQK involved only judgment – they do both involve judging the badness or acceptability of a particular moral violation.

Methods

The primary measure of interest for this chapter and analysis is the Moral Foundations Questionnaire for Kids (MFQK), described in detail in Chapter 2. As a brief reminder, 76 four-year-olds completed this task. They viewed vignettes depicting violations relevant to each of the five moral foundations (care, fairness, loyalty, authority, sanctity), as well as social norm violations. Children indicated a) if they thought the action was bad, b) how bad they thought it was, c) if the actor should be punished, and d) how the act made them feel. Summary scores were created for each foundation and the odd behaviours (social norms) subscales that ranged from 0-5, where higher scores indicated greater concern for a particular foundation. Responses to the punishment and emotion questions were recorded as categorical variables. All 4-year-olds completed the MFQK, and none were excluded. However, some missing data was produced due to child error (hitting the wrong button) or computer error (e.g., erroneously skipping questions). These errors were infrequent (2.72% of all MFQK data) and occurred randomly.

In addition to the MFQK, data from the Moral Foundations Puppet Task (MFPT, described in Chapters 2 and 3) was used to examine correspondence between the two measures. Specifically, we used the summary variables that were created for each foundation. These were ordinal level variables, ranging from 0-4, where higher scores indicated greater concern for the relevant moral foundation. The majority of 4-year-olds completed the MFPT before the MFQK; however, a small subset of participants (approximately 15%) completed it after due to the availability of daycares, parents, or lab volunteers.

Data Analysis

Given that children responded to items on the MFQK using a Likert scale, the data produced were ordinal in nature. Although summary scores were averaged across items to produce summary variables, they are still not truly continuous in nature – and, as shown in the next section, the data were certainly not normally distributed. Therefore, for all analyses conducted in this section, non-parametric tests were used.

Hypotheses 1 and 2 concerned how sensitive children would be to each of the 5 foundations, or how bad they would rate violations. These hypotheses were tested primarily by looking at the distributions, means, and other descriptive statistics (e.g., median) for each of the moral foundation summary scores produced by the MFQK. Children's responses to the odd behaviours subscale were also taken into consideration, as they served as a point of reference for how bad 4-year-olds rate actions in general.

Hypothesis 3 concerned comparisons between children's badness ratings for each foundation – specifically, between the care, fairness, loyalty, and authority foundation. To examine this hypothesis, we conducted a Friedman test, which is a nonparametric repeated measures test that uses and compares mean ranks rather than mean scores. This test served as an omnibus test indicating if children's mean badness ranks differed significantly across foundations on the MFQK. Follow-up pairwise comparisons were also conducted using related-samples Wilcoxon signed rank tests. Of note, although no predictions were made regarding how the sanctity foundation would compare to others, it was included in these analyses for exploratory purposes. Children's scores on the odd behaviours subscales were also included, again to serve as a point of reference. Furthermore, following the same procedure (i.e., Friedman's test as an omnibus test, and

Wilcoxon signed rank tests for follow-up pairwise comparisons), we tested if children's responses differed significantly between items within the same foundation or subscale. This analysis was exploratory and meant to help with interpretation of Hypotheses 1-3.

Hypothesis 4 involved predictions about correlations between foundations on the MFQK. To test this hypothesis, we conducted Spearman's correlations between each foundation or subscale. Of note, although we only made predictions regarding correlations between some foundations, all possible inter-subscale (care, fairness, loyalty, authority, sanctity, odd behaviours) correlations were examined for exploratory purposes and the sake of completeness. Similarly, Hypothesis 5 involved predictions regarding correlations between children's responses to moral foundations on the MFQK and MFPT. Spearman's correlations were used to test these hypotheses, and again we examined all possible correlations.

Finally, the last analyses pertained to the punishment and emotion questions on the MFQK. These analyses were purely descriptive and exploratory, and meant to potentially provide directions to future research. We sought to examine if there are patterns in which acts children rate as punishable, and which emotions they endorse after learning of a moral violation. Specifically, we examined the average percentage of response options endorsed for the emotion and punishment questions.

As in Chapter 3, the Benjamini and Hochberg (1995) procedure was again used to control for false discovery rates across multiple comparisons (see Chapter 3 for a complete description). Therefore, any corrected significance thresholds reported in the results section were calculated using this procedure.

Results

Examining 4-year-olds' moral concerns using the MFQK (H1-H2)

The distribution of 4-year-olds' average badness ratings for each of the moral foundations and odd behaviours subscales are shown in Figure 2, while descriptive statistics are shown in Table 13. The distribution for the care foundation was quite platykurtic, but with a peak in the highest values. It appeared to be negatively skewed, with evidence of a ceiling effect. The distributions of the authority and sanctity foundations were very similar, although slightly less platykurtic. The distribution for the fairness foundation appeared negatively skewed, with more endorsement of higher values. However, there was less evidence of a ceiling effect relative to care, authority, and sanctity. In contrast to the other foundations, both the loyalty and odd behaviours subscales showed very platykurtic distributions; while the loyalty foundation showed small peaks across the distribution, they are not consistently congregated towards higher, lower, or middle scores. Overall, none of the distributions appeared to be normally distributed.

Hypotheses 1 and 2 concerned how sensitive 4-year-olds would be to violations in each area. Considering the skewness of each distribution, when examining these hypotheses, we felt it would be prudent to consider the mean, median, and mode of each foundation (see Table 13). In general, it appears as though 4-year-olds rated care, authority, and sanctity violations as very bad (~4), while fairness violations were rated approximately 'quite' bad (~3). Loyalty and social norm behaviours were rated, on average, 'a little' bad (~2).

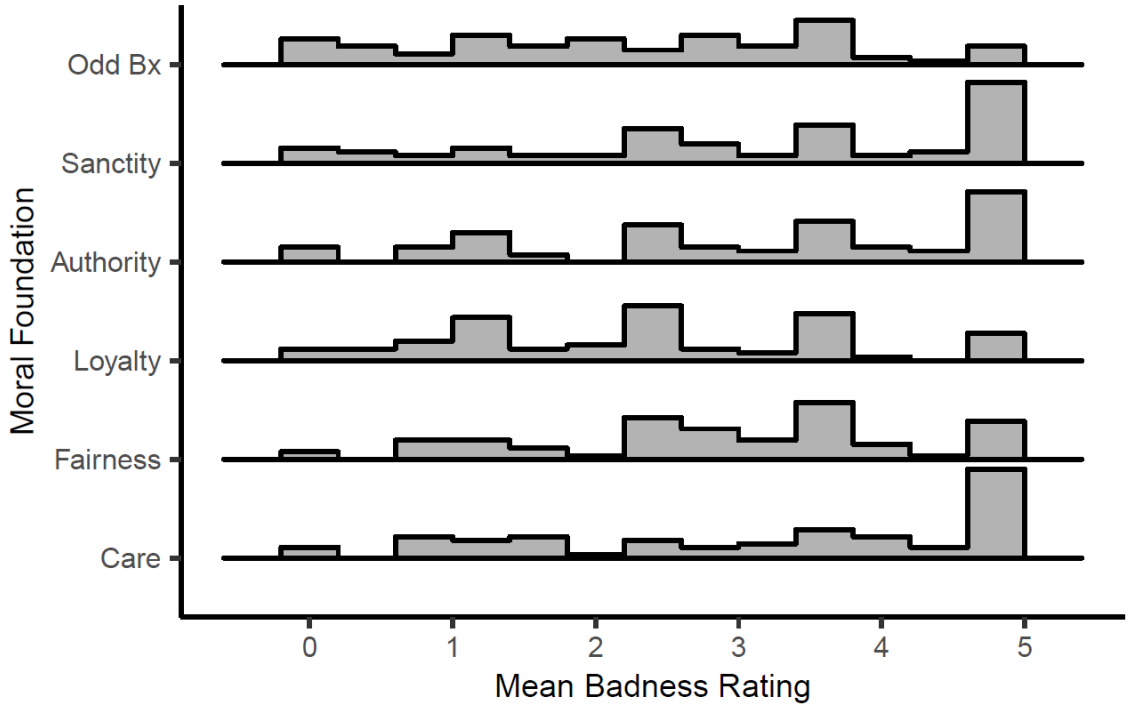


Figure 2. Distributions of average badness ratings for subscales of the Moral Foundations Questionnaire for Kids (0 = not bad, 5 = very very bad). Height of bars indicate relative frequency of scores. Sample sizes were 74 (care), 69 (fairness), 67 (loyalty), 71 (authority), 68 (sanctity), and 71 (odd behaviours).

Table 13

Descriptive statistics for summary variables representing 4-year-olds' responses to moral and social norm violations on the MFQK

Foundation	*N*	Mean	*SD*	Median	Mode	Min	Max
Care	74	3.29	1.58	3.75	5.00	0	5.00
Fairness	69	2.98	1.32	3.00	3.75	0	5.00
Loyalty	67	2.41	1.34	2.25	1.25	0	5.00
Authority	71	3.14	1.58	3.50	5.00	0	5.00
Sanctity	68	3.21	1.63	3.75	5.00	0	5.00
Odd Bx	71	2.35	1.43	2.59	1.25	0	5.00

Comparing Moral Foundations on the MFQK (H3)

To examine our predictions regarding differences in 4-year-olds' concern for the different moral foundations, we conducted a Friedman test. This revealed that children's mean badness ranks differed significantly across foundations, χ^2 (5) = 34.15, N = 57,

p<.001. To examine differences between specific foundations, we conducted pairwise related-samples Wilcoxon signed rank tests (corrected significance level = .03). The results of these pairwise comparisons are visualized in Figure 3. The care foundation was significantly different from the fairness (p = .021), loyalty (p<.001), and odd behaviours (p<.001) subscales, but was not different from the authority (p = .565) and sanctity (p = .877) foundations. The fairness foundation was significantly different from loyalty (p = .006) and odd behaviours (p<.001), but was not different from authority (p = .146) or sanctity (p = .086). Similarly, authority was significantly different from loyalty (p<.001) and odd behaviours (p<.001), but was not different from sanctity (p = .571). Finally, sanctity was significantly different from loyalty (p = .001) and odd behaviours (p<.001), while there was no difference between loyalty and the odd behaviours subscales (p = .949).

In Hypothesis 3, we predicted 4-year-olds would show significantly less concern for the authority and loyalty foundations relative to the care and fairness foundations. To summarize, 4-year-olds did show significantly less concern for the loyalty foundation relative to the care and fairness foundations. Unexpectedly, they showed greater concern for care than fairness, and greater concern for authority than loyalty. Their sensitivity to the authority foundation did not differ significantly from the care or fairness foundations. Other comparisons were exploratory; we also saw that 4-year-olds rated violations of loyalty and social norms significantly less bad than violations in any other foundation (including sanctity). No significant differences were found between how they rated violations of loyalty and social norms, or how they rated violations of sanctity and violations of care, fairness, or authority.

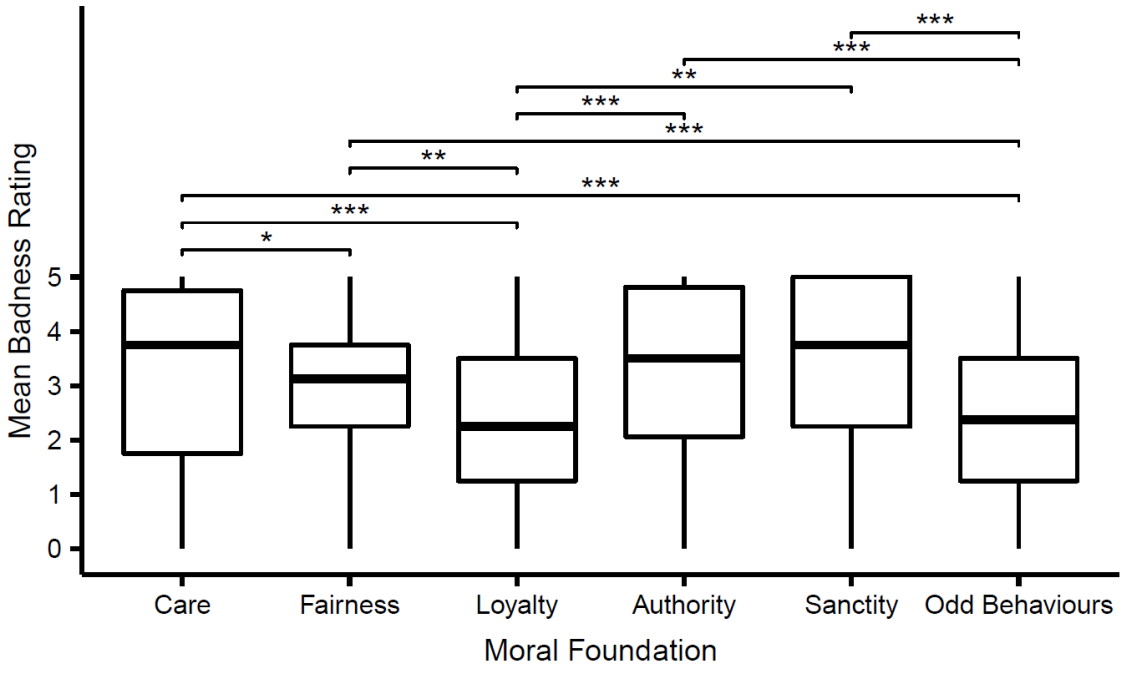


Figure 3. Results of pairwise comparisons between 4-year-olds' mean badness ratings on the Moral Foundations Questionnaire for Kids, using related samples Wilcoxon signed rank tests. Significant pairwise comparisons are shown; *p<.03, **p<.01, ***p<.001. Corrected significant threshold for these analyses was .03, following application of the Benjamini and Hochberg (1995) procedure.

Exploring 4-year-olds' responses to individual items on the MFQK

To further assist with interpretations Hypothesis 3, we examined children's responses to individual items within the same subscale or foundation. The percentage of 4-year-olds that selected each response option (ranging from 0 = not bad, to 5 = very very bad) across items is shown in Table 14, as well as the mean badness rating for each individual item. We conducted a series of Friedman tests to examine significant differences in item responses within subscales (care, fairness, loyalty, authority, sanctity, ad odd behaviours). Children's mean badness ranks did not differ significantly across items on the care foundation (χ^2 = 7.40, N = 75, DF = 3, p = .060), fairness foundation (χ^2 = 7.65, N = 70, DF = 3, p = .054), authority foundation (χ^2 = 3.20, N = 72, DF = 3, p = .362), or sanctity foundation (χ^2 = 6.46, N = 69, DF = 3, p = .091).

Table 14

Percentage of participants that endorsed each response option on each MFQK item

Foundation	Item	Response Options						Mean (*SD*)
		0	1	2	3	4	5	
Care	You see a girl punch another girl in the stomach	21.1	9.2	1.3	3.9	6.6	57.9	3.39 (2.14)
	You see a girl calling a girl stupid	23.7	3.9	0	1.3	9.2	61.8	3.54 (2.15)
	You see a girl stomp on the tail of her pet cat	20.0	13.3	5.3	9.3	14.7	37.3	2.97 (2.02)
	You see a girl kick a stray dog	26.3	5.3	2.6	6.6	7.9	51.3	3.18 (2.18)
Fairness	You see a girl cheating in a race by taking a shortcut	23.6	5.6	0	1.4	12.5	56.9	3.44 (2.15)
	You see a girl cheating in a board game	32.9	6.6	1.3	3.9	9.2	46.1	2.88 (2.29)
	You see a girl cut to the front of the line	33.3	9.3	4.0	9.3	16.0	28.0	2.49 (2.12)
	You see a girl taking all of the cookies, leaving none for others	29.3	4.0	4.0	6.7	6.7	49.3	3.05 (2.22)
Loyalty	You see a girl telling secrets about her sister to people her sister doesn't like	48.6	15.3	0	2.8	8.3	25.0	1.82 (2.18)
	You see a girl reading her sister's diary	31.5	4.1	1.4	4.1	5.5	53.4	3.08 (2.30)
	You see a girl teach a secret password to people who are not in her club	36.5	10.8	5.4	5.4	10.8	31.1	2.36 (2.19)
	You see a girl score a goal against her own team to help the other team win	43.2	9.5	1.4	4.1	8.1	33.8	2.26 (2.29)
Authority	You see a girl ignore her parents when they tell her to stop watching TV	30.3	5.3	1.3	5.3	11.8	46.1	3.01 (2.23)
	You see a girl wearing a hat at school even after the teachers asks her not to	30.7	13.3	0	2.7	1.3	52.0	2.87 (2.33)
	You see a girl calling her teacher bad words	22.7	4.0	1.3	12.0	8.0	52.0	3.35 (2.08)
	You see a girl calling her parents bad words	25.7	1.4	0	5.4	13.5	54.1	3.43 (2.12)
Sanctity	You see a girl rubbing poop on herself in the shower	27.4	1.4	1.4	2.7	13.7	53.4	3.34 (2.18)
	You see a girl using a dirty diaper as a pillow	25.3	6.7	1.3	4.0	6.7	56.0	3.28 (2.20)
	You see a girl drinking pee with her dinner	21.6	8.1	1.4	2.7	10.8	55.4	3.39 (2.13)
	You see a girl loudly burping and farting while eating	34.7	8.0	1.3	6.7	5.3	44.0	2.72 (2.29)
Odd Behaviour	You see a girl sleeping in her raincoat, instead of pajamas	36.5	6.8	2.7	2.7	8.1	43.2	2.69 (2.32)
	You see a girl wearing her pajamas to school	41.3	9.3	2.7	4.0	4.0	38.7	2.36 (2.32)
	You see a girl eating her soup with a fork	56.6	14.5	2.6	6.6	2.6	17.1	1.36 (1.94)
	You see a boy eat dessert before dinner is served, instead of eating it afterwards	25.3	9.3	1.3	6.7	10.7	46.7	3.08 (2.17)

Note. 0 = not bad, 1 = not very bad, 2 = a little bad, 3 = quite bad, 4 = very bad, 5 = very very bad

Children's mean badness ranks did differ significantly across items on the loyalty foundation ($\chi^2 = 11.08$, $N = 68$, $DF = 3$, $p = .011$). Specifically, using pairwise related-samples Wilcoxon signed rank tests (corrected significance threshold = .017), we saw that children rated Item 2 (you see a boy/girl reading their siblings' diary; $M = 3.08$, $SD = 2.30$) as significantly more bad than Items 1 (you see a boy/girl telling secrets about their sibling to someone their sibling doesn't like; $M = 1.82$, $SD = 2.18$, $p = .001$) and 4 (you see someone score a goal against their own team; $M = 2.26$, $SD = 2.29$, $p = .015$).

Similarly, children's mean badness ranks differed significantly across items on the odd behaviours scale, $\chi^2 = 20.63$, $N = 72$, $DF = 3$, $p < .001$. Pairwise comparisons revealed (corrected significance threshold = .025) that children rated Item 3 (you see a boy/girl eating soup with a fork, $M = 1.36$, $SD = 1.94$) significantly less bad than Item 1 (you see a boy/girl sleeping in a raincoat; $M = 2.69$, $SD = 2.32$, $p<.001$), Item 2 (you see a boy/girl wearing pajamas at school; $M = 2.36$, $SD = 2.32$, $p = .003$), and Item 4 (you see a boy/girl eating dessert before dinner; $M = 3.08$, $SD = 2.17$; $p<.001$).

Correlations between Moral Foundations (H4-H5)

To examine relationships between 4-year-olds' responses to violations of moral foundations and social norms (H4), we conducted Spearman correlations between subscale scores on the MFQK (significance threshold of .05 was unchanged). As seen in Table 15, there were significant positive correlations between all subscales on the MFQK. Therefore, children who rated a moral violation as wrong in one foundation were also more likely to rate violations as wrong in any other foundations, as well as violations of social norms. The strong positive correlations may suggest the presence of a confounding variable, such as shared method variance.

Table 15

Correlations between 4-year-olds' moral concerns on the Moral Foundations Questionnaire for Kids

	Care	Fairness	Loyalty	Authority	Sanctity	Odd Behaviours
Care	1					
Fairness	.55**	1				
Loyalty	.48**	.48**	1			
Authority	.63**	.55**	.46**	1		
Sanctity	.51**	.48**	.46**	.51**	1	
Odd Behaviours	.35*	.55**	.52**	.39*	.54**	1

Note. **p<.001, *p<.01. The Benjamini and Hochberg (1995) procedure was applied, there was no change to the significance threshold of .05.

In addition to correlations between foundations on the MFQK, we examined correlations between 4-year-olds' responses on the MFQK (2-year-olds did not complete the MFQK) and the Moral Foundations Puppet Task (MFPT; H5) using Spearman's correlations. After applying a corrected significance threshold of = .0017, there were no significant correlations between children's scores on the MFPT and MFQK (Table 16). However, it is important to note that there was very low variability in 4-year-olds' responses on the MFPT and MFQK, with the exception of the loyalty (MFQK and MFPT) and odd behaviours (MFQK) scales.

Table 16

Correlations between 4-year-olds' responses on the Moral Foundations Questionnaire for Kids (MFQK) and Moral Foundations Puppet Task (MFPT)

			MFPT		
MFQK	Care	Fairness	Loyalty	Authority	Sanctity
Care	.01	-.04	.23	.14	-.02
Fairness	.01	.05	.28	.12	-.00
Loyalty	.10	.01	.21	.02	-.16
Authority	.08	.09	.23	.10	-.10
Sanctity	.02	.06	.16	.01	.07
Odd Behaviours	-.06	-.12	.23	-.05	-.10

Note. Following the Benjamini and Hochberg (1995) procedure, corrected significance threshold for theses analyses was .0017.

Emotion and Punishment Responses

The final investigations of this chapter were exploratory and involved the emotion and punishment questions on the MFQK. Figure 4 shows the average percentage of times each emotion (happy, sad, angry, yucky, neutral) was chosen for a given foundation. Percentages were averaged across participants' responses to the four questions of each foundation. Happy was endorsed the most for the loyalty and odd behaviour vignettes. Sad was endorsed the most for the care and fairness vignettes, and the least for sanctity related violations. Angry was endorsed the most for care and authority, and the lowest for loyalty and social norm violations. Yucky was endorsed the most for the sanctity foundation, and relatively little for the remaining scales. Finally, neutral was endorsed rather evenly across the scales, but was slightly higher for the odd behaviours vignettes.

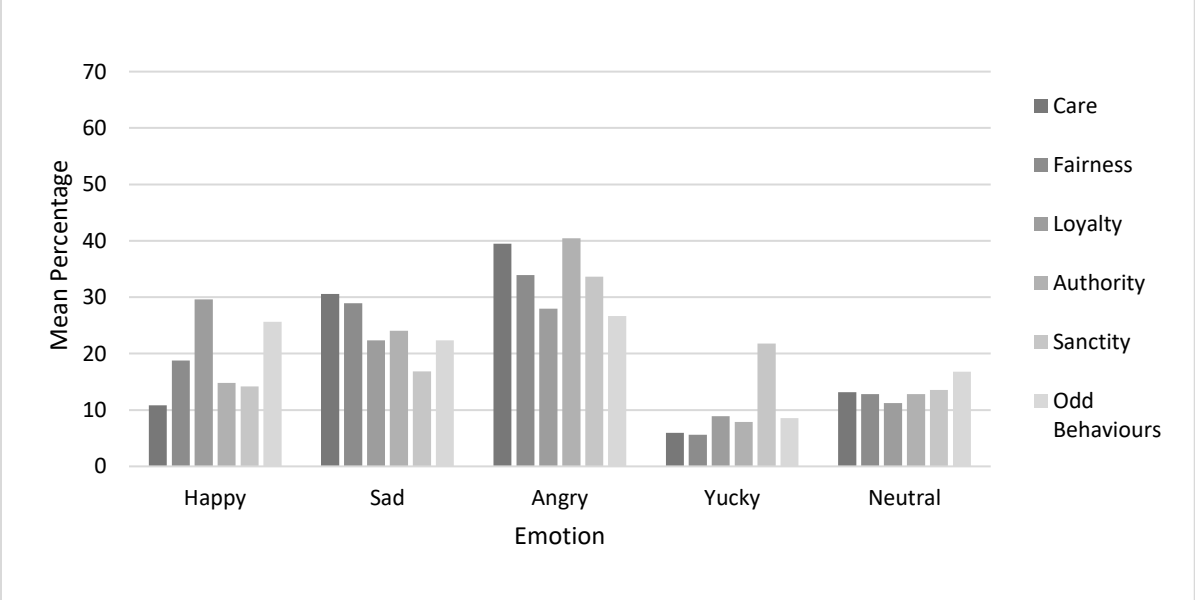


Figure 4. Average (across four vignettes) number of times each emotion was endorsed for a violations of moral foundations and social norms (odd behaviours).

Punishment responses were examined in a similar fashion as the emotion responses. We looked at the percentage of times 'yes' versus 'no' was chosen in response to the question "should he/she be punished?" following a moral foundation or social

norm violation. We also included the percentage of times an action was rated as not bad, because children who selected this response were not asked if the actor should be punished. These percentages were averaged for each scale of the MFQK (four vignettes per scale; see Figure 5). Of note, percentages add to 100 across all three response options (yes punish, no do not punish, not bad). For example, across all four items of the care foundation, on average participants responded that 'yes' the act was bad *and* 'yes' the actor should be punished 51.78% of the time. They responded that 'yes' the act was bad *but* no the actor should not be punished 25.45% of the time, while on average participants said that violations were *not* bad 22.78% of the time.

As shown in Figure 5, there were very few remarkable patterns in children's responses to the punishment questions. The loyalty and odd behaviour violations were rated as 'bad' less frequently and were also punished less frequently by our 4-year-old participants.

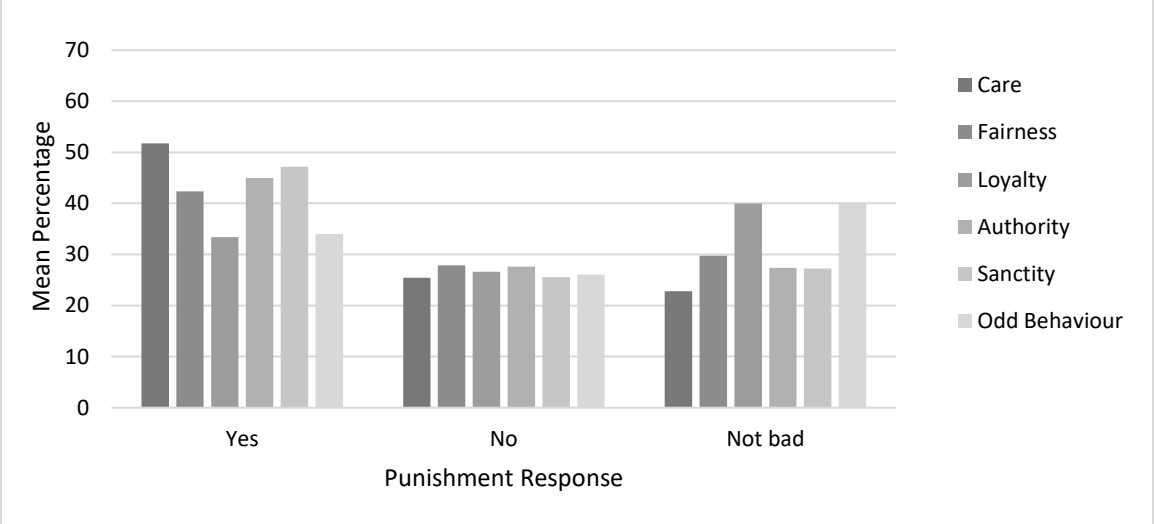


Figure 5. Average (across four vignettes) number of times 4-year-olds endorsed 'yes he/she should be punished' versus 'no he/she should not be punished' for violations of moral foundations and social norms (odd behaviours). Children who rated a particular violation as 'not bad' were not asked about punishment; average percentage of 'not bad' responses are also shown.

Discussion

The purpose of the current chapter was to obtain a broader picture of 4-year-olds' concern for moral foundations using a verbal measure, the Moral Foundations Questionnaire for Kids (MFQK). This measure allowed us to examine and compare degrees of concern for the different foundations. It also allowed us to examine 4-year-olds' moral concerns as assessed by two different measures: the MFQK, and the Moral Foundations Puppet Task (MFPT) utilized in the previous chapter. Regarding degree of concern for moral foundations, we found that, as expected, 4-year-olds rated violations of care and fairness as very bad. Interestingly, although authority violations were expected to be rated as less severe than care or fairness, both authority and sanctity violations were rated as equally bad as care and fairness violations (no comparison predictions were made regarding the sanctity foundation). In contrast, loyalty violations were rated significantly lower than violations in all other foundations, and were comparable to social norm violations (odd behaviours). The pattern of correlations found within the MFQK and between the MFQK and MFPT were unexpected; we found significant positive correlations between all foundations on the MFQK (rather than just between individualizing and binding foundations, as predicted), and no significant correlations between the MFQK and MFPT.

4-year-olds' concern for the five moral foundations

The results found for the care and fairness foundations are consistent with what has been found in previous research, and what was shown in the previous chapter on the MFPT. Four-year-olds from liberal, Western backgrounds are known to be sensitive to both physical and emotional harm to other humans (e.g., Smetana et al., 2012), as well as equality (e.g., Rakoczy et al., 2016). Given the scenarios included in the MFQK, the

current study also suggests that they are sensitive to physical harm to non-human animals, and cheating. Interestingly, we found that 4-year-olds rated care violations as significantly worse than violations of fairness. However, looking at an item level, some fairness items were rated higher (i.e., more bad) than care items, and vice versa. Upon further examination of Appendix G, it appears as though the item 'you see a boy/girl cut to the front of the line,' although not significantly different from other items, was rated lower and may have pulled down the average badness rating for fairness While this item was worded to be simple and concise, it is possible that not all 4-year-olds fully understood what 'cutting the line' meant. They may understand or be more likely to use more literal language such as 'went in front of me' or 'took my spot'. Overall, our primary take-away for these two foundations is that, as shown repeatedly by previous research, 4-year-olds are sensitive to different kinds of violations in the care and fairness foundations.

Interestingly, our sample of 4-year-olds rated authority and sanctity violations just as bad as violations of care and fairness. This was an unexpected result for the authority foundation, given previous research showing that 4-year-olds rate disobedience as less severe than physical or relational acts of harm (Smetana et al., 2012). However, it is consistent with other research showing that younger (i.e., 4- to 5-year-old) more likely to accept authority than older children, and generally rate more acts of all kinds as morally wrong (Jambon & Smetana, 2014; Laupa & Turiel, 1993). Depending on the framework employed, there are a few different interpretations of the finding that 4-year-olds ratings of care, fairness, authority, and sanctity violations were not significantly different. From a traditional, Kohlbergian point of view, this would be interpreted as 4-year-olds being at

an earlier, more primitive form of moral reasoning or development. At this stage, they view rules as absolute, and contamination concerns as moral - relative to higher moral stages where people form their own opinions and better distinguish between what is moral versus what is a social norm. However, from an MFT perspective, there are no higher or lower levels of moral reasoning – just different moral beliefs (Haidt, 2012). Following MFT, we might suggest that the 4-year-olds in our sample were sensitive to authority and sanctity due to both an evolutionary predisposition, as well as environmental influences. During the preschool years, a great deal of emphasis is placed on rules (e.g., rules at daycare, rules about shapes or colours, social norms, etc.) and contamination (e.g., toileting, proper hand washing, safe food). Therefore, although it would not be unreasonable to say that 4-year-olds are at an earlier developmental stage in many ways, it is also important to remember that younger and older children receive different types and amounts of messages regarding concerns of rules, respect, and purity.

Relative to all other violations (except for odd behaviours) 4-year-olds rated loyalty violations as less bad. On average, they did rate these violations as 'a little bad;' however, considering this was not significantly different from how they rated social norms, it does not strongly support concern for loyalty. While a lack of concern was unexpected, it is consistent with findings on the MFPT. One possible explanation is the evaluation of third parties, rather than participant's own teammates as in previous studies (e.g., Abrams et al., 2003). Other explanations include cognitive development and poor comprehension of others' ingroups or outgroups, as well as environmental influence – especially in liberal, Western societies, young children are often taught to include everyone and support other teams. Another possibility could be that, contrary to the

predictions of MFT, children are not predisposed to develop concerns in this area; however, this is impossible to conclude without more investigation. Interestingly, there were some differences in how children rated items within the loyalty foundation – specifically, they rated Item 2 (reading your sibling's diary) as worse than all other items. While other items specified an outgroup (e.g., people your sibling does not like, opposite team), this item did not. Therefore, although it was a form of family betrayal, it is possible that without a specific outgroup, 4-year-olds viewed this more in line with emotional harm to another person, and therefore rated it as worse. Overall, especially when taken together with the findings presented in the previous chapter, we think it is reasonable to suggest that our sample of 4-year-olds was not sensitive to third-party loyalty violations.

Like the loyalty foundation, odd behaviours (social norm violations) were rated as significantly less bad than other moral foundations. This is consistent with previous research showing that 4-year-olds are able to distinguish between moral and conventional concerns (Smetana et al., 2012). However, it is interesting that, in that line of research (e.g., Laupa & Turiel, 1993; Smetana et al., 2014), authority violations (e.g., like not going to bed when asked) were seen as conventional (rather than moral) concerns. Therefore, in the current study, the finding that 4-year-olds distinguished between concerns traditionally considered conventional (authority, sanctity) and social norm violations lends credence to MFT's assertion that such concerns are indeed morally relevant. However, similar to the loyalty foundation, there were significant differences in how 4-year-olds rated the items within the odd behaviours subscale. Most notably, 'eating soup with a fork' was rated significantly lower than all other items. The other

items (e.g., wearing pajamas to school, eating dessert before dinner, wearing raincoat to bed) could potentially, at least for young children, invoke feelings or concerns about authority. Young children are typically not allowed to eat dessert before dinner, or sleep in outdoor clothing. Therefore, while on average these items were rated lower than most other foundations, further revision may be required to reduce possible confounds, and to be more appropriate for use as a control or comparison scale.

In addition to showing how items within a specific subscale held together, the item analysis shown in Table 14 also raises some concerns regarding 4-year-old's use of the badness scale. Across all items, there are noticeable peaks at the 0 (not bad) and 5 (very, very bad) response options, while the middle items (a little bad, quite bad) were used the least. Gelman and Baillargeon (1983) noted that, following Piagetian reasoning about cognitive development, young children struggle with shades of grey and rely primarily on dichotomous thinking. Therefore, they may be more likely to focus on extremes when using Likert scales. Chambers and Johnston (2002) specifically studied children's use of Likert scales across subjective and objective items. They found that for social and emotional items (e.g., how would you or someone feel in X scenario), younger children (5- to 6-year-olds) used extreme response options more than older children. There was no difference in extreme scores as a function of response options – i.e., it did not matter if there were 3 or 5 options. Mellor and Moore (2014) found similar results, and questioned the use of Likert scales for all children under the age of 12. For our purposes, this suggests that children's scores on the MFQK may not represent their true concern for each scale or foundation. It is difficult to say what the true scores should be, as the impact would have been in both directions – extreme low scores might be slightly

higher, extreme high scores might be slightly lower. Overall, although we feel that the MFQK showed greater variability in children's values than the MFPT, given that use of middle options was not zero, it is important to recognize that extreme scores were the norm, and there may be finer gradations in children's concern for the moral foundations assessed. This is especially important when looking to conduct and compare research across different age groups.

Correlational and exploratory analyses

The two remaining primary analyses of interest in this chapter were correlational in nature. Firstly, we examined correlations between 4-year-olds responses to the six different subscales (care, fairness, loyalty, authority, sanctity, and odd behaviours) on the MFQK. The resulting pattern of correlations was certainly unexpected, instead of seeing specific positive correlations between the individualizing (care and fairness) and binding (loyalty, authority, sanctity) foundations, we saw consistently strong, positive correlations between all possible scales. There are a few possible interpretations of this. It is possible that, at a young age, foundations are less specialized, and being sensitive to one predicts being sensitive to another (as well as to social norm violations). It is also possible that these correlations are driven by developmental status, such as the ability to translate moral beliefs into moral judgments. However, as alluded to briefly in the results section, we feel that common method variance was likely a strong contributor, making it difficult to put too much stock into the correlations that were found. Podsakoff et al., (2003) defined common method variance as "variance that is attributable to the measurement method rather than to the constructs the measures represent" (p. 879). If participants respond in a similar manner across survey items, common method variance

can create false or inflated correlations. Anecdotally, children who rated one item as bad appeared to rate many other items as bad, and vice versa. Therefore, the correlations found for the MFQK are likely more reflective of response tendencies across identical question formats, and it is difficult to say what the true relationships are. The correlations observed between the moral foundations and odd behaviours subscales also suggest that common method variance is plausible.

One of the best ways to address common method variance is using different methods and different reporters to measure the same construct (Podsakoff et al., 2003). In the current study, we used two different measures to assess moral concerns in 4-year-olds: the MFQK and MFPT. However, as shown in Table 16, the correspondence between these measures was poor; no significant correlations were found between any of the foundations. This is interesting given that a similar pattern of results, where 4-year-olds were sensitive to all moral foundations except for loyalty, was found in both measures. It could suggest numerous interpretations – there may have been a difference in children's comprehension of the two measures, children may respond differently to more auditory versus more visual stimuli. However, as with the previously discussed correlations, we think there is likely a larger confound or issue at play – in this case, the lack of variation in 4-year-old's responses across both measures. Especially for the care, fairness, authority, and sanctity foundations, there were considerable ceiling effects in children's MFQK responses, and possibly even less variability in their MFPT responses. This interpretation is further strengthened by examining the pattern of correlations between the two measures; for example, the loyalty foundation on the MFPT showed the strongest correlations with other foundations on the MFQK, and this foundation provoked

the greatest variability in children's responses. The lack of variation and impact on correlations was likely compounded by issues previously discussed, such as 4-year-old's use of Likert scales.

In addition to primary analyses, we conducted two exploratory analyses to examine 4-year-old's responses to the emotion and punishment questions on the MFQK. Given that these analyses were purely descriptive, no concrete conclusions can be drawn, but some patterns are interesting nonetheless. Out of five possible emotions (happy, sad, angry, yucky, neutral), 'happy' was endorsed the most for loyalty and odd behaviour items, which further supports the notion that 4-year-olds were less sensitive to these violations and less likely to view them as moral concerns. 'Angry' was endorsed the most in general across all foundations, which fits with previous research suggesting that anger is a common, non-specific moral emotion (e.g., Russell & Giner-Sorolla, 2011). Finally, 'yucky' was endorsed by far the most for the sanctity foundation, which is similarly consistent with research suggesting that disgust is specific to purity violations (e.g., Horberg et al., 2009). Of note, the measure we used (MFQK) only included physical disgust, so it would be interesting to see if children respond similarly to other forms of sanctity violations. For the punishment question, there are fewer noticeable patterns. In general, for foundations that 4-year-olds showed greater concern for (i.e., care, fairness, authority, sanctity), they seemed to punish violations at relatively high rates. This is consistent with other moral research; Smetana et al. (2012) found that 2- to 4-year-olds rated moral violations as quite deserving of punishment, averaging approximately a 2.5/3 on their punishment scale. Therefore, it appears that when children think something is bad, they tend to believe it should be punished – which is perhaps consistent with the

previous discussion of dichotomous thinking styles.

Limitations and future directions

The limitations of the measures and analyses utilized in this chapter have been described throughout this discussion, such as the use of Likert scales with young children. However, there are still some that have not yet been addressed; most notably, we have not discussed the poor internal consistency shown for certain MFQK subscales. Fairness, loyalty, and odd behaviours showed internal consistencies ranging from .41 to .57. Interestingly, these were the same three subscales discussed above as having either items that were rated significantly different by participants (loyalty, odd behaviours), or items that, while not significantly different from others, did not appear to match how other items were rated (fairness). Possible reasons for why these items were rated differently have already been discussed, which included both poor wording and confounds with other foundations. Combined with the poor internal consistencies, these item analyses highlight that the results found for the fairness, loyalty, and odd behaviours scales should be interpreted with great caution. We hypothesize that, were these scales improved, fairness badness ratings would have been higher, and badness ratings for loyalty and odd behaviours would have been lower; however, it is impossible to say for certain. Overall, the overarching limitation of this chapter is the reliability and validity of the MFQK; further modification and more stringent psychometric testing appears to be required. That said, it was still a useful and interesting measure for a preliminary investigation into a broader range of moral concerns in 4-year-olds.

In addition to resolving potential measurement concerns, the findings presented in this chapter inspire numerous other interesting ideas for future research. The MFQK

allowed for a broader investigation of the types of moral concerns held by young children. However, both breadth and depth are needed to properly examine what, when, and why children are sensitive to particular concerns. For example, within the fairness foundation, it would be interesting to examine children's understanding of different concepts (e.g., equity, free-riding), how they judge such behaviours in personal versus third party interactions, or their emotions and actions after encountering fairness violations, as well as how each of these things interact and relate to each other. As highlighted by concerns with common method variance in the current chapter, the use of different methodologies is recommended; possibilities include real-life peer evaluations, behavioural observations, or observed emotional reactions. Emotions have been addressed very little in the current moral development literature, owing to a history of approaching moral development research from a rationalist, cognitive perspective. However, as highlighted by the current chapter, the study of moral emotions represents a particularly interesting area of future investigation.

In conclusion, the purpose of the current chapter was to conduct a broader study of 4-year-olds' moral beliefs, specifically in the areas of care, fairness, loyalty, authority, and sanctity. Children's reactions to social norm violations were also examined as a comparison for how 4-year-olds judge violations in non-moral areas. Although the reliability and validity of certain subscales was called into question, our findings still suggested that 4-year-olds from a Western, liberal background are sensitive to different violations of care, fairness, authority, and sanctity, but less sensitive to violations of loyalty and social norms. The observed distinction between authority, sanctity, and social norm violations lends further support to the notion that authority and sanctity are indeed

moral issues, as proposed by MFT (Haidt, 2012). Further discussion of what these results mean for MFT and moral development literature is addressed in Chapter 6.

Chapter 5: Parental Influences on Morality

The final goal of the current dissertation was to examine the tenet of cultural learning in Moral Foundations Theory (MFT). Although MFT is in part a nativist and evolutionary theory, it also highlights how morality is impacted by environmental influences (Haidt & Joseph, 2008). Through our experiences and culture, innate moral foundations can become expanded, refined, or even reduced in importance. As children develop, they encounter a variety of different events and activities that have the potential to influence their moral beliefs and virtues. Teachers, peers, experiences, television, advertisements, and more could potentially impact the way children think about a certain topic. In early childhood, arguably one of the biggest sources of environmental influence are the people with whom children spend the most time – their parents. Therefore, the current chapter focuses on parents' concern for moral foundations and their correspondence with 2- and 4-year-olds' responses to the Moral Foundations Questionnaire for Kids (MFQK) and Moral Foundations Puppet Task (MFPT).

The pioneers of moral development research, Jean Piaget (1977) and Lawrence Kohlberg (1963), viewed parents as largely inconsequential to moral development. Instead, peers were seen as more important, as they provided experiences and dilemmas through which children could rationally self-construct and develop their moral reasoning. Social Domain Theory (SDT), which extended and refined the work of Lawrence Kohlberg, also viewed morality as originating primarily through interactions with peers and siblings. However, according to SDT, parents can add to the lessons that children learn from these interactions, as outlined by Smetana (1999). Specifically, Smetana (1999) highlighted that parents are important to moral development in three ways. Firstly,

they help children make sense of their interactions and morality by forbidding actions, giving instructions, and responding to moral violations. Secondly, they provide emotional bonds and moral support to children as they develop. Finally, parents can provide developmentally appropriate reasoning and explanations. However, parents were not viewed as sources of moral beliefs or virtues in SDT. The influence of these three theories of moral development – the stage theories of Piaget and Kohlberg, and SDT – can be seen in the literature, as there is only a small amount of research that examines the impact of parents on moral development.

Of the research on parents that exists, the majority examines how parenting styles and interactions influence children's ability to reason about moral situations. Walker and Taylor (1991) rejected the notion of cognitive developmentalists that parents cannot influence moral development because they are in positions of authority. Instead, they argued that parents likely care more about their child's moral development relative to peers, teachers, or other persons. To study the role that parents play in moral development, they administered Kohlberg's moral scenarios individually to parents and school-aged children to assess their moral reasoning. They also conducted a family session, and coded the discussions into different categories (e.g., supportive, cognitively interfering, conflictual, informative, etc.). There was no relationship between parents' and children's levels of moral reasoning. However, Walker and Taylor (1991) did find that parents' discussion styles predicted moral development two years later. Representational (e.g., paraphrasing, clarifying) and supportive styles predicted greater development in moral reasoning. In contrast, operational (directly challenging or critiquing) and information (providing opinions) styles predicted less moral development.

Further research by Walker and colleagues examined other aspects or comparisons of parents' influence on moral development. Walker and Hennig (1999) used similar measures as Walker and Taylor (1991) to again study the influence of parents' interaction styles, but also measured parents' ego functioning. They found that parents who had poor ego functioning (e.g., defensive, insensitive, poor emotional expression) hindered the development of children's moral reasoning – in addition to parents that were highly opinionated, hostile, critical, or interfering, as previously found. Parents who scaffolded development and asked appropriate questions, and who provided appropriate emotional responsiveness, were more likely to facilitate moral reasoning. Therefore, this study highlighted the importance of parents, but also of emotions in moral reasoning.

Walker et al. (2000) compared the influence of parent and peer interaction styles on children's moral development. Interestingly, they found that interfering (i.e., more confrontational) peer interactions predicted higher moral reasoning, which is the opposite of what was found in parents. They hypothesized this was due children having a more egalitarian relationship with peers. Overall, the research of Walker and colleagues contradicted cognitive theories of moral development and highlighted the importance of parental influence on moral development in school-aged children.

More recent studies have expanded the research on parenting styles and moral development. In general, parents who are more controlling, give harsher discipline, and show poor behaviour monitoring have children (school-aged or early adolescence) who are less likely to internalize 'positive' moral values (i.e., caring, fairness; Campaert et al., 2018; Karmakar, 2015). In contrast, positive parent-child relationships in elementary

school students are related to higher scores for children's moral self (less preference for immoral behaviour; Sengsavang & Krettenauer, 2015). Mutually responsive orientations between parents and children have been particularly highlighted as important for early conscience development (Kochanska, 2002). That is, children who have parents that respond to their needs and show positive affect are more likely to embrace parental influence and adopt parental values and behaviours. The dynamics between family members may be important too; Groenendyk and Volling (2007) found that greater observed co-parenting behaviours were related to early conscience development in toddlers and preschoolers.

In general, there is evidence to suggest that parents impact their children's moral development, especially through their interaction styles and emotional responsiveness. However, this research is limited in numerous ways. First, it has focused exclusively on children's reasoning about care and fairness. Specifically, almost all studies used the following approach: the researchers decided what was less vs. more moral (where greater concern for care and fairness was more moral), and then examined parents' influence on children's 'level' of morality. This approach ignores other areas in which there is evidence of moral concerns, such as authority, loyalty, and sanctity. Furthermore, it fails to consider that parents with different styles may have different values and might transmit those values to their children. For example, a parent with a more controlling style may have strong authority values and instill those values in their children. Overall, there is little research examining the different moral beliefs that parents hold and how they correspond to children's beliefs – excepting one study which found that children's reactions to prosocial behaviours were influenced by parents' values of justice and

fairness (Cowell & Decety, 2015). Therefore, while existing research loosely supports MFT's notion of cultural learning (i.e., parents influence moral development), more research is needed to examine a broader range of moral beliefs, and to examine correspondence between parents and children using a non-prescriptive, non-judgmental approach.

The primary goal of this chapter was to examine the correspondence between parents' concern for moral foundations and children's concern for moral foundations. Specifically, we examined relationships between parents' responses on the Moral Foundations Questionnaire (MFQ), and their children's responses on the Moral Foundations Questionnaire for Kids (MFQK) and Moral Foundations Puppet Task (MFPT). In addition, we used our parent MFQ data to examine previous findings that have been shown using adult samples and the MFQ. Our hypotheses were as follows:

H1: We expected to see significant positive relationships between parents' care, fairness, and loyalty scores on the MFQ, and their children's care, fairness, and loyalty scores on the MFQK and MFPT, respectively.

Justification: For the care and fairness foundations, previous studies have shown that parents' virtues in these areas influence children's related behaviours (Cowell & Decety, 2015), and that both young children and adults from Western, liberal samples (similar to our sample demographic) highly value care and fairness (Graham et al., 2011, Helwig et al., 2001). For the loyalty foundation, there is no relevant literature to rely on. However, we thought it would be reasonable to expect a positive correlation between parents' and children's loyalty scores. Of note, no specific predictions were made for the authority or sanctity foundations. Although these relationships were still of interest, we

thought that, for developmental reasons, we might not see relationships between children's and parents' scores (and a conclusion of 'no relationship' cannot be tested with hypothesis testing). For authority, younger children are often taught the importance of listening and obeying, even if their parents do not place a lot of value on authority. Similarly, a lot of time is spent teaching toddlers and preschoolers about disgust and contamination (e.g., toileting), even though parents may not value adult forms of sanctity. Furthermore, the sanctity violations in the child measures focused largely on physical disgust, while adult measures focused more on chastity or religion. Therefore, we did not expect to see correlations between children's and parents' scores on these foundations.

H2: We expected to replicate relationships previously found between political orientation and moral beliefs using the MFQ. Specifically, in our sample of predominantly western, white, and liberal adults, we predicted that they would score significantly higher on the care and fairness foundations, and lower on the loyalty, authority, and sanctity foundations. Furthermore, we expected that self-identified liberal parents would be more sensitive to care and fairness than self-identified conservative parents, while conservative parents would show higher concern for the loyalty, authority, and sanctity foundations than liberal parents.

H3: We also expected to replicate previous correlational patterns found between moral foundations using the MFQ. Specifically, we predicted that care and fairness (the individualizing foundations) would show medium to large (i.e., .3 or above) significant positive correlations with each other, as would the three binding foundations (loyalty, authority, sanctity). Other correlations were expected to be positive, but small (less than .3).

Justification: Hypotheses 2 and 3 are based the number of existing studies that have used the MFQ. This research has shown that liberals place more emphasis on care and fairness, while conservatives place relatively more emphasis on loyalty, authority, and sanctity (although they still value care and fairness, e.g., Graham et al., 2011). Furthermore, stronger correlations have been found between the individualizing and binding foundations than between other combinations of correlations (Graham et al., 2011).

Methods

The primary measure of interest for this chapter and analysis was the Moral Foundations Questionnaire (MFQ; Graham et al. 2011), described in detail in Chapter 2. As a brief reminder, 178 parents of 2- and 4-year-olds completed this measure.
The MFQ is a 32-item questionnaire that assesses adults' concern for the five moral foundations of care, fairness, loyalty, authority, and sanctity. It yields total scores ranging from 0 to 5, with greater scores indicating greater concern for each foundation. Four parents did not see or complete the second page of the MFQ, meaning we were unable to compute summary scores for these parents. Therefore, the final parent sample is 174. Parents' responses to the political orientation question (categorical variable) on the demographics questionnaire (see Chapter 2) were also used in this chapter.

In addition to parent measures, child measures were used to examine the correspondence between parents' and children's morals. Four-year-olds' responses (N = 76) to the Moral Foundations Questionnaire for Kids (MFQK, see Chapter 2) were used. This questionnaire yields summary scores that represent children's badness ratings of violations in each moral foundation (care, fairness, loyalty, authority, and sanctity), as

well as social norms violations (odd behaviours). Summary scores ranged from 0-5, where greater scores represent greater concern for a particular foundation. In addition to the MFQK, data from the Moral Foundations Puppet Task (MFPT, described in Chapters 2 and 3) was used. Seventy-six four-year-olds and 102 two-year-olds completed this measure. The summary variables created for each of the five moral foundations were used. These were ordinal level variables, ranging from 0-4, where higher scores indicated greater concern for the relevant moral foundation.

Data Analysis

Although the primary hypotheses and analyses of interest in the current chapter involved the relationships between parents' and children's morals, we first conducted the analyses examining solely the MFQ data (H2-H3). This allowed us to examine and understand responses to the MFQ in our sample of parents in detail, prior to using them in more complicated predictive analyses. Hypothesis 2 concerned relationships between political orientation and moral beliefs and was tested through two separate analyses. First, considering the demographics of our sample of parents (i.e., predominantly liberal), we examined their overall pattern of moral beliefs. We considered the distribution and descriptive statistics of each foundation (as measured by the MFQ). We then conducted a repeated measures ANOVA to test for significant differences between foundations, with post-hoc pairwise comparisons. Mauchly's test of sphericity and standardized residuals were used to test assumptions of the ANOVA.

In addition to examining the overall sample, we also tested Hypothesis 2 by directly comparing the moral concerns of parents (as measured by the MFQ) based on their political orientation. On our demographics questionnaire, parents self-identified

their political orientation across a range of possible options. Based on their responses, we grouped parents into 'more liberal' (liberal, socialist, or environmentalist, $N = 108$) and 'more conservative' (moderate, libertarian, or conservative, $N = 40$) categories. Responses of 'other' ($N = 23$) were coded as missing data, in addition to the 5 parents who chose not to respond. Given the difference in sample size between the two groups, we conducted a series of Welch's t-tests (does not assume equal variances) to compare scores on each of the moral foundations between the two groups.

Hypothesis 3 concerned correlations between parents' moral foundation scores on the MFQ; we predicted that correlations between individualizing and binding foundations would be medium to large, while all others would be small. Therefore, to examine these predictions, we conducted Pearson correlations between all possible combinations of the five moral foundations. As in the previous two chapters, the Benjamini and Hochberg (1995) procedure was used to control for false discovery rates across multiple comparisons (see Chapter 3 for a complete description), for this and the previous analyses. Therefore, the corrected significance thresholds reported in the results section were calculated using this procedure.

Finally, the last group of analyses tested Hypothesis 1 and examined the relationships between parents' and children's moral concerns. We conducted a series of generalized linear models to examine how parents' scores on the MFQ separately predicted children's scores on the MFPT (2- and 4-year-olds) and MFQK (4-year-olds only). Of note, we only made predictions regarding relationships between the same moral foundations (e.g., parents' care scores with children's care scores, parents' fairness scores with children's fairness scores). However, in each model conducted, we included all

parent variables as predictors (i.e., parents' scores on all 5 foundations) for exploratory purposes, and also for predictive ability; it was possible that one parent variable alone would not predict children's responses, but multiple variables would. Similarly, although we only made predictions that involved children's responses to care, fairness, and loyalty, we examined parents' influence on children's responses to all moral foundations again for exploratory purposes, and for the sake of completion. In all generalized linear models conducted, children's moral foundation scores, as measured using either the MFPT or MFQK, were treated as the dependent variables. Therefore, for both the MFQK and MFPT, there were 5 separate models ran. That is, one model for children's scores for each foundation on each measure was run, with 10 models run in total. A summary of the models that were run is presented in Table 17.

For the MFQK, children's moral foundation summary scores on the MFQK (i.e., the dependent variables) showed skewed distributions (see Chapter 4, Figure 2). Therefore, for each of the 5 generalized linear models that were run, we compared gamma distributions (with a log link) and normal distributions (with an identity link) to examine which best represented the underlying distribution. However, because the MFQK variables were negatively skewed, and gamma distributions are positively skewed, each variable first had to be transformed. The MFQK subscale scores (for the care, fairness, loyalty, authority, and sanctity foundations) of each participant were multiplied by -1, added to the maximum possible value (5), and then added again by +1 (as all values for gamma distributions must be greater than zero). That is, the transformation employed the following formula: $(-1)x+5+1$, where x = participants' scores on a particular foundation. This created positively skewed continuous variables

ranging from 1-6. This reversed the interpretation for all MFQK variables; the transformed variables were labelled as disregard for care, disregard for fairness, disregard for loyalty, disregard for authority, and disregard for sanctity, representing the opposite of concern for care, fairness, loyalty, authority, and sanctity, respectively. For model comparison and selection, we examined the distributions of deviance residuals, as well as Bayesian Information Criterions (BICs). All models were conducted using the Huber-White Sandwich Estimator; this provides a more robust estimation as it does not assume homogeneity of variance. For each model we have reported the omnibus test, as well as main effects. McFadden's pseudo R squared was also calculated.

While only 4-year-olds responded to the MFQK, 2- and 4-year-olds responded to the MFPT. Therefore, for the 5 models that were run examining parents' influence on children's MFPT scores (one for each foundation), we entered age group (2 vs. 4) as a predictor, in addition to the 5 parent variables (care, fairness, loyalty, authority, sanctity). Of note, we did not enter interactions with age into these models; this produced rather unwieldy models, and we had no prior hypotheses regarding the impact of parents' influence according to age. As described in Chapters 2 and 3, moral foundation summary scores on the MFPT were ordinal-level variables. Therefore, for each generalized linear model, we compared an ordinal logistic (multinomial distribution with cumulative logit link) model with a normal distribution (identity link) model. Of note, the regular, untransformed MFPT variables were used. As ordinal logistic models do not produce deviance residuals, only the Bayesian Information Criterions (BICs) were used for model selection. As with the MFQK models, the Huber-White Sandwich Estimator was used, and we have reported the omnibus test, main effects, and McFadden's pseudo R squared.

Table 17

A summary of the generalized linear models used to examine parents' influence on children's moral beliefs

#	Measure for DV	Transf.?	DV	# of predict.	Predictors	Distributions compared
1	MFQK	Yes	Child disreg. for care	5	Parents' MFQ scores (care, fairness, loyalty, authority, sanctity)	Gamma vs. normal
2	MFQK	Yes	Child disreg. for fairness	5		
3	MFQK	Yes	Child disreg. for loyalty	5		
4	MFQK	Yes	Child disreg. for authority	5		
5	MFQK	Yes	Child disreg. for sanctity	5		
6	MFPT	No	Child care	6	Age group & parents' MFQ scores (care, fairness, loyalty, authority, sanctity)	Ordinal vs. normal
7	MFPT	No	Child fairness	6		
8	MFPT	No	Child loyalty	6		
9	MFPT	No	Child authority	6		
10	MFPT	No	Child sanctity	6		

Note. MFQK = Moral Foundations Questionnaire for Kids; MFPT = Moral Foundations Puppet Task; MFQ = Moral Foundations Questionnaire; transf. = transformation; DV = dependent variable, predict. = predictors, disreg. = disregard. Transformation of original MFQK variables (which created the disregard for [insert foundation] variables) was (-1)x+5+1. Full description of distributions compared: gamma with a log link, normal with an identity link, and ordinal logistic (multinomial distribution with cumulative logit link).

Results

Comparing parents' moral concerns on the MFQ (H2)

The distribution of parent/guardians' scores for each moral foundation on the MFQ are shown in Figure 6. All variables appeared to be largely normally distributed, although with varying amounts of spread. The care and fairness distributions were quite narrow and concentrated at higher scores. The care distribution showed some evidence of a ceiling effect, although the frequency of scores did still taper at the upper end. The loyalty, authority, and sanctity foundations showed more variability in responses, with scores spread out across the distributions, and peaks near the middle.

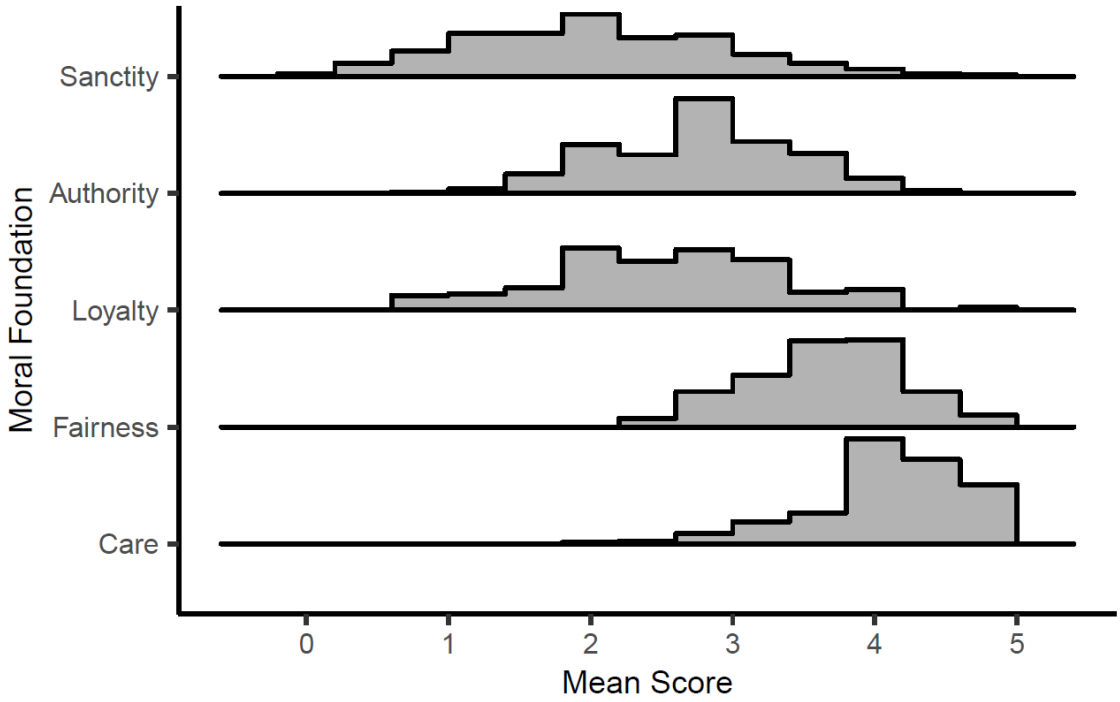


Figure 6. Distributions of parents' scores on the Moral Foundations Questionnaire (MFQ) according to foundation. Higher scores represent greater concern for a particular foundation. Height of bars indicate relative frequency of scores. $N = 174$.

To test our hypothesis (H2) that our predominantly liberal sample would be more sensitive to care and fairness, and less sensitive to the other three foundations, we compared parents' moral foundation scores on the MFQ. Mauchly's test of sphericity revealed that the assumption of sphericity was not met across the moral foundations, Mauchly's W = .603, p <.001. Therefore, the Greenhouse-Geisser correction was used to conduct a repeated-measures ANOVA. Visual examination of standardized residuals for each foundation suggested that the assumption of normality was met. The analysis revealed that parents' summary scores differed significantly across the different moral foundations, $F(3.14, 543.08) = 363.23$, p<.001. To examine differences between specific foundations, we conducted follow-up pairwise comparisons (no change to significance threshold of .05). All foundations were significantly different from all other foundations (all p's <.001). As shown in Figure 6, the parents in our sample were most sensitive to

the care foundation ($M = 4.10$, $SD = 0.56$), followed by fairness ($M = 3.64$, $SD = 0.55$), authority ($M = 2.77$, $SD = 0.66$), loyalty ($M = 2.54$, $SD = 0.85$), and sanctity ($M = 2.01$, $SD = 0.88$).

Comparing the moral beliefs of liberals and conservatives (H2)

In addition to comparing moral foundations scores across the entire sample, we also tested Hypothesis 2 by examining differences in moral concerns according to political orientations. The average scores for each moral foundation on the MFQ for the 'more liberal' versus 'more conservative' parent participants are shown in Figure 7.

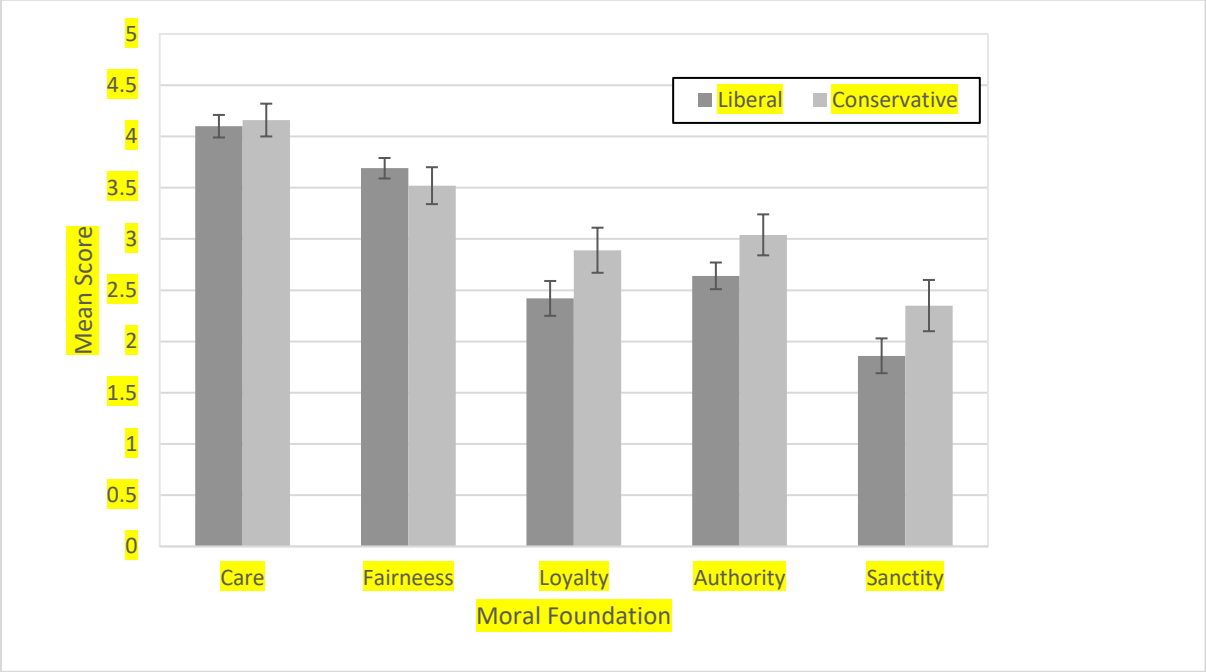


Figure 7. Average moral foundation scores for parent participants on the Moral Foundations Questionnaire, according to political orientation. Parents were grouped into the categories of 'more liberal' (liberal, socialist, or environmentalist, N = 108) and 'more conservative' (moderate or conservative, N = 40). Error bars represent 95% confidence intervals.

We conducted a series of Welch's t-tests between each of the moral foundations (corrected significance threshold = .03). Liberals and conservatives showed no significant differences in their concern for the care ($t(77.53) = -0.56$, $p = .574$, $d = 0.10$) and fairness ($t(68.08) = 1.67$, $p = .100$, $d = 0.31$) foundations. However, relative to liberals,

conservatives scored significantly higher (and showed greater concern for) the loyalty ($t(91.11) = -3.46$, $p = .001$, $d = 0.60$), authority ($t(77.19) = -3.36$, $p = .001$, $d = 0.62$), and sanctity ($t(76.64) = -3.25$, $p = .002$, $d = 0.59$) foundations.

Correlations Between Moral Foundations (H3)

To test Hypothesis 3, we conducted Pearson correlations between parents' scores for the five moral foundations (corrected significance threshold = .035). Results are shown in Table 18. Significant, positive relationships were seen between a number of foundations: care and fairness, care and loyalty, care and sanctity, fairness and loyalty, loyalty and authority, loyalty and sanctity, and authority and sanctity. For these pairings, as concern for one foundation increased, so did concern for the other. No significant correlations were found between care and authority, fairness and authority, or fairness and sanctity.

Table 18

Correlations between parents' moral foundation scores on the Moral Foundations Questionnaire

Foundation	Care	Fairness	Loyalty	Authority	Sanctity
Care	1				
Fairness	.44**	1			
Loyalty	.23*	.23*	1		
Authority	.09	.12	.55**	1	
Sanctity	.20*	.16	.55**	.60**	1

Note. *p<.01, **p<.001. Following application of the Benjamini and Hochberg (1995) procedure, corrected significance threshold was .035.

In Hypothesis 3, we specifically predicted that that care and fairness (the individualizing foundations) would show medium to large (i.e., .3 or above) positive correlations with each other, as would the three binding foundations (loyalty, authority, sanctity) – while other correlations were expected to be small (but positive). This is indeed the pattern we saw; the positive correlations found between the loyalty, authority,

and sanctity foundations were large (above .5), and the positive correlation between the care and fairness foundations was in the upper end of the medium range (above .4). All other correlations were positive but small (.23 or smaller).

Parents' Influence on Children's Morals (H1)

The primary purpose of this chapter was to examine the impact of parents' moral beliefs and virtues on their children's moral concerns (Hypothesis 1). As described above, we conducted a series of generalized linear models to examine the influence of these predictors on children's concern for moral foundations on both the Moral Foundations Questionnaire for Kids (MFQK) and Moral Foundations Puppet Task (MFPT).

Moral Foundations Questionnaire for Kids (MFQK)

Each of the moral foundations on the MFQK was examined using a separate generalized linear model. As a reminder, the MFQK was only completed by 4-year-olds, therefore this analysis only pertains to the correspondence between parents' morals and 4-year-olds' moral concerns. For each generalized linear model, we compared a gamma distribution (with a log link) and a normal distribution (with an identity link) to see which model best represented the underlying distribution. To test the gamma distributions, MFQK variables first had to be transformed (see Data Analysis). Transformations yielded positively skewed continuous variables ranging from 1-6. They were labelled as disregard for care, disregard for fairness, disregard for loyalty, disregard for authority, and disregard for sanctity, representing the opposite of concern for care, fairness, loyalty, authority, and sanctity, respectively. Regardless of model selection, these variables are used for each MFQK model to maintain consistency and ease of interpretation.

To select the best model for each MFQK dependent variables, we compared the deviance residuals and Bayesian Information Criterions (BICs) for the normal and gamma distributions. The distribution of deviance residuals for each model and foundation are shown in Figure 8. Across all foundations, the distributions of deviance residuals were difficult to compare between the normal and gamma distributions, as they showed rather large differences in dispersion. The deviance residuals for the gamma distributions showed a much narrower range than those for the normal distributions. If we ignore range and consider the shape of each distribution, all do show peaks in the middle with less frequency of scores at either end. However, none are perfect bell curves. Therefore, we relied heavily on the BICs for model selection (see Table 19).

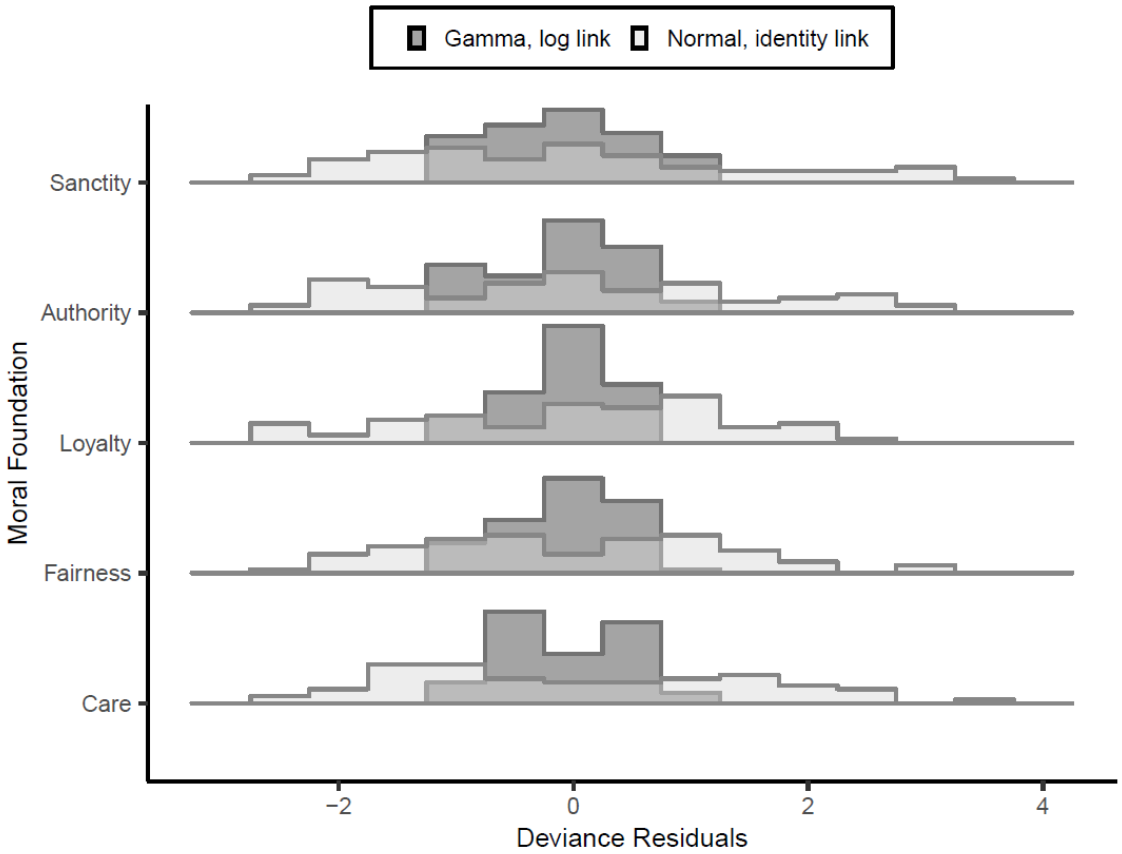


Figure 8. Distributions of deviance residuals from generalized linear models of 4-year-olds' responses to the Moral Foundations Questionnaire for Kids. A separate model was run for each foundation; models tested included a gamma distribution with a log link, and a normal distribution with an identity link. Height of bars indicate relative frequency of scores.

According to criteria put forth by (Raftery, 1995), there was strong to very strong evidence that a gamma distribution was the best model for the disregard for care, disregard for authority, and disregard for sanctity variables. There was also very strong evidence that a normal distribution was the best model for the disregard for loyalty variable. For disregard for fairness, the BICs showed weak evidence that a gamma distribution was the better model. Therefore, it made more sense to apply the normal distribution with an identity link, as it was the most simple and parsimonious model.

Table 19

Bayesian Information Criterions (BIC) for generalized linear model distributions of 4-year-olds' scores on the Moral Foundations Questionnaire for Kids

MFQK dependent	Normal distribution, identity link	Gamma distribution, log link	Model selected
Disregard for care	287.96	269.57	Gamma
Disregard for fairness	250.77	248.31	Normal
Disregard for loyalty	246.73	254.93	Normal
Disregard for authority	277.13	267.65	Gamma
Disregard for sanctity	275.48	260.37	Gamma

Results from each generalized linear model conducted are shown in Table 20. For each model, a different MFQK subscale (e.g., disregard for care, disregard for fairness, etc.) served as the dependent variable, while parents' MFQ scores for care, fairness, loyalty, authority, and sanctity were entered as predictors. For disregard for care, the omnibus test indicated that the combination of these five predictors was a better model than a model with just the intercept. Testing of main effects revealed that when controlling for the effects of all other predictors, the main effect of parents' concern for care, fairness, and sanctity on children's disregard for care were not significant. There was a small, significant main effect of loyalty; as parents' loyalty scores increased, children's disregard for care decreased (or, as parents' loyalty scores increased,

children's concern for care increased). There was also a medium, significant main effect of authority; as parents' authority scores increased, children's disregard for care increased (or, as parents' authority scores increased, children's concern for care decreased).

Table 20

Results of Generalized Linear Modelling to examine the impact of parents' moral beliefs on 4-year-olds' moral concerns, as measured by the Moral Foundations Questionnaire for Kids

Overall Model						Main Effects					
DV	Model	*N*	*χ2*	*p*	R^{2*}	Parent predictors	*β*	*SE*	95% CI	Wald *χ2*	*p*
Child disregard for care	Gamma log link	72	13.97	.016	.079	Care	-0.09	.12	[-0.32, 0.14]	0.62	.431
						Fairness	0.17	.12	[-0.07, 0.41]	1.88	.170
						Loyalty	-0.19	.09	[-0.37, -0.10]	4.27	.039
						Authority	0.42	.13	[0.17, 0.67]	10.66	.001
						Sanctity	-0.04	.13	[-0.29, 0.20]	0.12	.728
Child disregard for fairness	Normal identity link	67	4.66	.458	.054	Care	-0.20	.25	[-30, 0.70]	0.62	.431
						Fairness	-0.16	.27	[-0.67, 0.37]	0.33	.564
						Loyalty	-0.06	.25	[-0.55, 0.42]	0.06	.801
						Authority	0.28	.31	[-0.32, 0.88]	0.84	.359
						Sanctity	0.18	.22	[-0.25, 0.61]	0.64	.422
Child disregard for loyalty	Normal identity link	65	4.85	.435	.044	Care	-0.01	.28	[-0.55, 0.53]	0.00	.967
						Fairness	0.26	.29	[-0.31, 0.84]	0.81	.370
						Loyalty	0.25	.23	[-0.19, 0.70]	1.24	.266
						Authority	0.32	.31	[-0.29, 0.93]	1.07	.301
						Sanctity	-0.17	.25	[-0.67, 0.33]	0.47	.494
Child disregard for authority	Gamma log link	69	10.58	.060	.067	Care	-0.06	.12	[-0.30, 0.17]	0.28	.599
						Fairness	0.19	.10	[-0.01, 0.40]	3.38	.066
						Loyalty	-0.11	.09	[-0.29, 0.08]	1.27	.260
						Authority	0.19	.11	[-0.03, 0.42]	2.93	.087
						Sanctity	0.13	.10	[-0.07, 0.32]	1.70	.193
Child disregard for sanctity	Gamma log link	66	6.78	.238	.055	Care	0.02	.12	[-0.22, 0.25]	0.02	.878
						Fairness	0.21	.16	[-0.10, 0.53]	1.74	.188
						Loyalty	-0.00	.11	[-0.23, 0.22]	0.00	.974
						Authority	0.30	.14	[0.04, 0.57]	4.93	.026
						Sanctity	-0.15	.11	[-0.36, 0.06]	1.97	.160

Note. R^2 is McFadden's Pseudo R squared. Parents' moral foundations were measured using the Moral Foundations Questionnaire (MFQ).

For the remaining foundations (disregard for fairness, loyalty, authority, and

sanctity), the omnibus tests revealed that the combination of the five parent predictors were no better than models with just the intercept; main effects are presented solely for the sake of completeness and to allow for potential examination of trends.

Moral Foundations Puppet Task (MFPT)

Each of the moral foundations on the MFPT was also examined using a separate generalized linear model. As a reminder, the MFPT was completed by both 2- and 4-year-olds. Parents' scores on the MFQ were again entered as predictors, as well as age group. For each generalized linear model, we compared an ordinal logistic model with a normal distribution model. Of note, the regular, untransformed MFPT variables were used (where higher scores indicate greater concern for the relevant foundation). To select the best model for each MFPT dependent variable, we compared the Bayesian Information Criterions (BICs) for the normal and ordinal logistic distributions (see Table 21). According to criteria put forth by Raftery (1995), there was very strong evidence that an ordinal logistic distribution was the best model for all MFPT variables.

Table 21

Bayesian Information Criterions (BIC) for generalized linear model distributions of summary scores on the Moral Foundations Puppet Task

MFPT dependent	Normal distribution, identity link	Ordinal logistic	Model selected
Care	579.60	490.16	Ordinal
Fairness	585.57	487.24	Ordinal
Loyalty	568.64	454.11	Ordinal
Authority	575.35	484.24	Ordinal
Sanctity	532.27	475.27	Ordinal

Note. Ordinal logistic refers to a multinomial distribution with a cumulative logit link.

Results from each generalized linear model selected are shown in Table 22. For each model, a different MFPT summary score (e.g., care, fairness, etc.) served as the dependent variable, while age group (coded as 0 = 2-year-olds, 1 = 4-year-olds) and

parents' MFQ scores for care, fairness, loyalty, authority, and sanctity were entered as predictors. Of note, exponentiated coefficients are displayed, which represent odds ratios. For continuous predictors, values less than 1 indicate a negative relationship between the predictor and outcome, while values greater than 1 indicate a positive relationship. For the age group predictor, values less than 1 indicate that 4-year-olds had higher scores on the outcome variable, and values greater than 1 indicate that 4-year-olds had lower scores on the outcome variable.

For the care foundation, the omnibus test indicated that the combination of these six predictors was a better model than a model with just the intercept. When controlling for the effects of all other predictors, the main effects of parents' concern for fairness, loyalty, authority, and sanctity on children's concern for care were not significant. There was a small, significant main effect of care; as parents' care scores increased, children's concern for care decreased. There was also a medium, significant effect of age group; as shown in Chapter 3, 4-year-olds showed greater concern for care than 2-year-olds.

For the fairness foundation, the omnibus test also indicated that the combination of the six predictors entered was a better model than a model with just the intercept. When controlling for the effects of all other predictors, the main effects of parents' concern for care, fairness, and authority on children's concern for fairness were not significant. There was a small, significant main effect of loyalty; as parents' loyalty scores increased, children's fairness scores decreased. There was also a small, significant effect of sanctity; as parents' sanctity scores increased, children's concern for fairness also increased. Finally, there was again a medium, significant effect of age group; as shown in Chapter 3, 4-year-olds showed greater concern for fairness than 2-year-olds.

Table 22

Results of Generalized Linear Modelling[2] to examine the impact of parents' moral beliefs on 2- and 4-year-olds' moral concerns, measured by the Moral Foundations Puppet Task

Overall Model						Main Effects					
DV	Model	*N*	$\chi 2$	*p*	R^{2*}	Predictors	e^{β}	*SE*	95% CI	Wald $\chi 2$	*p*
Child care	Ordinal logistic	167	32.17	<.001	.23	Age	0.25	.30	[0.14, 0.45]	21.16	<.001
						Care	0.42	.31	[0.23, 0.76]	8.21	.004
						Fairness	1.22	.34	[0.62, 2.38]	0.32	.570
						Loyalty	0.79	.21	[0.52, 1.19]	1.28	.258
						Authority	1.03	.32	[0.55, 1.91]	0.01	.929
						Sanctity	0.97	.22	[0.63, 1.48]	0.02	.880
Child fairness	Ordinal logistic	166	40.54	<.001	.085	Age	0.18	.32	[0.09, 0.33]	29.23	<.001
						Care	0.68	.27	[0.40, 1.16]	2.00	.158
						Fairness	0.91	.28	[0.52, 1.58]	0.12	.735
						Loyalty	0.63	.23	[0.40, 0.98]	4.15	.042
						Authority	1.01	.27	[0.59, 1.71]	0.00	.982
						Sanctity	1.69	.22	[1.10, 2.60]	5.76	.016
Child loyalty	Ordinal logistic	154	5.10	.531	.013	Age	0.92	.33	[0.48, 1.77]	0.06	.806
						Care	0.59	.36	[0.29, 1.19]	2.20	.138
						Fairness	1.69	.32	[0.90, 3.16]	2.64	.104
						Loyalty	1.19	.22	[0.77, 1.84]	0.59	.442
						Authority	0.99	.34	[0.51, 1.94]	0.00	.982
						Sanctity	0.81	.31	[0.44, 1.48]	0.48	.489
Child authority	Ordinal logistic	167	37.98	<.001	.081	Age	0.17	.33	[0.09, 0.33]	28.31	<.001
						Care	0.88	.32	[0.47, 1.66]	0.15	.703
						Fairness	0.90	.33	[0.47, 1.72]	0.10	.757
						Loyalty	1.45	.21	[0.97, 2.18]	3.21	.073
						Authority	0.84	.30	[0.47, 1.50]	0.34	.560
						Sanctity	1.14	.22	[0.73, 1.77]	0.32	.571
Child sanctity	Ordinal logistic	167	50.50	<.001	.11	Age	0.11	.34	[0.06, 0.22]	40.40	<.001
						Care	0.81	.24	[0.51, 1.29]	0.78	.379
						Fairness	1.05	.30	[0.59, 1.90]	0.03	.863
						Loyalty	1.05	.26	[0.64, 1.73]	0.04	.849
						Authority	0.91	.32	[0.48, 1.71]	0.09	.762
						Sanctity	1.26	.22	[0.81, 1.95]	1.08	.300

Note. R^2 is McFadden's Pseudo R squared. Foundations listed in the 'predictors' column refer to parents' moral foundation scores on the Moral Foundations Questionnaire (MFQ). The age predictor variable was coded as 0 (2-year-olds) and 1 (4-year-olds). Exponentiated coefficients (e^{β}) represent odds ratios. For continuous predictors, values less than 1 indicate a negative relationship between the predictor and outcome, while values greater than 1 indicate a positive relationship. For the age group predictor, values less than 1 indicate that 4-year-olds had higher scores on the outcome variable, and values greater than 1 indicate that 4-year-olds had lower scores on the outcome variable.

[2]The analyses run in Table 22 were also conducted using only the children who passed the relevant manipulation checks on the MFPT. This led to a significant reduction in the sample size for each analysis; and meant that the models for children's concern for loyalty and authority did not converge or were invalid. For the remaining foundations, no changes to the pattern of results shown in Table 22 were observed.

For both the authority and sanctity foundations, the omnibus tests indicated that the six predictors entered were better models than models with just the intercept. However, testing of main effects revealed that when controlling for the effects of all other predictors, there were no significant relationships between any of the parent predictors and children's authority or sanctity scores. There was a significant effect of age group on both foundations; 4-year-olds again showed greater concern for authority and sanctity than 2-year-olds. These effects were medium and large for the authority and sanctity foundations, respectively.

Finally, for the loyalty foundation, the omnibus test indicated that the combination of the six predictors entered was no different than a model with just the intercept. This result is consistent with the absence of age group difference found in children's concern for loyalty on the MFPT in Chapter 2. Main effects are included in Table 22 for the sake of consistency and to allow for potential examination of trends.

Summary of Parents' Influence

A visual summary of the relationships between parents' and children's morals, on both the MFPT and MFQK, is shown in Figure 9. Relationships found for the transformed MFQK variables (e.g., disregard for care, etc.) have been reversed for ease of interpretation; they are shown as relationships with children's concern for (rather than disregard for) each moral foundation. In Hypothesis 1, we predicted significant positive relationships between parents' care, fairness, and loyalty scores on the MFQ, and their children's care, fairness, and loyalty scores on the MFQK and MFPT, respectively. Of these, the only significant relationship found was for the care foundation on the MFPT. However, it was unexpectedly negative; children's care scores were negatively related to

parents' concern for care. Other exploratory analyses revealed that children's MFQK care scores were negatively related to parents' concern for authority, and positively related to parents' concern for loyalty. For the fairness foundation, children's MFPT fairness scores were negatively related to parents' loyalty scores, and positively related to parents' sanctity scores. For the loyalty, authority, and sanctity foundations, there were no significant influences of parents' morals on children's responses to the MFQK or MFPT.

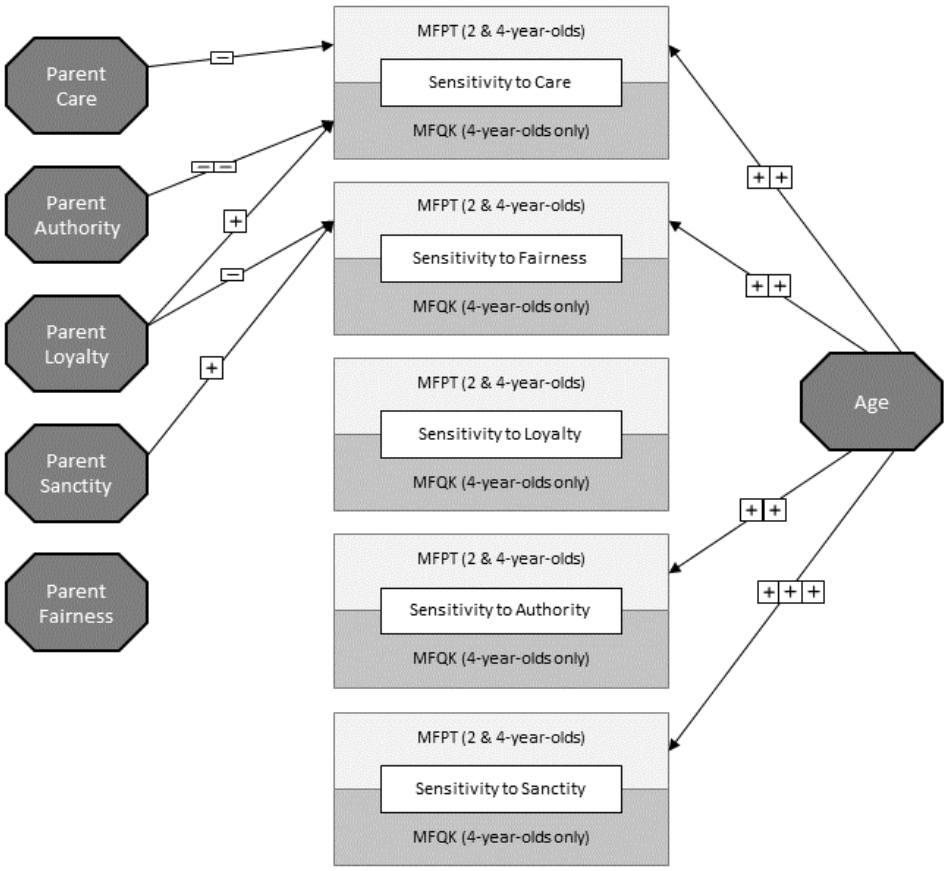


Figure 9. A summary of the results of generalized linear modelling to examine the relationships between parents' and children's moral concerns. Children's scores on moral foundation were measured using two different measures, the Moral Foundations Puppet Task (MFPT; completed by 2- and 4-year-olds) and the Moral Foundations Questionnaire for Kids (MFQK; completed by 4-year-olds only). For the MFPT, the influence of age group was also considered. Symbols represent direction and size of effect (small, medium, or large).

Discussion

The primary goal of the current chapter was to examine relationships between

parents' and children's moral foundations. Specifically, we analysed relationships between parents' responses on the Moral Foundations Questionnaire (MFQ), and their children's responses on the Moral Foundations Questionnaire for Kids (MFQK; 4-year-olds only), and the Moral Foundations Puppet Task (MFPT; 2- and 4-year-olds). We predicted that we would find significant, positive relationships between parents' scores and their children's' scores on the care foundation, parents' and children's' scores on the fairness foundation, and parents' and children's' scores on the loyalty foundation. Unexpectedly, we found a negative relationship between parents' care scores on the MFQ and children's' care scores on the MFPT, and no significant relationship between parents' care scores and children's' care scores on the MFQK. No significant relationships were found between parents' fairness scores and children's fairness scores, on either the MFQK or MFPT. However, children's' care and fairness scores did show some relationships with parents' scores on other moral foundations, such as authority and loyalty. The effect sizes for these relationships were predominantly small. No significant relationships were found between parents' scores on any foundations (care, fairness, loyalty, authority, or sanctity) and children's' loyalty, authority, or sanctity scores (on the MFQK or MFPT).

Relationships between children's and parents' moral beliefs

Taken together, the overarching pattern of the parent-child analyses was that very few relationships were found, and those that were found were often unexpected. We will first briefly discuss some possible interpretations of individual findings; however, the discussion will subsequently focus on interpretations and explanations for the general pattern. Starting with the care foundation (for children), we saw that parents' concern for

authority was negatively related to children's care scores on the MFQK. Interestingly, this fits with previous research showing that more punitive and authoritarian parenting is negatively related to child internalization of care and fairness (Campaert et al., 2018; Karmakar, 2015). We also saw that parents' loyalty scores were positively correlated with children's care scores on the MFQK. This was surprising, as the care and loyalty foundations are typically not strongly linked to each other (Graham et al., 2011). One possibility is that it speaks to the influence of family cohesion or structure, which have been found to relate to children's moral values (White & Matawie, 2004). However, the most surprising finding was that parents' care scores were negatively related to children's care scores on the MFPT. This was puzzling to interpret; one guess is that parents with very high care scores have stricter rules (e.g., regarding how to treat others), and so this functions similar to authoritarian parenting. However, it is also worth considering that, as 4-year-olds showed very little variability in their responses to the MFPT scenarios, this relationship was driven more by 2-year-olds – who may not yet have internalized parent morals (or understand their parents' beliefs), and whose responses were no different from chance (see Chapter 3). Furthermore, as discussed in Chapter 3, the MFPT produced dichotomous data and was likely less sensitive to degrees of concern for moral foundations.

It is likely also important to question the reliability of relationships with MFPT data when looking at the fairness foundation, where we saw that parents' loyalty and sanctity scores were negatively and positively related with children's MFPT fairness scores, respectively. Both relationships are difficult to explain, especially given the positive relationship found between parents' loyalty scores and children's care scores on

the MFQK. If these relationships were to hold up in future research, our best guesses would be that parents who strongly value loyalty are less likely to promote being fair to everyone (as they care more about ingroups), while parents who strongly value sanctity may encourage the importance or sacredness of sharing resources (although admittedly, both feel like a stretch). Regarding other foundations, no relationships were found between parents' morals and children's scores on the loyalty, authority, or sanctity foundations. Considering the pattern of results shown in previous chapters, where child participants were not sensitive to loyalty violations and their responses were largely random, it makes sense that we did not see significant, distinguishable relationships between children's loyalty scores and parents' moral beliefs. For the authority and sanctity foundations, we did not expect to find relationships between parents' and children's morals; as previously outlined, the salience of these foundations during the preschool years regardless of parents' values likely diminishes potential relationships. Of note, a non-significant result does not confirm that there are no relationships, but does lend some support to the suggestion.

Overall, as we previously noted, the analyses examining relationships between parents' and children's morals largely produced unexpected, non-significant, and difficult to explain results, suggesting that there is something else at play. One possible explanation is that there is less parental influence on moral beliefs during the toddler and preschool years relative to later in childhood. However, this would be largely inconsistent with developmental research as a whole; we know parent factors do impact child behaviours and feelings from an early age (e.g., Slatcher & Trentacosta, 2012). Furthermore, although previous research has not examined correspondence between

parents' and children's moral beliefs per se, there is evidence to suggest that parent factors influence morally relevant variables, such as guilt or correcting the behaviours of others (Kochanska, 2002). Therefore, alternatively, we believe the most likely explanation is that we are not seeing the full picture and are missing potential mediators or moderators.

Re-visiting literature on relationships between parents and their children provides some possible candidates for mediators and moderators. As highlighted at the beginning of this chapter, parents' interaction styles have been found to influence children's morals. Specifically, parents who are hostile, opinionated, controlling, and harsher disciplinarians (i.e., to summarize, those with an authoritarian parenting style) tend to have children with 'worse' moral reasoning or values (i.e., less concern for care and fairness; Campaert et al., 2018; Karmakar, 2015; Walker & Taylor, 1991). Leman (2005) further found that parents with authoritative (rather than authoritarian) parenting style facilitate 'good' moral development not because of the amount of moral justifications provided, but because of the types of justifications – suggesting that the types of rationales provided for moral beliefs are important for parent to child transmission. Similarly, Spinrad et al. (1999) found that parents' emotional reactions to stressful situations were related to children's moral behaviours, suggesting that how parents react to violations may influence the transmission of moral values. Relationships are important too; White and Matawie (2004) found that parent-teen cohesion moderated the influence of parents' moral values on adolescent moral values (parents with more cohesion had greater influence). This not only highlights the importance of good relationships, but also supports our hypothesis of mediating or moderating variables in parent-child transmission

of moral values.

In addition to parent or family factors, child factors are also likely to impact the extent to which they inherit or adopt their parents' moral virtues. For example, Slatcher and Trentacosta (2012) found that child negative emotionality (in 3- to 5-year-old children) moderated the relationship between parents' negative emotionality and children's problem behaviours. Similarly, Wachs et al. (2004) found that child temperament moderated the influence of environmental chaos on child compliance at daycares. In adults, attachment avoidance has been associated with less concern with harm and unfairness, and attachment anxiety has been associated with more concern for harm, unfairness, and impurity (Koleva et al., 2014). Therefore, it is also possible that different child attachment styles may make children more or less sensitive to messages about certain moral values. Overall, the previous research described at the beginning of this chapter demonstrated that parenting style was important for the development of 'good' moral virtues. In the current study, we took a less prescriptive or less judgmental approach, and instead simply attempted to examine correspondence of moral beliefs between parents and their children. However, in reality, it is probable that neither approach paints a complete picture. Instead, parent to child transmission of moral values likely follows a complex pattern and is influenced by beliefs, parenting styles, conversations, emotions, child temperament, self-regulation, and more.

The replication of previous adult research using the MFQ

While the primary goal of this chapter was to examine the correspondence between children's and parents' morals, given that we administered the MFQ to a large sample of adults, we also sought to replicate previous findings using this measure.

Previous research has consistently established that politically liberal adults score slightly higher than conservatives on care and fairness measures, while conservatives score significantly higher on loyalty, authority, and sanctity (e.g., Graham et al., 2011). The findings of the current study supported this phenomenon in two different ways. First, we saw that our sample, comprised predominantly of self-identified liberal participants, placed the most value on care and fairness, with significantly lower scores on loyalty, authority, and sanctity. Furthermore, even in our predominantly liberal sample, we saw differences in participants who identified as more liberal versus more conservative. Specifically, our participants who identified as more conservative scored significantly higher on loyalty, authority, and sanctity. Of note, we did not see differences in care and fairness scores, likely because our 'conservatives' included moderates who were not that different from our 'liberals.' This is reflected in the patterns shown in Figure 7 as well – both groups still showed more concern for care and fairness, and less to loyalty, authority, and sanctity. However, overall, the fact that we found differences even in our predominantly liberal sample highlights the robustness of differences in moral beliefs according to political orientation.

A few studies have examined possible underlying causes, predictors, or explanations for the moral divide between liberals and conservatives. Nilsson and Erlandsson (2015) found that the preference for equality mediated the relationship between individualizing foundations (care, fairness) and political identity, while resistance to change and system justification mediated the relationship between binding foundations (loyalty, authority, sanctity) and political identity. Similarly, Jost et al. (2003) found that conservatives resist change and justify inequality more, but also have

more anxiety from perceived threat and uncertainty. Therefore, they concluded that conservatives resist change and justify inequality *because* they feel more worried by uncertainty. Neuroimaging studies also support that conservatives are especially sensitive to threats (Rempala et al., 2016). Furthermore, the divide in moral beliefs between liberals and conservatives is not merely a descriptive or interesting phenomenon, it also predicts voting behaviour. Franks and Scherr (2015) found that participants' endorsement of moral foundations consistently predicted their voting behaviour in the 2012 U.S. Presidential Election, above and beyond other demographic variables. It would be interesting to examine the relationship between moral beliefs and voting behaviour in a Canadian sample, especially given the number of different parties that are strong contenders in Canadian elections.

The second area of adult research we sought to examine or replicate was the strength of correlations between individualizing and binding foundations. Previous research has found stronger correlations between the individualizing (care, fairness) and binding (authority, loyalty, sanctity) foundations than between other combinations of foundations. As discussed previously, the pattern shown in Table 18 certainly supports this effect. It was again interesting to see that this effect was found even in a predominantly liberal sample that placed less emphasis, on average, on the three binding foundations. In terms of what this means for MFT, some researchers have suggested that the five moral foundations could be reduced to only the individualizing/binding categories. In a Swedish sample, Nilsson and Erlandsson (2015) found that the advantage of a five-factor model was quite small over a two-factor model, and suggested it might be more useful and parsimonious to think of moral foundations as 2 categories. That said,

there are some critiques of the individualizing versus binding framework. For example, Janoff-Bulman and Carnes (2013) asserted that care and fairness can concern the collective, and referred to this as social justice – in contrast to social order, which represents the traditional 'binding' concerns about conformity and group homogeneity. Malka et al. (2016) asserted that collective care or fairness concerns do not invalidate the individualizing versus binding conceptualization, but do necessitate clarifying that 'binding' in MFT does not just mean concern with groups, but deals specifically with concern for social order. Overall, the primary message here is that relationships between individualizing and binding foundations appear to be robust, and these categories represent potential predictors in areas of future study.

Limitations and future directions

Both limitations and future research pertaining to this chapter's analysis have already been discussed, either here or in previous chapters (e.g., reliability and validity of MFQK data). Therefore, our discussion here will be brief. One limitation that has not been considered, and that pertains especially to this chapter, is the homogeneity of our sample. Our hope was that by recruiting a large number of participants through community daycares (i.e., rather than solely through our lab), we would see a diverse sample in terms of socioeconomic status, background, and political orientation. However, the sample was still largely one of convenience, and more targeted recruitment measures were likely necessary for sample diversity. For most analyses, this meant that our results and findings only apply to other people like our participants – western, wealthy, White, and liberal. However, it creates a bigger problem when conducting predictive analyses; to have more predictive power, it is preferable to have greater diversity in responses.

Therefore, especially when considering factors that influence the development of children's moral beliefs, it will be important in future research to ensure greater variability in variables of interest.

In conclusion, the primary goal of this chapter was to examine relationships between parents' and children's morals. In doing so, we hoped to investigate the tenet of cultural learning in MFT; that is, that culture and environment influence that strength of children's moral concerns in different areas. We also sought to study these relationships and influences from a non-judgmental, non-prescriptive approach, without specifying what types of moral values are 'good' or indicate 'more' development. In general, we found very few relationships, and those we did find were largely unexpected or difficult to explain. This highlighted that a more complex investigation of not only parental beliefs, but the way in which those beliefs are communicated, as well as child receptivity to those beliefs, is required to thoroughly examine the impact of parents on moral development. A more in-depth discussion of these results as they pertain to MFT is covered in Chapter 6.

Chapter 6: General Discussion

Together, the analyses included in this dissertation examined children's moral beliefs from a broader perspective and at a younger age than had been previously studied. We considered moral beliefs in the areas of care, fairness, loyalty, authority, and sanctity, as outlined by Moral Foundations Theory (MFT). The measures administered in the current study addressed the following questions: 1) How do 2- and 4-year-olds respond to violations of moral foundations (care, fairness, loyalty, authority, and sanctity) on a behavioural puppet task?; 2) Are there age differences in 2- and 4-year-olds' moral concerns?; 3) How do 4-year-olds respond to moral and social norm violations on a verbal questionnaire?; 4) What are the relationships between children's scores on each moral foundation?; and 5) How do parents' moral beliefs influence or relate to children's moral concerns?.

In Chapter 3, we examined 2- and 4-year-olds' moral concerns by measuring their puppet preferences and judgments following scenes of moral violations and moral obedience (Moral Foundations Puppet Task; MFPT). Two-year-olds did not appear to be sensitive to our depictions of moral foundations; their puppet choices were no different from chance across all scenarios and foundations. This was highly unexpected for the care and fairness foundations, as previous literature has consistently shown that children are sensitive to such issues. However, closer examination of the literature revealed methodological differences that may explain this difference – and suggested that 2-year-olds may be sensitive to helping vs. hindering, but less so to physical caring vs. harm. The influence of weaker comprehension was also not ruled out. In contrast to 2-year-olds, 4-year-olds showed greater and significant concern (as shown by significant puppet

preferences) for the care, fairness, authority, and sanctity foundations. They were surprisingly not sensitive to our loyalty scenarios, which was possibly influenced by the use of third-party disloyalty rather than acts performed by ingroup members, as in previous research. We saw some positive correlations between foundations, predominantly involving the sanctity and authority foundations – a different pattern from what has been shown in adults. Overall, the primary conclusion from Chapter 3 was that 4-year-olds were sensitive to all foundations but loyalty, while 2-year-olds were not sensitive to any of the moral foundations, using the current methods.

In Chapter 4, we further examined 4-year-olds' moral concerns using a verbal questionnaire (Moral Foundations Questionnaire for Kids; MFQK). Like Chapter 3, we saw that 4-year-olds were sensitive to the care, fairness, authority, and sanctity foundations, but not particularly sensitive to loyalty or social norm violations. When directly comparing their ratings of violations in each foundation, we saw that care, sanctity, and authority violations were rated as very bad, with no differences between them. Fairness violations were rated slightly less bad, but still quite bad. Loyalty and social norm violations were considered only a little bad, and significantly less bad than all other foundations. Reliability and validity were a concern with the MFQK, particularly for the fairness, loyalty, and odd behaviours (social norms) subscales. Furthermore, we saw significant positive correlations between all foundations on the MFQK; however, we believe that common method variance likely inflated these relationships. Overall, despite methodological and psychometric concerns, when paired with the results from Chapter 3, we feel that Chapter 4 supported the conclusion that our sample of 4-year-olds was sensitive to care, fairness, authority, and sanctity, but not

loyalty.

Finally, in Chapter 5, we examined the influence of parents' morals on their children's responses to moral foundations. We administered the Moral Foundations Questionnaire (MFQ; Graham et al., 2011) to examine parents' moral foundation scores and examined how they related to children's scores on the MFQK and MFPT. In general, we saw very few relationships between parents' and children's morals, and those that we did find were largely unexpected. One relationship – that parents' authority scores were negatively related to children's concern for care – was consistent with previous literature (e.g., Leman, 2005). However, the remainder were difficult to explain. As discussed in Chapter 5, we believe that the relationship between parents' and children's morals is more complex than the picture painted here, and that a number of mediating and moderating factors (e.g., parenting style, parent communication of beliefs) need to be considered in future research. In addition to the influence of parents, we examined previous phenomena discovered using the MFQ in adults, such as relationships between foundations and political groups. Similar to previous research, we found strong correlations within the individualizing (care and fairness) and binding (loyalty, authority, and sanctity) foundations. We also replicated the finding that conservatives place more emphasis on the binding foundations relative to liberals. Considering we replicated these findings in a predominantly liberal sample, this highlighted the robustness of these effects.

Novel Contributions

Altogether, this dissertation represents one the first examinations of MFT in child development literature. It provides insight into the development of children's moral

beliefs, especially outside of the areas of care, justice, and fairness. The contributions of this research pertain to two primary areas: MFT, and the existing moral development literature.

Theoretical Contributions

Despite making claims regarding the partly innate, partly environmentally influenced development of moral foundations, MFT had yet to be examined in young children. In particular, while there was some relevant literature for most moral foundations in children, there was none on loyalty, authority, and sanctity in children younger than the ages of 3 or 4. Interestingly, in our study, we saw that 2-year-olds were not consistently sensitive to any of the moral issues they were presented with. In terms of MFT, the most natural question that stems from this result is likely, "But what does this mean for nativism?" Indeed, even Haidt (2012) presented the infant studies of Hamlin et al. (2007) as evidence that children are born with innate moral intuitions and capabilities. However, it is wrong to equate innateness with evident at birth or early childhood (Haidt, 2013). In MFT, innate means that the genetic and biological underpinnings of moral foundations are present at birth, and we are primed to pay more attention to concerns in those areas. It does not mean that we will show particular behaviours or judgments in infancy. For example, language is commonly discussed as having innate influences (e.g., Yang et al., 2017) – but, as evidenced regularly by the participants of this study, it still takes children a long time to speak correctly and effectively.

From an MFT perspective, there are several other reasons why the 2-year-olds in our study did not show significant puppet preferences following the moral violations that were presented. Although moral foundations are believed to be innate, they also become

more developed and refined through exposure to cultural messages, the behaviours of others, and other environmental factors (Haidt & Bjorklund, 2007). Therefore, it is possible that at the age of 2, children have not had enough time to learn, absorb, and apply concepts to their innate moral framework. It is also possible that, because their moral foundations are still developing, their responses in many contexts are still inappropriate and ineffective. Consequently, 2-year-olds may show beliefs or judgments in some scenarios (e.g., looking time behaviour following care or fairness violations; DesChamps et al., 2016; Hamlin et al., 2011), but not others. In our study, children had to translate concerns into pointing or verbal responses, which may be too complex for their immature moral foundations to process. Overall, with null results, we cannot conclusively say that 2-year-olds were or were not sensitive to the moral foundations. However, it is important to remember that such results do not rule out nativism as an influence on moral development.

Although 2-year-olds did not show significant puppet preferences in response to violations, 4-year-olds did – and overwhelmingly so – for the care, fairness, authority, and sanctity foundations. If we take these findings at face value (i.e., that 4-year-olds are sensitive to these foundations but 2-year-olds are not), there are a few different implications for MFT. One would be that it supports the tenet of cultural learning and environmental influence on the development of morality. Relative to 2-year-olds, 4-year-olds have certainly had more cultural exposure, and more time to develop their moral foundations accordingly. We might also consider that it supports the combination of both innate and environmental influences; to develop such strong preferences in such a short amount of time could be suggestive of a predisposition to having concerns in those areas.

However, other factors (i.e., outside of MFT) may also explain the differences between 2- and 4-year-olds. For example, a great deal of development occurs during the preschool years that could impact 4-year-olds ability to understand videos, ability to respond to questions, and more. Therefore, age differences, especially cross-sectional ones, are not a particularly strong argument for cultural learning.

One way in which we were hoping to more formally study MFT's tenet of cultural learning was through relationships between parents' and children's morals. However, although some relationships were found, these relationships were particularly hard to understand through the lens of cultural learning. We expected that parents' beliefs in a particular foundation would directly correspond to their children's beliefs in that same foundation; instead, the few relationships we found were predominantly between dissimilar foundations (e.g., parents' sanctity scores with children's fairness scores). Previous research has shown that parenting style and parents' empathy influences children's development of moral reasoning about care and fairness (e.g., Walker & Hennig, 1999). Rather than parenting style, we chose to study parents' moral beliefs, as this seemed like a simple and more direct way of investigating cultural learning and the transmission of beliefs between generations. As we highlighted in Chapter 5, there is likely a more complex relationship at play. This is perhaps an area where MFT could be more specific; MFT is clear that morality is influenced by cultural learning, but is less clear regarding specifically what constitutes cultural learning, or if there are particular factors (e.g., beliefs vs. parenting style, parents vs. society) that matter more than others. Going forward, it will be important to consider how cultural learning is defined and what aspects of the environment are important.

Although the current study could not directly support either nativism or cultural learning, we believe we did show evidence for MFT's tenet of moral pluralism in young children – or, at the very least, that their morality extends beyond care and fairness. On the MFQK, we saw that 4-year-olds were significantly more sensitive to authority and sanctity violations than social norm violations. For these social norm violations, there was some concern that items were confounded with the authority foundation. However, if anything, we feel the fact that we still saw differences with the authority and sanctity violations really highlights that these foundations are indeed qualitatively different and morally relevant. This was further supported by anecdotal evidence. When confronted by social norm violations, many children seemed hesitant or unsure how to respond. In comparison, authority and sanctity violations were often met with gasps or verbal displays of disbelief. Overall, we believe our study refutes previous conceptualizations of morality as monist, or based only on care and fairness (Kohlberg, 1971). Instead, it supports the application of MFT's moral pluralism (Graham et al., 2013) to a much younger population than previously studied.

Four-year-olds' strong concern for issues of authority and sanctity were particularly interesting in the context of sample demographics; they belonged to a Western, liberal society that typically places less emphasis on such concerns (Graham et al., 2011). Graham et al. (2013) highlighted that the development and refinement of moral foundations enables people to navigate and function optimally in their culture or society. It is worth considering that what is optimal may vary by age as well. For example, young children typically do not make great decisions, and are not very good at avoiding contaminants. Therefore, it is possible that adults need to invoke their

predisposition towards authority and sanctity in order for them to function better in society. As children age, these concerns might decline in a Western society, but stay the same or increase in others (e.g., war-ridden countries, cultures with strong hierarchies). This strong concern for authority and sanctity in early childhood may not be universal; for example, children who grow up exclusively at home, with many family members close by to protect them, may not need to be as sensitive to authority and sanctity from such a young age. In general, the fact that our sample of 4-year-olds was sensitive to authority and sanctity highlights that what is morally relevant to individuals from the same society can vary with age, and more descriptive research is needed on moral beliefs across different ages and cultures.

In sum, although the lack of support for nativism or cultural learning in the current study may prove frustrating for some readers, proper testing of these tenets was not our primary goal. Our goal was to provide novel, descriptive research on children's moral development using MFT as a guiding framework. In doing so, we believe that we showed evidence of moral pluralism in young children and the extension of the moral domain beyond care and fairness. Furthermore, this study supports the pragmatic validity of MFT; as Graham et al. (2013) highlighted, the usefulness of a theory is not only in its accuracy, but in the novel research it inspires. The novel contributions of our study are further highlighted in the next section, where we see how it extended existing research on moral development.

Contributions to Moral Development Literature

In his book, Haidt (2012) outlined how adult morality research has been heavily biased towards a liberal perspective. In essence, researchers have focused on trying to

explain how people come to adopt conservative moral beliefs, and why they are morally wrong. In addition to creating larger societal and cultural divides, this type of approach is bad scientifically, as it creates less objective and more biased conclusions. This phenomenon is not exclusive to adult research. For the past half-century, moral development literature has been conducted largely from a rationalist, cognitive perspective. In such theories, concerns outside of care and fairness are viewed as conventional and not morally relevant (Kohlberg & Hersh, 1977; Piaget, 1977; Smetana, 2013). Researchers defined what was 'more' or 'less' moral, where greater morality was seen as putting exclusive emphasis on care and fairness – essentially, being more liberal. Studies then highlighted the range of negative qualities and factors (e.g., problem behaviours) that were related to such moral beliefs. Therefore, one of the biggest contributions of our study to moral development literature is that it was conducted from a non-judgmental, non-prescriptive approach. Using this approach, we were able to broaden the range of moral concerns previously studied in young children, and highlight that there is more to morality than care and fairness.

Although concerns about loyalty and authority were not considered to be moral, they had still been studied in children as young as 4 years old. Therefore, the contribution of breadth was especially true for very young children (i.e., toddlers), as well as concerns about sanctity. There was no existing research indicating what toddlers think of disobedience or disloyalty, and no research in general on what children think about purity or sanctity. To understand the full moral spectrum, when moral concerns first emerge, and how they are influenced by environmental or other factors, we need descriptive research on diverse moral areas across all ages. In the current study, given our null results

in 2-year-olds, we were unable to conclusively determine what toddlers thought of our broad spectrum of moral issues. However, we did show that sanctity and authority are moral concerns in 4-year-olds. Furthermore, our hope is that we have highlighted the importance of a wider moral palate, and that this will inspire future research. Over time, as research in this area grows, it will be able to become less exploratory and more hypothesis driven, generating more conclusive evidence regarding toddlers' beliefs about loyalty, authority, and sanctity.

In addition to broadening the range of concerns considered moral, the current study broadened the methodologies that have been used to study moral concerns in the existing literature. In previous chapters, we discussed how methodological differences may explain some inconsistencies between this study and previous ones, such as the lack of concern for care and fairness in 2-year-olds. Although we did not consider this prior to our study's outcomes, we now believe that the homogeneity of previous studies, especially in infants and toddlers, is a limitation of the existing body of research. In children younger than 3, we saw that studies have focused exclusively on helping or hindering the completion of a goal, and not on other depictions of care or harm (e.g., Hamlin et al., 2007). Similarly, toddler fairness research has focused solely on looking time reactions to distributive equality (e.g., DesChamps et al., 2016). Consistency is important in research, as is replication. They allow us to compare and confirm our conclusions and interpretations. However, breadth and variation both have their place. Even just within the domains of care and fairness, moral concerns encompass a broad spectrum of values and beliefs. Therefore, a contribution of the current study is that the moral scenarios presented to children were novel and diverse. This made it difficult to

compare findings, but really highlighted just how important methodological variance is for a complete understanding of morality.

A significant strength of the current study, relative to existing moral development research, is the sample size. To our knowledge, it is one of the largest studies conducted on the development of morality in children younger than 3; we had 102 two-year-olds in our sample alone. For issues of care and fairness, the infant and toddler research presented in this paper has predominantly been conducted with samples of 50 participants or less. Most employed samples of approximately 30 participants, while some had less than 20 per experiment or study. In Chapter 3, we discussed the issues small sample sizes present specifically in looking time research. These issues are not specific to looking time research; rather, they tend to occur because looking time studies are conducted with infants and toddlers, a notoriously difficult population to recruit and study. As outlined by Oakes (2017), small sample sizes can lead to both false positives and false negatives. When we consider, for example, the puppet preference studies conducted by Hamlin and colleagues (e.g., Hamlin et al., 2011; Hamlin & Wynn, 2011), we see very large effects in very small samples. Considering the amount of error and noise in toddler and infant data, it is perhaps worth questioning the influence of false positives on such research. Overall, the current study was one of the largest conducted in its field. Combined with the use of blinded experimenters and third-party coding, we believe this reinforces its significance, and indicates that it merits proper consideration amongst the array of large, significant effects found in rather small samples.

Finally, a discussion of how this study contributes to existing moral development literature would be remiss without touching on cross-cultural research. The current study

was certainly conducted using a western, educated, industrialized, rich, and democratic sample (WEIRD; Henrich et al., 2010). However, MFT and the approach used in the current study provide an excellent framework for future cross-cultural research. Moral development has been examined in different cultures, but has typically been approached from a Kohlbergian perspective. Therefore, the resulting findings have suggested that other cultures are behind in their moral reasoning or development (Shweder et al., 1987) – indicating a very biased perspective. Using a non-judgmental, non-prescriptive approach, researchers can conduct research that is less biased and paints a more comprehensive view of morality. Furthermore, the methods employed in the current study lend themselves well to future cross-cultural research. In particular, the scenarios depicted in the MFPT relied very little on language, and were designed to be very simple, basic representations of moral foundations. Therefore, we believe they could be easily adapted to different languages and cultures. Overall, it is our hope that the approach and methods used in the current study will provide a springboard for descriptive moral research across ages, cultures, and countries.

Implications

There are numerous implications that stem from a broader, less judgmental approach to studying morality in young children and across the lifespan. Many of these implications originate in the fact that differences in moral beliefs are often at the root of inter-societal and inter-nation conflicts. These conflicts can vary from arguments about policies, to terrifying acts of war, genocide, or human rights violations. The traditional approach to morality, where people are viewed as being more or less moral, does little to help explain or resolve such conflicts. We are certainly not proposing that MFT, or

understanding how diverse moral concerns emerge over childhood, are going to miraculously resolve world problems. However, we think that with MFT, and the approach to moral research taken in the current study, we can paint a more comprehensive picture of how such conflicts arise, and perhaps even highlight some areas of common moral ground.

Differences in moral beliefs can cause smaller conflicts too, such as conflicts between families, friends, colleagues, or peers. Understanding people's concern for moral foundations may have particular importance for children and families. Societal change occurs rapidly, as do the beliefs and opinions that young people are exposed to, especially in the age of the internet. As such, children often grow up in different moral climates than their parents, and may adopt different moral beliefs. Even in the current study with very young children, we saw that children's moral profiles differed from those of their parents. This is not to say that parents do not influence their children's morals; we believe parents are important players in cultural learning. However, the foundations that children build likely differ, to varying degrees, from their parents. Therefore, on top of the variables previously established (e.g., social support; Rodriguez et al., 2018), we believe that moral foundations might provide additional insight and further predict why some families have larger conflicts than others.

In addition to better understanding the roots of conflicts, having a broader and better picture of people's moral values may be useful for socioemotional, occupational, or other areas of functioning. Value-based living is currently a popular topic in psychological research, and living according to one's values has a number of positive effects such as improved psychological well-being and better academic performance

(Chase et al., 2013; Ciarrochi et al., 2011). Values that are commonly assessed include achievement, family, adventure, and more. We believe that moral values may also contribute to better values-based living, as they guide how people act and how they make judgments. For example, someone who values loyalty may thrive in a team sport, rather than an individual one. Similarly, moral values may be useful to predict or determine what types of jobs or roles people are most suited for; someone who highly values authority may fit well in a military role, while someone who values it less may prefer running their own business or freelancing. In general, we believe that continued study of moral values across the lifespan could provide more insight into how, where, and why people thrive the most.

In sum, on its own, the current study helps to explain the moral beliefs held by young children and how those may change with age. However, taking into account the theory on which it is based, the contributions to the literature, and the directions it provides for future research, there are many important implications for groups and individuals.

Limitations

Overall, this study provided a broad investigation into children's moral beliefs using a large sample of toddlers, preschoolers, and their parents. Limitations pertinent to each individual chapter have already been discussed. However, across all three chapters and analyses, there are some important shared limitations to consider. First of all, the population we recruited was very WEIRD (Henrich et al., 2010). As shown in Chapter 2, Table 3, the parents in our sample were predominantly white, liberal, wealthy, and educated. It is worth noting that the population from which our sample was drawn (i.e., a

small Canadian province) is chiefly White and liberal (Statistics Canada, 2017). However, our sample was still much more educated and wealthier than average. Therefore, many of the conclusions and interpretations of the current study are limited in their application to groups without similar characteristics. Furthermore, because the characteristics of our sample were quite narrow, there were likely less-than-desired levels of variability in the moral beliefs of both children and parents. This may have limited our investigations of relationships between moral foundations (within children, and between children and their parents), and contributed to the lack of expected relationships, particularly in Chapter 5.

The limitations of the MFQK, most notably the reliability and validity of the measure, have already been discussed at length in Chapter 4. However, it is worth re-iterating that these limitations applied to all analyses where this measure was utilized. For the MFPT, we believe that the scenarios depicted and language used were appropriate (although the use of videos may have interfered with comprehension, as discussed in Chapter 3). However, the way in which we assessed concern for the scenarios, particularly for 2-year-olds, may have been too complex. That is, rather than observing which puppet children automatically pick up or play with (i.e., like Hamlin et al., 2007), we asked them two questions that they had to respond to using gestures or words. They had to remember what they saw, how they felt, and translate that into an appropriate response. This likely required more working memory and attention than some 2-year-olds were capable of. Therefore, in future research, a simple, less complex assessment of toddler's reactions to moral violations may be warranted.

Another methodological consideration involved the presence of experimenters,

and occasionally parents or daycare teachers, in the room while children were completing the study. When possible, we tried to limit children being accompanied by parents or teachers. However, this was not always feasible due to child comfort or daycare policy (especially for 2-year-olds). Furthermore, there were always 2 experimenters present for the MFPT, and at least one present for the MFQK. Therefore, it is both possible that children adjusted their responses, or were influenced by experimenter or parent bias. Anecdotally, this seemed unlikely for 2-year-olds, who appeared to be largely indifferent and unaware of other's opinions or biases. However, 4-year-olds were certainly more in tune with people in their environment and may have been more greatly influenced.

Finally, the last limitation that will be discussed here pertains to the ages of participants recruited. We chose to recruit 2-year-olds to extend the study of certain moral foundations to younger ages (e.g., loyalty, authority). We also recruited 4-year-olds to examine cross-sectional age differences, and because they were the youngest age group that could feasibly complete the MFQK. We opted for two distinct age groups as it allowed us to recruit more participants within each group and focus on what moral concerns look like at those particular ages. However, this limited our ability to determine how moral concerns emerge and develop with age. For example, 2-year-olds showed no significant puppet preferences to moral situations, while 4-year-olds did – however, as we did not recruit any 3-year-olds, we could not determine when this change occurs. Therefore, in future research, it will be important to study morality continuously across the lifespan.

Directions for Future Research

Despite decades of research on moral development, the results of the current

study highlight that there is still so much we do not know. In the current study, we attempted to examine a broader array of moral concerns than previously considered in developmental research. However, within one moral foundation, there are an immense number of things that can become moralized or relevant. Therefore, future studies will need to study moral breadth within foundations, as well as between. Diversity in methods is important too; although the MFPT was labelled as a behavioural task, in reality it was still a type of self-report. As highlighted by Graham et al. (2011), going beyond self-report data is important; they highlighted that moral behaviours, life events, and implicit attitudes are key to studying the full moral domain. Graham et al. (2011) also noted that, while MFT provides a great starting point, it is likely an incomplete map of morality. To be truly non-prescriptive and non-judgmental, researchers should not only consider issues within the five foundations. Other moral concerns should be specifically looked for and studied, to refine both the theory and the study of children's morality.

In addition to the moral foundations themselves, the tenets of MFT – most notably cultural learning and nativism – require further study and evaluation. For nativism, there is no concrete way to confirm or deny if an ability or foundation is indeed innately influenced, although there are ways to provide more support for the claim. Evolutionary and anthropological analyses play key roles in providing this support. However, empirical approaches are still useful; for example, a study of neurobiological data and how it corresponds to children's reactions to moral situations may shed light on the biological underpinnings of morality. Furthermore, following the notion of psychological preparedness, future research could examine how easily children learn to moralize novel behaviours in different domains. For example, similar to how Rottman and Kelemen

(2012) studied children's acquisition of novel purity concerns, researchers could examine how children learn novel concerns in each foundation and other areas. Testing to see if there are differences in how readily they acquire new morals may shed light on potential predispositions (although, this is not a perfect approach, as it will still be difficult to disentangle the influence of values they hold already).

Finally, for cultural learning, a wider consideration of culture and environment will be important in future investigations. Parent-child relationships are significant, but other factors matter too – such as education, society, or media consumed. Furthermore, in order to truly examine predictive relationships, longitudinal research is essential. Large-scale, longitudinal studies will inform when children's moral beliefs emerge, how they change over time, and what factors influence them at different ages. They will also indicate if early moral beliefs relate to later ones, or if moral foundations pass through certain developmental stages. As noted previously, conducting said research across multiple societies, countries, and cultures will also be important in understanding the development of moral foundations. In sum, MFT has opened a number of interesting avenues of research on moral development. The current study represents the beginning of said research; we look forward to seeing how this research progresses in the coming years.

Conclusion

In conclusion, the current study examined the development of moral foundations in toddlers and preschoolers, and the correspondence of children's moral foundations with their parents' moral beliefs. In his book *The Righteous Mind*, Jonathan Haidt (2012) rejected the notion that some people are good, while others are evil. Instead, he argued

that people around the world take the same innate moral predispositions and build different moral matrices on top of them. This makes it difficult for people to understand each other, but does not mean that one matrix is better or more mature than another. In the current study, we employed Haidt's Moral Foundations Theory to broaden the study of morality in children. In addition to the traditionally studied concerns about care and fairness, we examined 2- and 4-year-olds' responses to violations of loyalty, authority, and sanctity. In our predominantly white, Western, liberal, and wealthy sample, we found that 2-year-olds did not show consistent preferences or judgments in response to any of the moral foundations. In contrast, on two different measures, 4-year-olds were sensitive to issues of care, fairness, authority, and sanctity, but not loyalty. We saw little evidence of parent influence on children's morals, however, this was likely due to the impact of moderators or mediators that were not studied, as well as the narrow diversity of beliefs in our sample. This study was novel and important due to the non-prescriptive, non-judgmental approach used to study morality. It demonstrated support for the existence of moral beliefs outside of care and fairness in young children and provided useful future directions in the study of both MFT and moral development in general. Such research has the potential to better elucidate the causes of political and societal conflicts, and, although it will not solve the world's problems, it might add a little more compassion and understanding.

References

Abrams, D., Rutland, A., & Cameron, L. (2003). The Development of Subjective Group Dynamics: Children's Judgments of Normative and Deviant In-Group and Out-Group Individuals. *Child Development*, *74*(6), 1840–1856. https://doi.org/10.1046/j.1467-8624.2003.00641.x

Abrams, D., Rutland, A., Ferrell, J. M., & Pelletier, J. (2008). Children's Judgments of Disloyal and Immoral Peer Behavior: Subjective Group Dynamics in Minimal Intergroup Contexts. *Child Development*, *79*(2), 444–461. https://doi.org/10.1111/j.1467-8624.2007.01135.x

Anderson, D. R., & Pempek, T. A. (2005). Television and Very Young Children. *American Behavioral Scientist*, *48*(5), 505–522. https://doi.org/10.1177/0002764204271506

Askew, C., Çakır, K., Põldsam, L., & Reynolds, G. (2014). The effect of disgust and fear modeling on children's disgust and fear for animals. *Journal of Abnormal Psychology*, *123*(3), 566–577. https://doi.org/10.1037/a0037228

Atkin, L. M., & Gummerum, M. (2012). Moral and Group-based Criteria in Children's Evaluation of Peers after Transgressions: Peer Evaluations after Transgressions. *Infant and Child Development*, *21*(2), 189–197. https://doi.org/10.1002/icd.744

Barr, R., & Hayne, H. (1999). Developmental Changes in Imitation from Television during Infancy. *Child Development*, *70*(5), 1067–1081. https://doi.org/10.1111/1467-8624.00079

Baumard, N., Mascaro, O., & Chevallier, C. (2012). Preschoolers are able to take merit into account when distributing goods. *Developmental Psychology*, *48*(2), 492–498. https://doi.org/10.1037/a0026598

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300. JSTOR.

Benozio, A., & Diesendruck, G. (2015). Parochialism in preschool boys' resource allocation. *Evolution and Human Behavior*, *36*(4), 256–264. https://doi.org/10.1016/j.evolhumbehav.2014.12.002

Benozio, A., & Diesendruck, G. (2017). Parochial Compliance: Young Children's Biased Consideration of Authorities' Preferences Regarding Intergroup Interactions. *Child Development*, *88*(5), 1527–1535. https://doi.org/10.1111/cdev.12654

Buttelmann, D., & Böhm, R. (2014). The Ontogeny of the Motivation That Underlies In-Group Bias. *Psychological Science*, *25*(4), 921–927. https://doi.org/10.1177/0956797613516802

Campaert, K., Nocentini, A., & Menesini, E. (2018). The role of poor parenting and parental approval for children's moral disengagement. *Journal of Child and Family Studies*, *27*(8), 2656–2667. https://doi.org/10.1007/s10826-018-1097-1

Chambers, C. T., & Johnston, C. (2002). Developmental Differences in Children's Use of Rating Scales. *Journal of Pediatric Psychology*, *27*(1), 27–36. https://doi.org/10.1093/jpepsy/27.1.27

Chase, J. A., Houmanfar, R., Hayes, S. C., Ward, T. A., Vilardaga, J. P., & Follette, V. (2013). Values are not just goals: Online ACT-based values training adds to goal setting in improving undergraduate college student performance. *Journal of Contextual Behavioral Science*, *2*(3), 79–84. https://doi.org/10.1016/j.jcbs.2013.08.002

Chernyak, N., Sandham, B., Harris, P. L., & Cordes, S. (2016). Numerical cognition explains age-related changes in third-party fairness. *Developmental Psychology*, *52*(10), 1555–1562. https://doi.org/10.1037/dev0000196

Ciarrochi, J., Fisher, D., & Lane, L. (2011). The link between value motives, value success, and well-being among people diagnosed with cancer. *Psycho-Oncology*, *20*(11), 1184–1192. https://doi.org/10.1002/pon.1832

Clifford, S., Iyengar, V., Cabeza, R., & Sinnott-Armstrong, W. (2015). Moral foundations vignettes: A standardized stimulus database of scenarios based on moral foundations theory. *Behavior Research Methods*, *47*(4), 1178–1198. https://doi.org/10.3758/s13428-014-0551-2

Cowell, J. M., & Decety, J. (2015). Precursors to morality in development as a complex interplay between neural, socioenvironmental, and behavioral facets. *Proceedings of the National Academy of Sciences*, *112*(41), 12657–12662. https://doi.org/10.1073/pnas.1508832112

Curtis, R., & Moore, C. (2019, unpublished study). *Moral Foundations Questionnaire for Kids*.

Danovitch, J., & Bloom, P. (2009). Children's extension of disgust to physical and moral events. *Emotion*, *9*(1), 107–112. https://doi.org/10.1037/a0014113

DesChamps, T. D., Eason, A. E., & Sommerville, J. A. (2016). Infants Associate Praise and Admonishment with Fair and Unfair Individuals. *Infancy*, *21*(4), 478–504. https://doi.org/10.1111/infa.12117

Dewey, J. (1892). *Psychology* (3rd ed.). Harper Brothers.

Dien, D. S. (1982). A Chinese perspective on Kohlberg's theory of moral development. *Developmental Review*, *2*(4), 331–341. https://doi.org/10.1016/0273-2297(82)90017-X

Dingman, D., Wu, J., & Murphy, H. M. (2020). School-based, blacklight handwashing program can improve handwashing quality and knowledge among pre-school aged children. *Evaluation and Program Planning*, *78*, 101731. https://doi.org/10.1016/j.evalprogplan.2019.101731

Dixon, S. V., Graber, J. A., & Brooks-Gunn, J. (2008). The roles of respect for parental authority and parenting practices in parent-child conflict among African American, Latino, and European American families. *Journal of Family Psychology*, *22*(1), 1–10. https://doi.org/10.1037/0893-3200.22.1.1

Durkheim, E. (1951). *Suicide* (J. A. S. G. Simpson, Trans.). The Free Press
.

Durkheim, E. (1973). *Moral education* (E. K. Wilson & H. Schnurer, Eds.). The Free Press.

Enesco, I., Guerrero, S., Lago, M. O., & Rodríguez, P. (2011). ¿Es estable la orientación afectiva de los niños hacia los diferentes grupos étnicos? Un estudio longitudinal de las actitudes intragrupo y exogrupo en niños de Escuela Infantil. *Anales de Psicología / Annals of Psychology*, *27*(3), 631–638.

Engelmann, J. M., Herrmann, E., Rapp, D. J., & Tomasello, M. (2016). Young children (sometimes) do the right thing even when their peers do not. *Cognitive Development*, *39*, 86–92. https://doi.org/10.1016/j.cogdev.2016.04.004

Erdmann, K., Kavšek, M., & Heil, M. (2018). Infants' looking times in a dynamic mental rotation task: Clarifying inconsistent results. *Cognitive Development*, *48*, 279–285. https://doi.org/10.1016/j.cogdev.2018.10.002

Fang, G., Fang, F., Keller, M., Edelstein, W., Kehle, T. J., & Bray, M. A. (2003). Social moral reasoning in Chinese children: A developmental study. *Psychology in the Schools*, *40*(1), 125–138. https://doi.org/10.1002/pits.10074

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, *454*(7208), 1079–1083. https://doi.org/10.1038/nature07155

Franks, A. S., & Scherr, K. C. (2015). Using Moral Foundations to Predict Voting Behavior: Regression Models from the 2012 U.S. Presidential Election. *Analyses of Social Issues and Public Policy*, *15*(1), 213–232. https://doi.org/10.1111/asap.12074

Frimer, J. A., Biesanz, J. C., Walker, L. J., & MacKinlay, C. W. (2013). Liberals and conservatives rely on common moral foundations when making moral judgments about influential people. *Journal of Personality and Social Psychology*, *104*(6), 1040–1059. https://doi.org/10.1037/a0032277

Fu, G., Luo, Y. C., Heyman, G. D., Wang, B., Cameron, C. A., & Lee, K. (2016). Moral Evaluations of Lying for One's Own Group: Lying for One's Own Group. *Infant and Child Development*, *25*(5), 355–370. https://doi.org/10.1002/icd.1941

Garmon, L. C., Basinger, K. S., Gregg, V. R., & Gibbs, J. C. (1996). Gender Differences in Stage and Expression of Moral Judgment. *Merrill-Palmer Quarterly*, *42*(3), 418–437. JSTOR.

Gelman, R., & Baillargeon, R. (1983). A review of some Piagetian concepts. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology, volume 3: Cognitive development*. Wiley Press.

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources: Evaluation of fairness of distributive actions. *Developmental Science*, *14*(5), 1012–1020. https://doi.org/10.1111/j.1467-7687.2011.01048.x

Gilligan, C. (1982). *In A Different Voice: Psychological Theory and Women's Development* (Vol. 326).

Gilligan, C. (1987). Adolescent Development Reconsidered. *New Directions for Child Development*, 63–92.

Gilligan, C., & Attanucci, J. (1988). Two Moral Orientations: Gender Differences and Similarities. *Merrill-Palmer Quarterly*, *34*(3), 223–237. JSTOR.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Chapter Two - Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism. In P. Devine & A. Plant (Eds.), *Advances in Experimental Social Psychology* (Vol. 47, pp. 55–130). Academic Press. https://doi.org/10.1016/B978-0-12-407236-7.00002-4

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029–1046. https://doi.org/10.1037/a0015141

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*(2), 366–385. https://doi.org/10.1037/a0021847

Groenendyk, A. E., & Volling, B. L. (2007). Coparenting and early conscience development in the family. *The Journal of Genetic Psychology: Research and Theory on Human Development*, *168*(2), 201–224. https://doi.org/10.3200/GNTP.168.2.201-224

Hahn, E. R., & Garrett, M. K. (2017). Preschoolers' moral judgments of environmental harm and the influence of perspective taking. *Journal of Environmental Psychology*, *53*, 11–19. https://doi.org/10.1016/j.jenvp.2017.05.004

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), 613–628. https://doi.org/10.1037//0022-3514.65.4.613

Haidt, Jonathan. (2012). *The righteous mind: Why good people are divided by politics and religion*. Vintage Books.

Haidt, Jonathan. (2013). Moral psychology for the twenty-first century. *Journal of Moral Education*, *42*(3), 281–297. https://doi.org/10.1080/03057240.2013.817327

Haidt, Jonathan, & Bjorklund, F. (2007). Social Intuitionists Answer Six Questions About Morality. In W. Sinnott-Armstrong (Ed.), *Moral Psychology, Vol.2* (pp. 181–217). Oxford University Press. https://papers.ssrn.com/abstract=855164

Haidt, Jonathan, & Graham, J. (2009). Planet of the Durkheimians, Where Community, Authority, and Sacredness are Foundations of Morality. In *Social and Psychological Bases of Ideology and System Justification*. Oxford University Press.

Haidt, Jonathan, & Joseph, C. (2008). The Moral Mind: How Five Sets of Innate Intuitions Guide the Development of Many Culture- Specific Virtues, and Perhaps Even Modules. In *The Innate Mind, Volume 3: Foundations and the Future*. Oxford University Press.

Haidt, Jonathan, & Joseph, C. (2011). How Moral Foundations Theory Succeeded in Building on Sand: A Response to Suhler and Churchland. *Journal of Cognitive Neuroscience*, *23*(9), 2117–2122. https://doi.org/10.1162/jocn.2011.21638

Hamlin, J. K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences*, *108*(50), 19931–19936. https://doi.org/10.1073/pnas.1110306108

Hamlin, J. Kiley. (2013). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition*, *128*(3), 451–474. https://doi.org/10.1016/j.cognition.2013.04.004

Hamlin, J. Kiley, Mahajan, N., Liberman, Z., & Wynn, K. (2013). Not Like Me = Bad: Infants Prefer Those Who Harm Dissimilar Others. *Psychological Science*, *24*(4), 589–594. https://doi.org/10.1177/0956797612457785

Hamlin, J. Kiley, & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development*, *26*(1), 30–39. https://doi.org/10.1016/j.cogdev.2010.09.001

Hamlin, J. Kiley, & Wynn, K. (2012). Who knows what's good to eat? Infants fail to match the food preferences of antisocial others. *Cognitive Development*, *27*(3), 227–239. https://doi.org/10.1016/j.cogdev.2012.05.005

Hamlin, J. Kiley, Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557–559. https://doi.org/10.1038/nature06288

Hayne, H., Herbert, J., & Simcock, G. (2003). Imitation from television by 24- and 30-month-olds. *Developmental Science*, *6*(3), 254–261. https://doi.org/10.1111/1467-7687.00281

Helwig, C. C., & Prencipe, A. (1999). Children's judgments of flags and flag-burning. *Child Development*, *70*(1), 132–143. https://doi.org/10.1111/1467-8624.00010

Helwig, C. C., Zelazo, P. D., & Wilson, M. (2001). Children's Judgments of Psychological Harm in Normal and Noncanonical Situations. *Child Development*, *72*(1), 66–81. https://doi.org/10.1111/1467-8624.00266

Helzer, E. G., & Pizarro, D. A. (2011). Dirty Liberals!: Reminders of Physical Cleanliness Influence Moral and Political Attitudes. *Psychological Science*, *22*(4), 517–522. https://doi.org/10.1177/0956797611402514

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, *33*(2–3), 61–83; discussion 83-135. https://doi.org/10.1017/S0140525X0999152X

Hepach, R., Haberl, K., Lambert, S., & Tomasello, M. (2017). Toddlers Help Anonymously. *Infancy*, *22*(1), 130–145. https://doi.org/10.1111/infa.12143

Hepach, R., Vaish, A., Grossmann, T., & Tomasello, M. (2016). Young Children Want to See Others Get the Help They Need. *Child Development*, *87*(6), 1703–1714. https://doi.org/10.1111/cdev.12633

Hepach, R., Vaish, A., & Tomasello, M. (2012). Young Children Are Intrinsically Motivated to See Others Helped. *Psychological Science*, *23*(9), 967–972. https://doi.org/10.1177/0956797612440571

Hopkins, J. R. (2011). The Enduring Influence of Jean Piaget. *APS Observer*, *24*(10). https://www.psychologicalscience.org/observer/jean-piaget

Horberg, E. J., Oveis, C., Keltner, D., & Cohen, A. B. (2009). Disgust and the moralization of purity. *Journal of Personality and Social Psychology*, *97*(6), 963–976. https://doi.org/10.1037/a0017423

Inbar, Y., Pizarro, D., Iyer, R., & Haidt, J. (2012). Disgust Sensitivity, Political Conservatism, and Voting. *Social Psychological and Personality Science*, *3*(5), 537–544. https://doi.org/10.1177/1948550611429024

Jambon, M., & Smetana, J. G. (2014). Moral complexity in middle childhood: Children's evaluations of necessary harm. *Developmental Psychology*, *50*(1), 22–33. https://doi.org/10.1037/a0032992

Janoff-Bulman, R., & Carnes, N. C. (2013). Surveying the Moral Landscape: Moral Motives and Group-Based Moralities. *Personality and Social Psychology Review*, *17*(3), 219–236. https://doi.org/10.1177/1088868313480274

Johnson, D. J., Cheung, F., & Donnellan, M. B. (20140602). Does cleanliness influence moral judgments? A direct replication of Schnall, Benton, and Harvey (2008). *Social Psychology*, *45*(3), 209. https://doi.org/10.1027/1864-9335/a000186

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, *129*(3), 339–375. https://doi.org/10.1037/0033-2909.129.3.339

Karmakar, R. (2015). Does Parenting Style Influence the Internalization of Moral Values in Children and Adolescents? *Psychological Studies*, *60*(4), 438–446. https://doi.org/10.1007/s12646-015-0338-2

Kim, J. M. (1998). Korean Children's Concepts of Adult and Peer Authority and Moral Reasoning. *Developmental Psychology*, *34*(5), 947–955.

Kim, K. R., Kang, J.-S., & Yun, S. (2012). Moral Intuitions and Political Orientation: Similarities and Differences between South Korea and the United States. *Psychological Reports*, *111*(1), 173–185. https://doi.org/10.2466/17.09.21.PR0.111.4.173-185

Kochanska, G. (2002). Mutually Responsive Orientation Between Mothers and Their Young Children: A Context for the Early Development of Conscience. *Current Directions in Psychological Science*, *11*(6), 191–195. https://doi.org/10.1111/1467-8721.00198

Kohlberg, L. (1963). The Development of Children's Orientations Toward a Moral Order. I. Sequence in the Development of Moral Thought. *Vita Humana*, *6*(1–2), 11–33. https://doi.org/10.1159/000269667

Kohlberg, L. (1971). From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In *Psychology and Genetic Epistemology*. Academic Press.

Kohlberg, L., & Hersh, R. H. (1977). Moral development: A review of the theory. *Theory Into Practice*, *16*(2), 53–59. https://doi.org/10.1080/00405847709542675

Koleva, S., Selterman, D., Iyer, R., Ditto, P., & Graham, J. (2014). The Moral Compass of Insecurity: Anxious and Avoidant Attachment Predict Moral Judgment. *Social Psychological and Personality Science*, *5*(2), 185–194. https://doi.org/10.1177/1948550613490965

Krcmar, M. (2010). Can Social Meaningfulness and Repeat Exposure Help Infants and Toddlers Overcome the Video Deficit? *Media Psychology*, *13*(1), 31–53. https://doi.org/10.1080/15213260903562917

Kuhn, E. S., & Laird, R. D. (2011). Individual differences in early adolescents' beliefs in the legitimacy of parental authority. *Developmental Psychology*, *47*(5), 1353–1365. https://doi.org/10.1037/a0024050

Lagattuta, K. H., Nucci, L., & Bosacki, S. L. (2010). Bridging Theory of Mind and the Personal Domain: Children's Reasoning About Resistance to Parental Control. *Child Development*, *81*(2), 616–635. https://doi.org/10.1111/j.1467-8624.2009.01419.x

Laupa, M., & Turiel, E. (1993). Children's concepts of authority and social contexts. *Journal of Educational Psychology*, *85*(1), 191–197. https://doi.org/10.1037/0022-0663.85.1.191

Leman, P. J. (2005). Authority and moral reasons: Parenting style and children's perceptions of adult rule justifications. *International Journal of Behavioral Development*, *29*(4), 265–270. https://doi.org/10.1177/01650250544000044

Lewis, G. J., & Bates, T. C. (2011). From left to right: How the personality system allows basic traits to influence politics via characteristic moral adaptations. *British Journal of Psychology*, *102*(3), 546–558. https://doi.org/10.1111/j.2044-8295.2011.02016.x

Lewis, G. J., Kanai, R., Bates, T. C., & Rees, G. (2012). Moral values are associated with individual differences in regional brain volume. *Journal of Cognitive Neuroscience*, *24*(8), 1657–1663. https://doi.org/10.1162/jocn_a_00239

Li, J., Wang, W., Yu, J., & Zhu, L. (2016). Young Children's Development of Fairness Preference. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01274

Linebarger, D. L., & Walker, D. (2016). Infants' and Toddlers' Television Viewing and Language Outcomes: *American Behavioral Scientist*. https://doi.org/10.1177/0002764204271505

LoBue, V., Nishida, T., Chiong, C., DeLoache, J. S., & Haidt, J. (2011). When Getting Something Good is Bad: Even Three-year-olds React to Inequality: Children's Reactions to Inequality. *Social Development*, *20*(1), 154–170. https://doi.org/10.1111/j.1467-9507.2009.00560.x

Lu, H. J., & Chang, L. (2016). Resource allocation to kin, friends, and strangers by 3- to 6-year-old children. *Journal of Experimental Child Psychology*, *150*, 194–206. https://doi.org/10.1016/j.jecp.2016.05.018

Malka, A., Osborne, D., Soto, C. J., Greaves, L. M., Sibley, C. G., & Lelkes, Y. (2016). Binding Moral Foundations and the Narrowing of Ideological Conflict to the Traditional Morality Domain. *Personality and Social Psychology Bulletin*, *42*(9), 1243–1257. https://doi.org/10.1177/0146167216653936

Marcus, G. (2004). *The Birth of the Mind: How a Tiny Number of Genes Creates The Complexities of Human Thought* (Reprint edition). Basic Books.

Martín-Antón, L. J., Monjas, M. I., García Bacete, F. J., & Jiménez-Lagares, I. (2016). Problematic Social Situations for Peer-Rejected Students in the First Year of Elementary School. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01925

Melis, A. P., Altrichter, K., & Tomasello, M. (2013). Allocation of resources to collaborators and free-riders in 3-year-olds. *Journal of Experimental Child Psychology*, *114*(2), 364–370. https://doi.org/10.1016/j.jecp.2012.08.006

Mellor, D., & Moore, K. A. (2014). The Use of Likert Scales With Children. *Journal of Pediatric Psychology*, *39*(3), 369–379. https://doi.org/10.1093/jpepsy/jst079

Mulvey, K. L. (2016). Evaluations of moral and conventional intergroup transgressions. *British Journal of Developmental Psychology*, *34*(4), 489–501. https://doi.org/10.1111/bjdp.12145

Muris, P., Huijding, J., Mayer, B., & de Vries, H. (2012). Does 'Yuck' mean 'Eek'? Fear responses in children after a disgust manipulation. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*(2), 765–769. https://doi.org/10.1016/j.jbtep.2011.10.002

Nelson, S. A. (1980). Factors Influencing Young Children's Use of Motives and Outcomes as Moral Criteria. *Child Development*, *51*(3), 823–829. JSTOR. https://doi.org/10.2307/1129470

Nilsson, A., & Erlandsson, A. (2015). The Moral Foundations taxonomy: Structural validity and relation to political ideology in Sweden. *Personality and Individual Differences*, *76*, 28–32. https://doi.org/10.1016/j.paid.2014.11.049

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy*, *22*(4), 436–469. https://doi.org/10.1111/infa.12186

Olson, K. R., & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition*, *108*(1), 222–231. https://doi.org/10.1016/j.cognition.2007.12.003

Osborn, F. W. (1894). The ethical contents of children's minds. *Educational Review*, *8*, 143–146.

Parkinson, C., Sinnott-Armstrong, W., Koralus, P. E., Mendelovici, A., McGeer, V., & Wheatley, T. (2011). Is morality unified? Evidence that distinct neural systems underlie moral judgments of harm, dishonesty, and disgust. *Journal of Cognitive Neuroscience*, *23*(10), 3162–3180. https://doi.org/10.1162/jocn_a_00017

Piaget, J. (1948). *The moral judgment of the child* (pp. ix, 418). Free Press.

Piaget, J. (1977). *The moral judgment of the child*. Penguin Books.

Plötner, M., Over, H., Carpenter, M., & Tomasello, M. (2016). What Is a Group? Young Children's Perceptions of Different Types of Groups and Group Entitativity. *PLOS ONE*, *11*(3), e0152001. https://doi.org/10.1371/journal.pone.0152001

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology*, *25*, 111–163. JSTOR. https://doi.org/10.2307/271063

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, *118*(1), 57–75. https://doi.org/10.1037/a0021867

Rakoczy, H., Kaufmann, M., & Lohse, K. (2016). Young children understand the normative force of standards of equal resource distribution. *Journal of Experimental Child Psychology*, *150*, 396–403. https://doi.org/10.1016/j.jecp.2016.05.015

Raman, L., & Gelman, S. A. (2005). Children's understanding of the transmission of genetic disorders and contagious illnesses. *Developmental Psychology*, *41*(1), 171–182. https://doi.org/10.1037/0012-1649.41.1.171

Raman, L., & Gelman, S. A. (2008). Do children endorse psychosocial factors in the transmission of illness and disgust? *Developmental Psychology*, *44*(3), 801–813. https://doi.org/10.1037/0012-1649.44.3.801

Rempala, D. M., Okdie, B. M., & Garvey, K. J. (2016). Articulating ideology: How liberals and conservatives justify political affiliations using morality-based explanations. *Motivation and Emotion*, *40*(5), 703–719. https://doi.org/10.1007/s11031-016-9563-9

Rhodes, M., & Chalik, L. (2013). Social Categories as Markers of Intrinsic Interpersonal Obligations. *Psychological Science*, *24*(6), 999–1006. https://doi.org/10.1177/0956797612466267

Rodriguez, C. M., Silvia, P. J., & Pu, D. F. (2018). Predictors of change in mothers' and fathers' parent-child aggression risk. *Child Abuse & Neglect*, *86*, 247–256. https://doi.org/10.1016/j.chiabu.2018.09.008

Ross, H. (1984). *Footloose*. Paramount Pictures.

Rottman, J., & Kelemen, D. (2012). Aliens behaving badly: Children's acquisition of novel purity-based morals. *Cognition*, *124*(3), 356–360. https://doi.org/10.1016/j.cognition.2012.06.001

Russell, P. S., & Giner-Sorolla, R. (2011). Moral anger, but not moral disgust, responds to intentionality. *Emotion (Washington, D.C.)*, *11*(2), 233–240. https://doi.org/10.1037/a0022598

Samuelson, S. (1980). The Cooties Complex. *Western Folklore*, *39*(3), 198–210. JSTOR. https://doi.org/10.2307/1499801

Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition*, *124*(3), 325–333. https://doi.org/10.1016/j.cognition.2012.06.004

Schmidt, M. F. H., & Sommerville, J. A. (2011). Fairness Expectations and Altruistic Sharing in 15-Month-Old Human Infants. *PLoS ONE*, *6*(10), e23223. https://doi.org/10.1371/journal.pone.0023223

Seligman, M. E. (1971). Phobias and preparedness. *Behavior Therapy*, *2*(3), 307–320. https://doi.org/10.1016/S0005-7894(71)80064-3

Sengsavang, S., & Krettenauer, T. (2015). Children's moral self-concept: The role of aggression and parent–child relationships. *Merrill-Palmer Quarterly*, *61*(2), 213–235. https://doi.org/10.13110/merrpalmquar1982.61.2.0213

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

Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The "big three" of morality (autonomy, community, divinity) and the "big three" explanations of suffering. In A. M. Brandt & P. Rozin (Eds.), *Morality and health.* (1997-05770-006; pp. 119–169). Taylor & Frances/Routledge.

Slatcher, R. B., & Trentacosta, C. J. (2012). Influences of parent and child negative emotionality on young children's everyday behaviors. *Emotion*, *12*(5), 932–942. https://doi.org/10.1037/a0027148

Sloane, S., Baillargeon, R., & Premack, D. (2012). Do Infants Have a Sense of Fairness? *Psychological Science*, *23*(2), 196–204. https://doi.org/10.1177/0956797611422072

Smetana, J. G. (1999). The Role of Parents in Moral Development: A social domain analysis. *Journal of Moral Education*, *28*(3), 311–321. https://doi.org/10.1080/030572499103106

Smetana, J. G. (2013). Moral development: The social domain theory view. In *The Oxford handbook of developmental psychology (Vol 1): Body and mind* (pp. 832–863). Oxford University 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: **Young Children's Moral Judgments**. *Child Development*, *83*(2), 683–696. https://doi.org/10.1111/j.1467-8624.2011.01714.x

Smetana, J. G., Wong, M., Ball, C., & Yau, J. (2014). American and Chinese Children's Evaluations of Personal Domain Events and Resistance to Parental Authority. *Child Development*, *85*(2), 626–642. https://doi.org/10.1111/cdev.12140

Smith, C. E., Blake, P. R., & Harris, P. L. (2013). I Should but I Won't: Why Young Children Endorse Norms of Fair Sharing but Do Not Follow Them. *PLOS ONE*, *8*(3), e59510. https://doi.org/10.1371/journal.pone.0059510

Sommerville, J. A., Schmidt, M. F. H., Yun, J., & Burns, M. (2013). The Development of Fairness Expectations and Prosocial Behavior in the Second Year of Life: FAIRNESS EXPECTATIONS AND PROSOCIAL BEHAVIOR. *Infancy*, *18*(1), 40–66. https://doi.org/10.1111/j.1532-7078.2012.00129.x

Sperber, D. (1994). The modularity of thought and the epidemiology of representations. In *Mapping the mind: Domain specificity in cognition and culture* (pp. 39–67). Cambridge University Press. https://doi.org/10.1017/CBO9780511752902.003

Spinrad, T. L., Losoya, S. H., Eisenberg, N., Fabes, R. A., Shepard, S. A., Cumberland, A., Guthrie, I. K., & Murphy, B. C. (1999). The Relations of Parental Affect and Encouragement to Children's Moral Emotions and Behaviour. *Journal of Moral Education*, *28*(3), 323–337. https://doi.org/10.1080/030572499103115

Starks, B., & Robinson, R. V. (2005). Who Values the Obedient Child Now? The Religious Factor in Adult Values for Children, 1986-2002. *Social Forces*, *84*(1), 343–359. https://doi.org/10.1353/sof.2005.0130

Statistics Canada. (2017, February 8). *Census Profile, 2016 Census—Nova Scotia [Province] and Canada [Country]*. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/Page.cfm?Lang=E&Geo1=PR&Code1=12&Geo2=&Code2=&SearchText=Nova%20Scotia&SearchType=Begins&SearchPR=01&B1=All&GeoLevel=PR&GeoCode=12&type=0

Stephens, J. M. (2018). Bridging the Divide: The Role of Motivation and Self-Regulation in Explaining the Judgment-Action Gap Related to Academic Dishonesty. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00246

Stevenson, R. J., Oaten, M. J., Case, T. I., Repacholi, B. M., & Wagland, P. (2010). Children's response to adult disgust elicitors: Development and acquisition. *Developmental Psychology*, *46*(1), 165–177. https://doi.org/10.1037/a0016692

Stevenson, R., Oaten, M., Case, T., & Repacholi, B. (2014). Is Disgust Prepared? A Preliminary Examination in Young Children. *The Journal of General Psychology*, *141*(4), 326–347. https://doi.org/10.1080/00221309.2014.938720

Suddendorf, T. (2003). Early Representational Insight: Twenty-Four-Month-Olds Can Use a Photo to Find an Object in the World. *Child Development*, *74*(3), 896–904. JSTOR.

Suhler, C. L., & Churchland, P. (2011). *Can Innate, Modular "Foundations" Explain Morality? Challenges for Haidt's Moral Foundations Theory*. *23*(9), 15.

Surian, L., & Franchin, L. (2017). Toddlers Selectively Help Fair Agents. *Frontiers in Psychology*, *8*, 944. https://doi.org/10.3389/fpsyg.2017.00944

Tajfel, H., & Turner, J. (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). Brooks/Cole.

Triandis, H., Bontempo, R., Villareal, M., Asai, M., & Lucca, N. (1988). Individualism and Collectivism: Cross-Cultural Perspectives on Self-Ingroup Relationship. *Journal of Personality and Social Psychology*, *54*, 323–338. https://doi.org/10.1037/0022-3514.54.2.323

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, *46*(1), 35–57.

Troseth, G. L., & DeLoache, J. S. (1998). The Medium Can Obscure the Message: Young Children's Understanding of Video. *Child Development*, *69*(4), 950–965. JSTOR. https://doi.org/10.2307/1132355

Turiel, E. (1983). *The Development of Social Knowledge. Morality and Convention*. Cambridge University Press.

Turiel, E. (2008). Thought about actions in social domains: Morality, social conventions, and social interactions. *Cognitive Development*, *23*(1), 136–154. https://doi.org/10.1016/j.cogdev.2007.04.001

Ulber, J., Hamann, K., & Tomasello, M. (2015). How 18- and 24-month-old peers divide resources among themselves. *Journal of Experimental Child Psychology*, *140*, 228–244. https://doi.org/10.1016/j.jecp.2015.07.009

Vaish, A., Carpenter, M., & Tomasello, M. (2010). Young Children Selectively Avoid Helping People With Harmful Intentions: Selectivity in Helping. *Child Development*, *81*(6), 1661–1669. https://doi.org/10.1111/j.1467-8624.2010.01500.x

Vaish, A., Missana, M., & Tomasello, M. (2011). Three-year-old children intervene in third-party moral transgressions: Children intervene in moral transgressions. *British Journal of Developmental Psychology*, *29*(1), 124–130. https://doi.org/10.1348/026151010X532888

van Leeuwen, F., & Park, J. H. (2009). Perceptions of social dangers, moral foundations, and political orientation. *Personality and Individual Differences*, *47*(3), 169–173. https://doi.org/10.1016/j.paid.2009.02.017

van Leeuwen, F., Park, J. H., Koenig, B. L., & Graham, J. (2012). Regional variation in pathogen prevalence predicts endorsement of group-focused moral concerns. *Evolution and Human Behavior*, *33*(5), 429–437. https://doi.org/10.1016/j.evolhumbehav.2011.12.005

Vauclair, C.-M., & Fischer, R. (2011). Do cultural values predict individuals' moral attitudes? A cross-cultural multilevel approach. *European Journal of Social Psychology*, *41*(5), 645–657. https://doi.org/10.1002/ejsp.794

Wachs, T. D., Gurkas, P., & Kontos, S. (2004). Predictors of preschool children's compliance behavior in early childhood classroom settings. *Journal of Applied Developmental Psychology*, *25*(4), 439–457. https://doi.org/10.1016/j.appdev.2004.06.003

Walker, L. J., & Hennig, K. H. (1999). Parenting Style and the Development of Moral Reasoning. *Journal of Moral Education*, *28*(3), 359–374. https://doi.org/10.1080/030572499103133

Walker, L. J., Hennig, K. H., & Krettenauer, T. (2000). Parent and peer contexts for children's moral reasoning development. *Child Development*, *71*(4), 1033–1048. https://doi.org/10.1111/1467-8624.00207

Walker, L. J., & Taylor, J. H. (1991). Family interactions and the development of moral reasoning. *Child Development*, *62*(2), 264–283. https://doi.org/10.2307/1131002

Warneken, F., Lohse, K., Melis, A. P., & Tomasello, M. (2011). Young Children Share the Spoils After Collaboration. *Psychological Science*, *22*(2), 267–273. https://doi.org/10.1177/0956797610395392

Wendorf, C. A. (2001). History of American morality research, 1894–1932. *History of Psychology*, *4*(3), 272–288. https://doi.org/10.1037/1093-4510.4.3.272

White, F. A., & Matawie, K. M. (2004). Parental Morality and Family Processes as Predictors of Adolescent Morality. *Journal of Child and Family Studies*, *13*(2), 219–233. https://doi.org/10.1023/B:JCFS.0000015709.81116.ce

Williams, R. N., & Gantt, E. E. (2012). Felt moral obligation and the moral judgement–moral action gap: Toward a phenomenology of moral life. *Journal of Moral Education*, *41*(4), 417–435. https://doi.org/10.1080/03057240.2012.665587

Wittig, M., Jensen, K., & Tomasello, M. (2013). Five-year-olds understand fair as equal in a mini-ultimatum game. *Journal of Experimental Child Psychology*, *116*(2), 324–337. https://doi.org/10.1016/j.jecp.2013.06.004

Yang, C., Crain, S., Berwick, R. C., Chomsky, N., & Bolhuis, J. J. (2017). The growth of language: Universal Grammar, experience, and principles of computation. *Neuroscience & Biobehavioral Reviews*, *81*, 103–119. https://doi.org/10.1016/j.neubiorev.2016.12.023

Young, R. D., & Avdzej, A. (1979). The effects of obedience/disobedience and obese/nonobese body type on social acceptance by peers. *The Journal of Genetic Psychology: Research and Theory on Human Development*, *134*(1), 43–49. https://doi.org/10.1080/00221325.1979.10533397

Appendix A: Moral Foundations Questionnaire

Moral Foundations Questionnaire

Part 1. When you decide whether something is right or wrong, to what extent are the following considerations relevant to your thinking? Please rate each statement using this scale:

[0] = not at all relevant (This consideration has nothing to do with my judgments of right and wrong)
[1] = not very relevant
[2] = slightly relevant
[3] = somewhat relevant
[4] = very relevant
[5] = extremely relevant (This is one of the most important factors when I judge right and wrong)

1. _____Whether or not someone suffered emotionally
2. _____Whether or not some people were treated differently than others
3. _____Whether or not someone's action showed love for his or her country
4. _____Whether or not someone showed a lack of respect for authority
5. _____Whether or not someone violated standards of purity and decency
6. _____Whether or not someone was good at math
7. _____Whether or not someone cared for someone weak or vulnerable
8. _____Whether or not someone acted unfairly
9. _____Whether or not someone did something to betray his or her group
10. _____Whether or not someone conformed to the traditions of society
11. _____Whether or not someone did something disgusting
12. _____Whether or not someone was cruel
13. _____Whether or not someone was denied his or her rights
14. _____Whether or not someone showed a lack of loyalty
15. _____Whether or not an action caused chaos or disorder
16. _____Whether or not someone acted in a way that God would approve of

Part 2. Please read the following sentences and indicate your agreement or disagreement:

[0]	[1]	[2]	[3]	[4]	[5]
Strongly disagree	Moderately disagree	Slightly disagree	Slightly agree	Moderately agree	Strongly agree

17. _____Compassion for those who are suffering is the most crucial virtue.
18. _____When the government makes laws, the number one principle should be ensuring that everyone is treated fairly.
19. _____I am proud of my country's history.
20. _____Respect for authority is something all children need to learn.
21. _____People should not do things that are disgusting, even if no one is harmed.
22. _____It is better to do good than to do bad.
23. _____One of the worst things a person could do is hurt a defenseless animal.
24. _____Justice is the most important requirement for a society.
25. _____People should be loyal to their family members, even when they have done something wrong.
26. _____Men and women each have different roles to play in society.
27. _____I would call some acts wrong on the grounds that they are unnatural.
28. _____It can never be right to kill a human being.
29. _____ I think it's morally wrong that rich children inherit a lot of money while poor children inherit nothing.
30. _____ It is more important to be a team player than to express oneself.
31. _____ If I were a soldier and disagreed with my commanding officer's orders, I would obey anyway because that is my duty.
32. _____ Chastity is an important and valuable virtue.

The Moral Foundations Questionnaire (full version, July 2008) by Jesse Graham, Jonathan Haidt, and Brian Nosek.

For more information about Moral Foundations Theory and scoring this form, see: www.MoralFoundations.org

Scoring: Item scores are summed for each foundation to create total scores. Items 1, 7, 12, 17, 23, and 28 are care/harm. Items 2, 8, 13, 18, 24, and 29 are fairness/cheating. Items 3, 9, 14, 19, 25, and 30 are loyalty/disloyalty. Items 4, 10, 15, 20, 26, and 31 are authority/subversion. Items 5, 11, 16, 21, 27, and 32 are sanctity/purity.

Appendix B: Demographics Questionnaire

Please have one parent/guardian complete the following survey and return it to your child's daycare with the signed consent form. Any parent/guardian can complete the survey. You can stop or refrain from answering a question on this survey at any time.

Demographic Questionnaire

Child's Date of Birth: month_____ day _____ year _____

With which gender do you (the parent/guardian) identify?

___Female
___Male
___Transgender male
___Transgender female
___Agender/non-binary
___Other: _____

What is the gender of your child?

___Female
___Male
___Transgender male
___Transgender female
___Agender/non-binary
___Other: _____

Which category best describes your total family annual income (before deductions)?

___Less than $25000
___$25000-less than $40000
___$40000-less than $75000
___$75000-less than $100000
___$100000 or more

What is your highest level of education attained?

___Some high school
___High school diploma or equivalent
___Some college, trade school, or equivalent
___Some university
___Diploma from college, trade school, or equivalent
___Bachelor or undergraduate degree
___Some graduate school
___Master's (e.g., MA, M.Sc., M.Ed)
___Degree in medicine, dentistry, veterinary medicine

___Earned doctorate
___Other: _____

How many siblings does your child have?

___None
___One
___Two
___Three
___Four
___More than four

What is your cultural background, race, or ethnicity? Select all that apply.

___White
___Black or African Nova Scotian
___Hispanic, Latino, or Spanish
___Asian
___Asian Indian
___Hawaiian Native
___Pacific Islander
___Middle Eastern
___Alaskan Native
___First Nations, Inuit, or Metis
___Other: _____

Which category best describes your religious orientation?

___Christian
___Muslim
___Buddhist
___Hindu
___Jewish
___Sikh
___Atheist
___Agnostic
___Non-religious
___Other: _____

Which category best describes your political orientation?

___Liberal
___Conservative
___Moderate
___Socialist
___Libertarian
___Environmentalist
___Other: _____

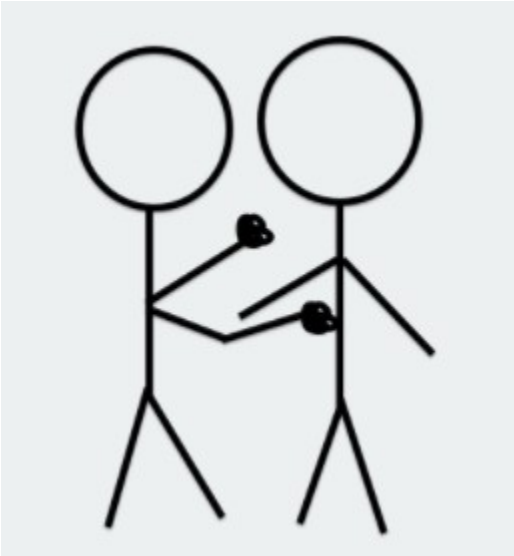
Appendix C: MFQK Items and Pictures

Male Items

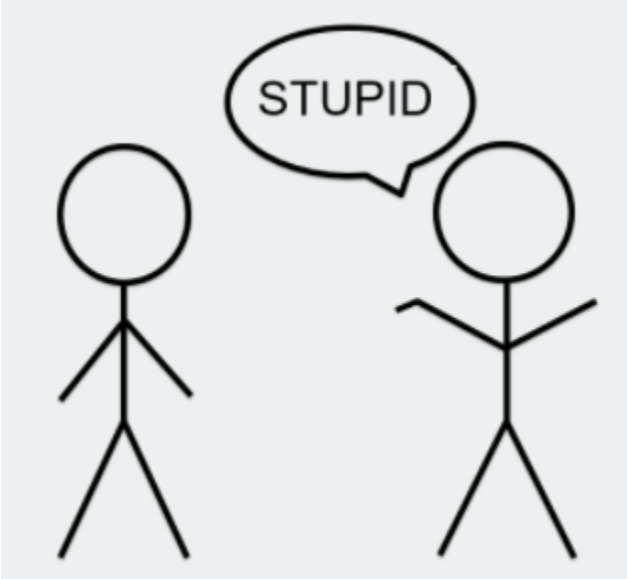
I. CARE/HARM

A. Harm to Human

1. You see a boy punch another boy in the stomach.

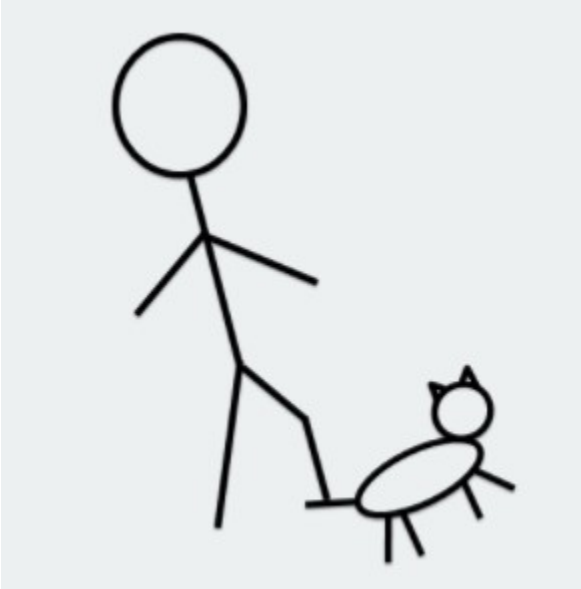


2. You see a boy calling a boy stupid.

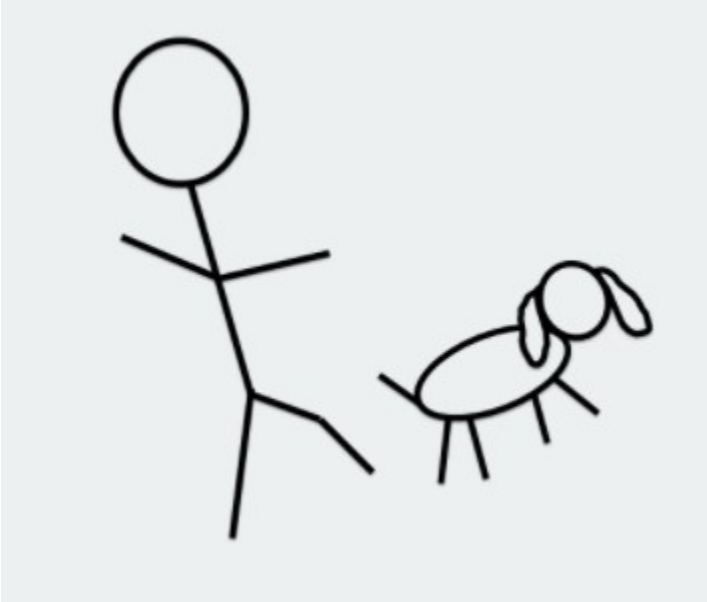


B. Harm to Animal

3. You see a boy stomp on the tail of his pet cat.



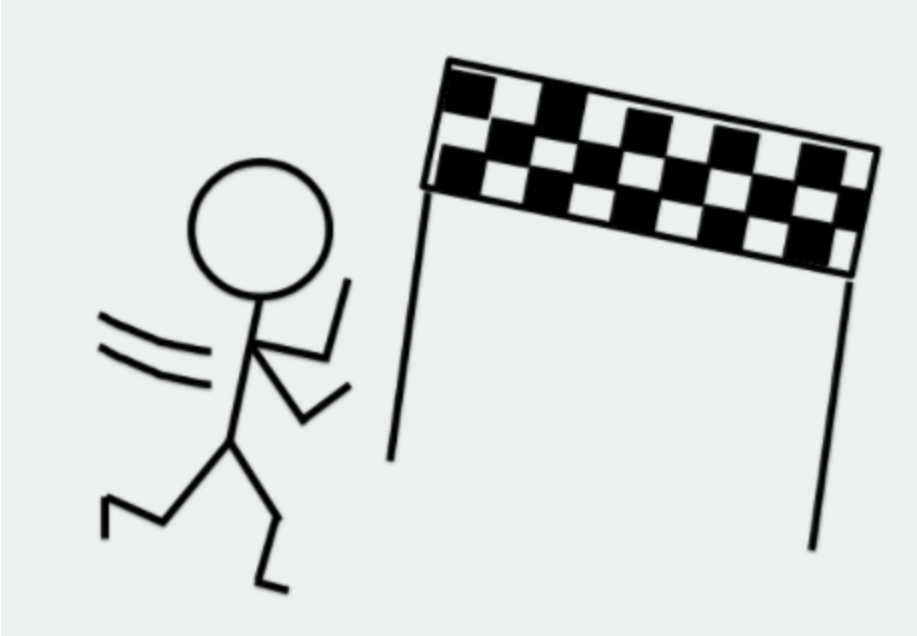
4. You see a boy kick a stray dog



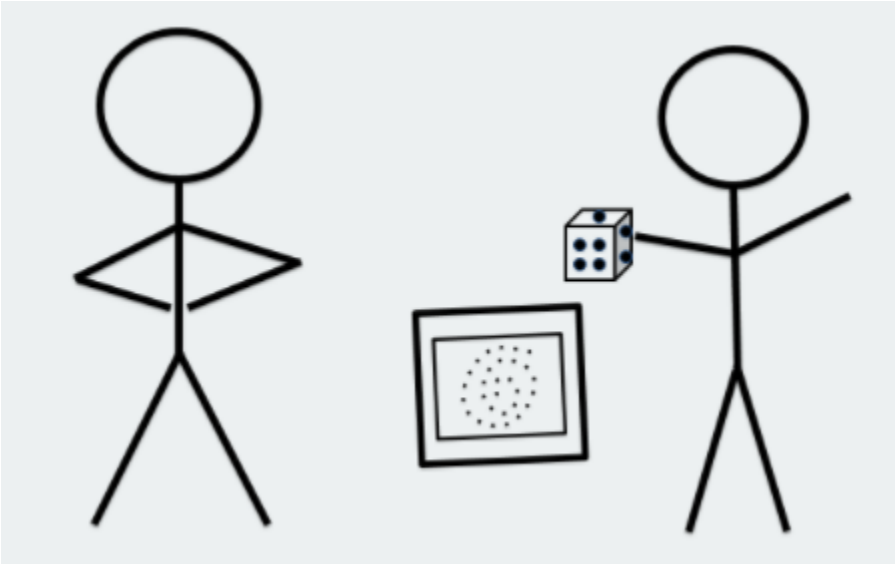
II. FAIRNESS

A. Cheating

5. You see a boy cheating in a race by taking a shortcut.

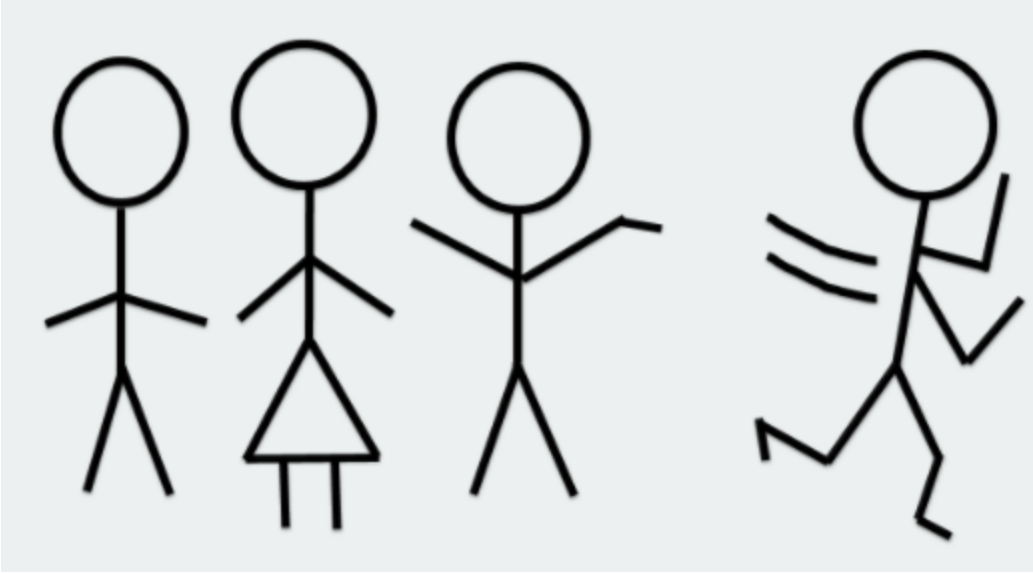


6. You see a boy cheating in a board game.

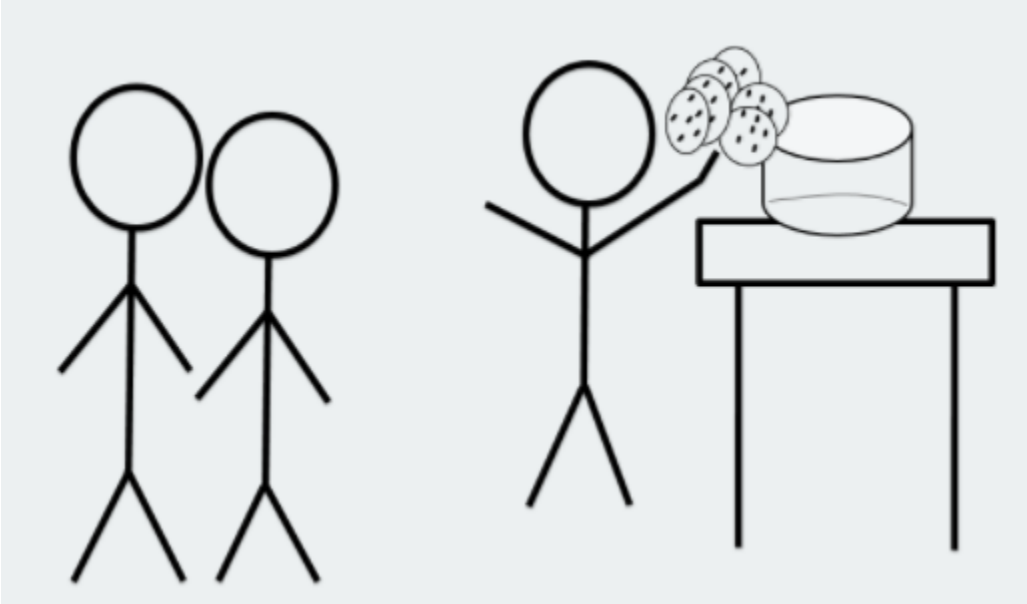


B. Inequality

7. You see a boy cut to the front of the line.



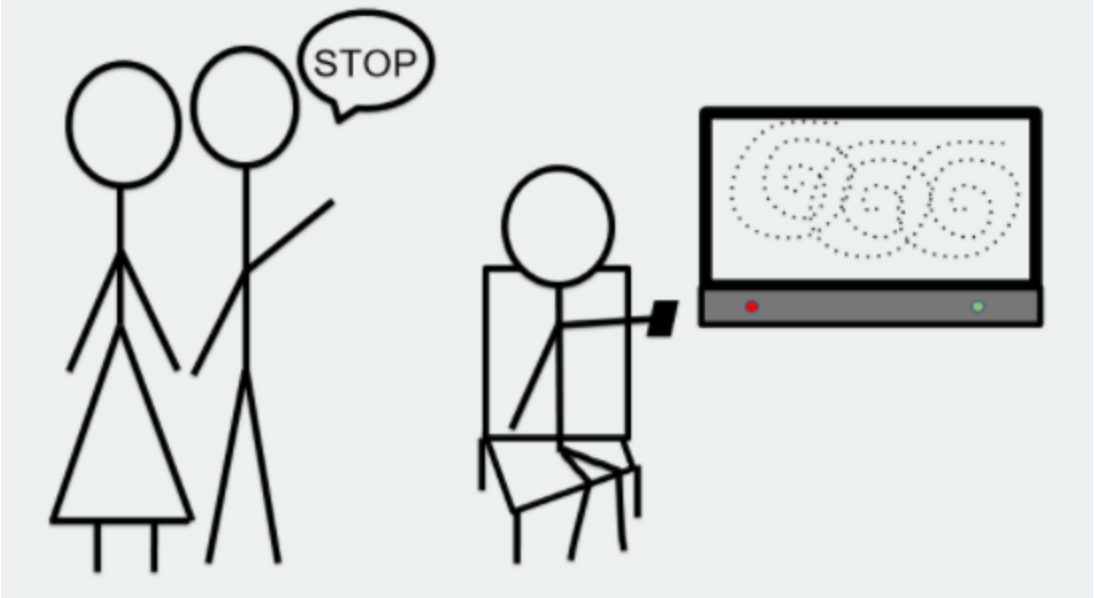
8. You see a boy taking all of the cookies, and leaving none for others.



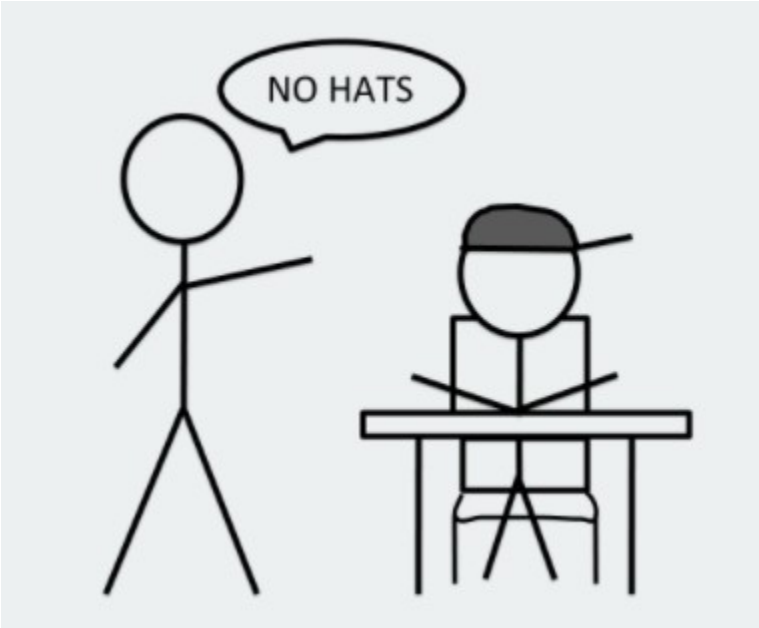
III. AUTHORITY

A. Disobedience

9. You see a boy ignore his parents when they tell him to stop watching TV.

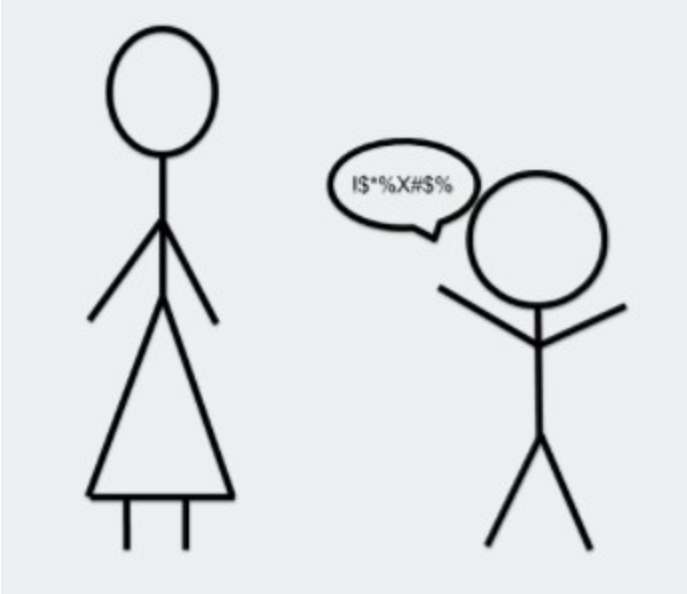


10. You see a boy wearing a hat at school, even after the teacher asks him not to.

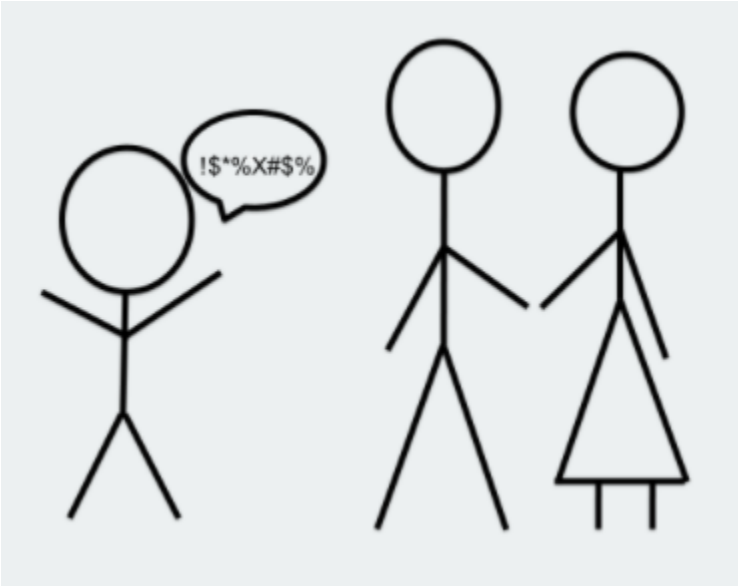


B. Disrespect

11. You see a boy calling his teacher bad words.



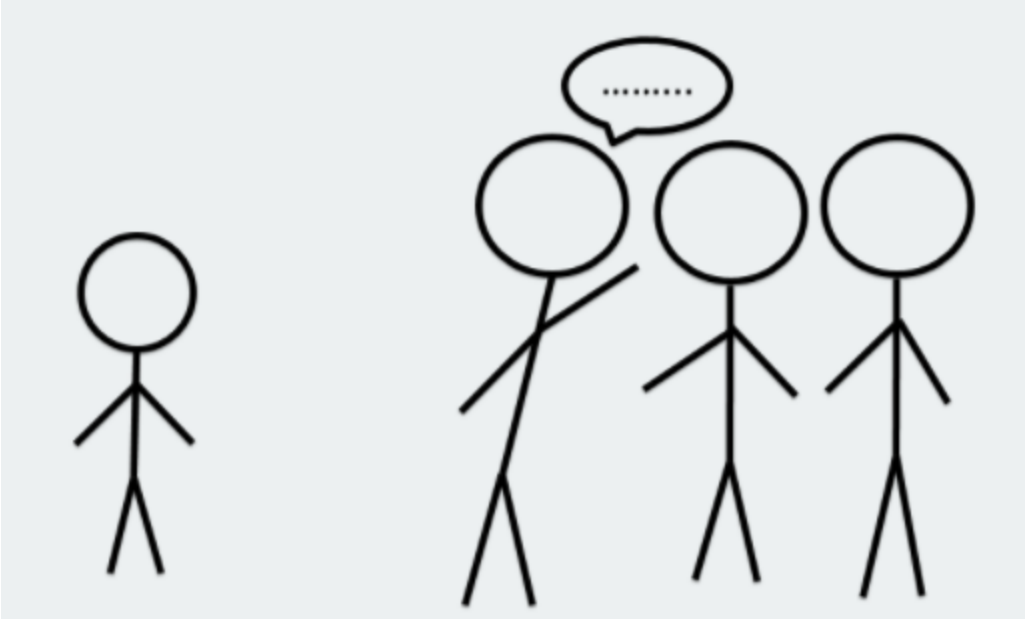
12. You see a boy calling his parents bad words.



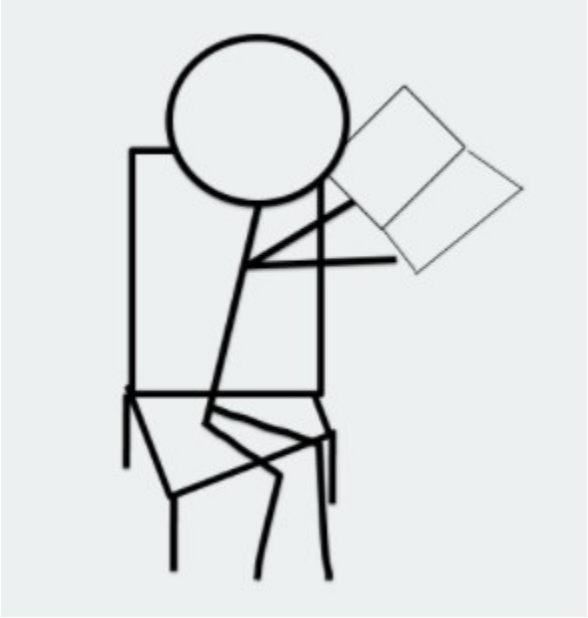
IV. LOYALTY

A. Disloyalty to Sibling

13. You see a boy telling secrets about his brother to people his brother doesn't like.

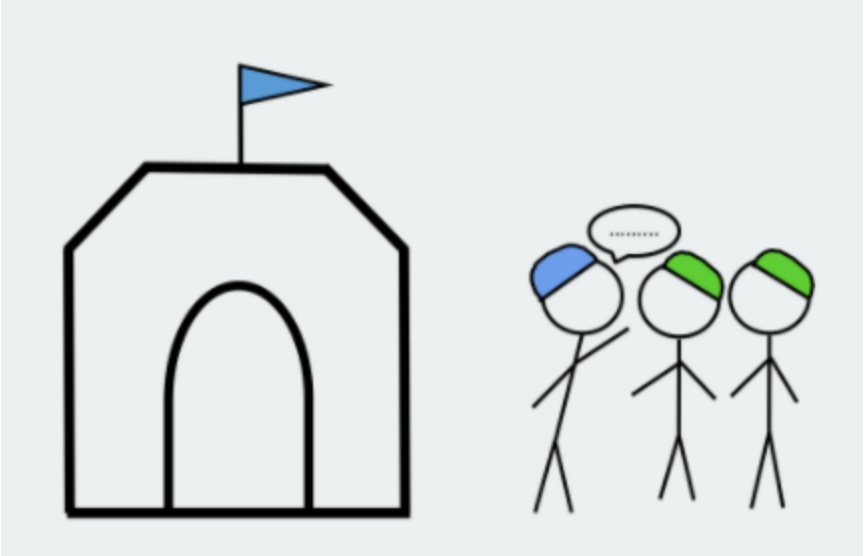


14. You see a boy reading his brother's diary.

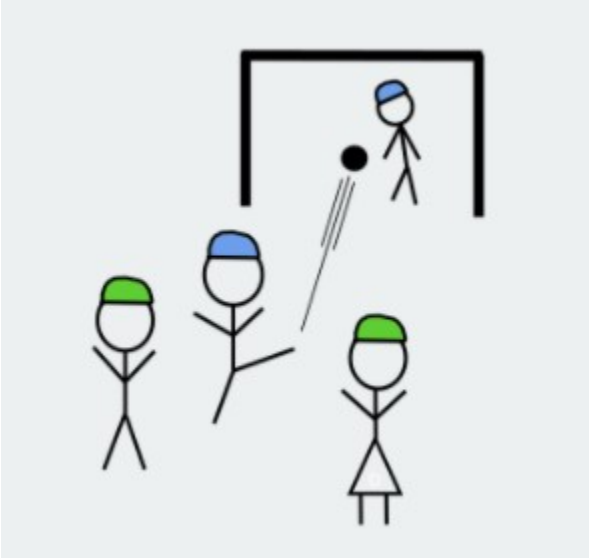


B. Disloyalty to Group

15. You see a boy teach a secret password to people who are not in his club.



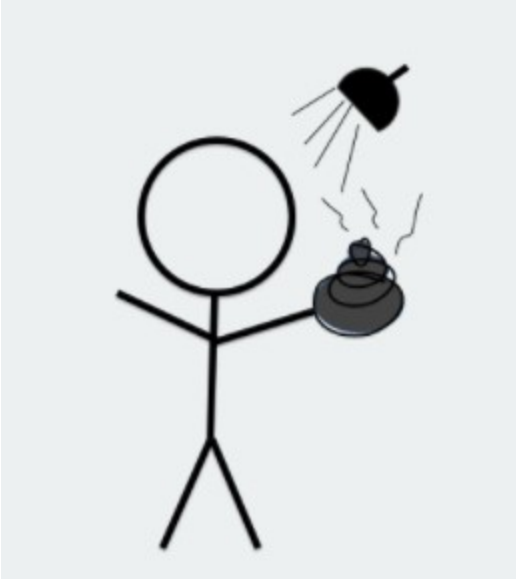
16. You see a boy score a goal against his own team to help the other team win.



V. SANCTITY/PURITY

A. Hygiene

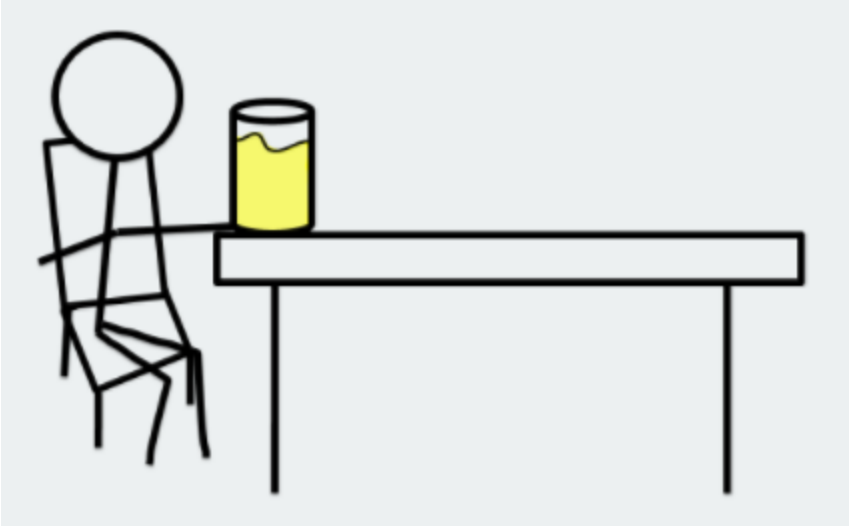
17. You see a boy rubbing poop on himself in the shower.



18. You see a boy using a dirty diaper as a pillow.

B. Food

19. You see a boy drinking pee with his dinner.

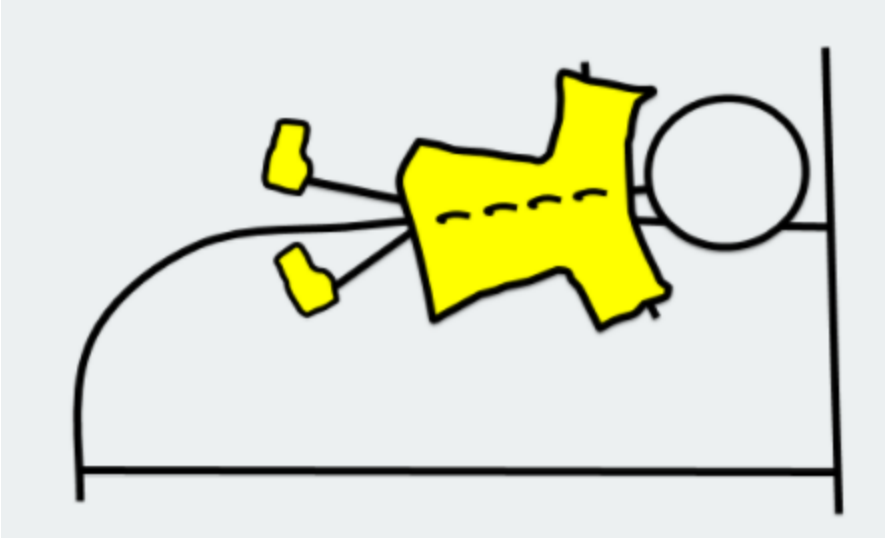


20. You see a boy loudly burping and farting while eating.

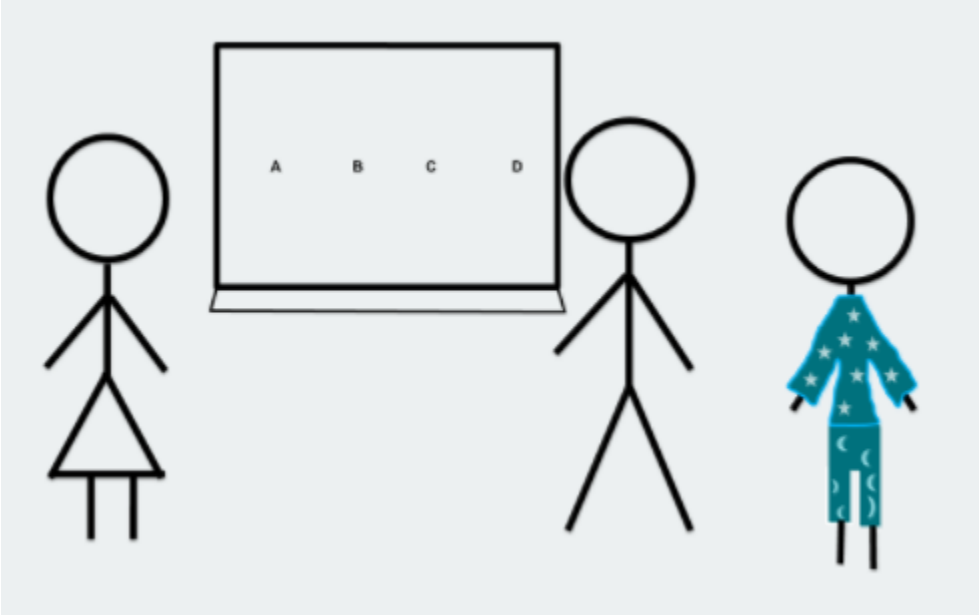
VI. ODD BEHAVIOUR

A. Clothing

21. You see a boy sleeping in his raincoat, instead of sleeping in pajamas.

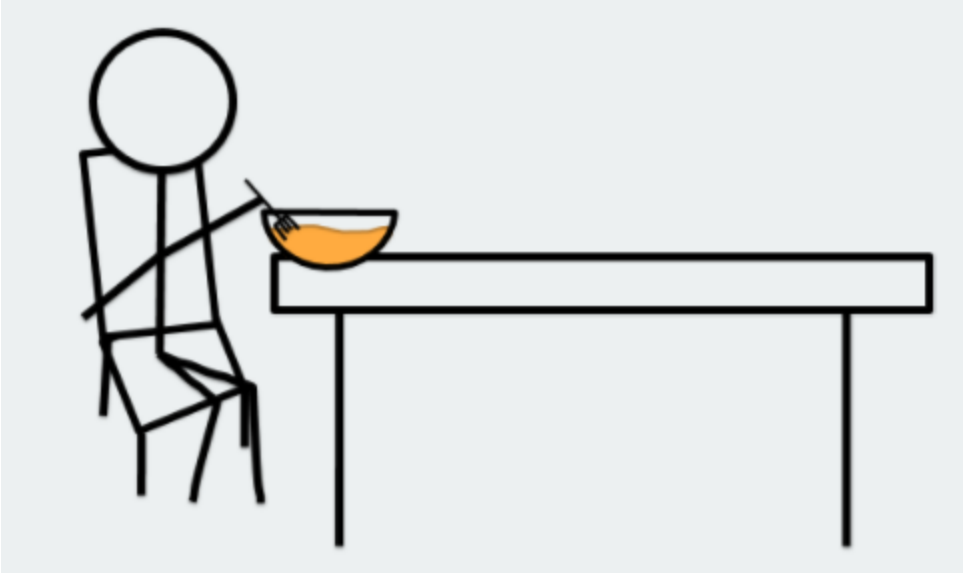


22. You see a boy wearing his pajamas to school.

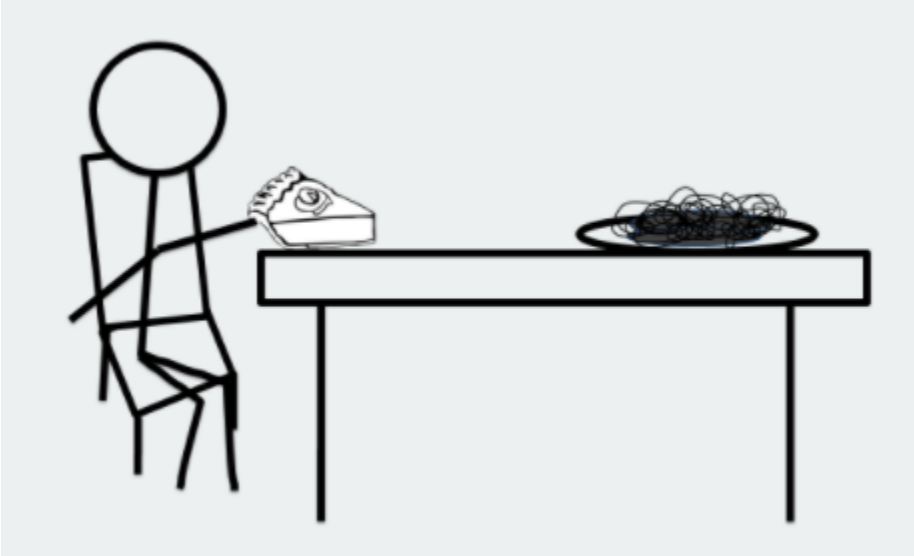


B. Food

23. You see a boy eating his soup with a fork.



24. You see a boy eat dessert before dinner is served, instead of eating it afterwards.

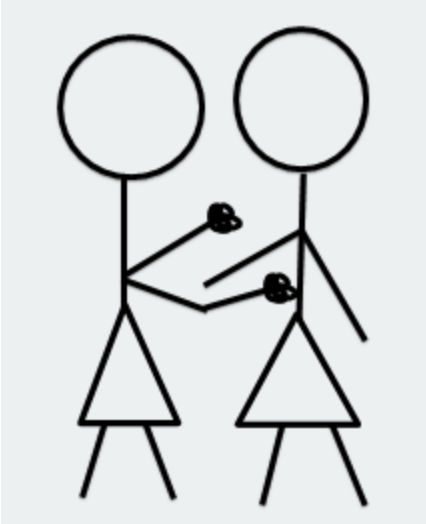


Female Items

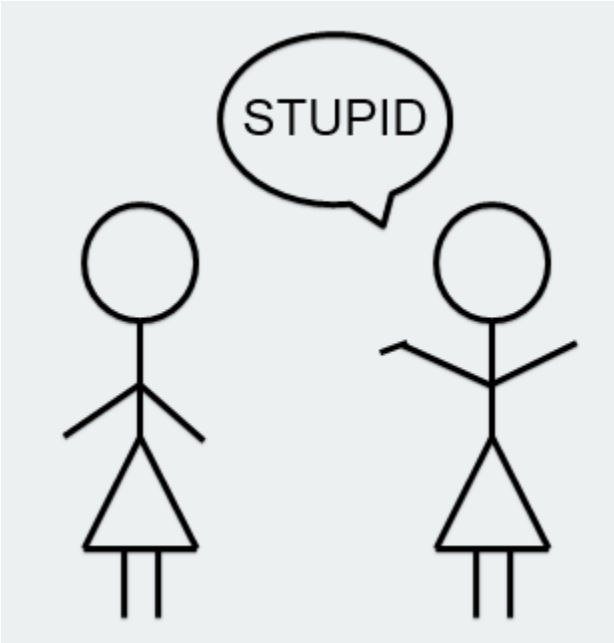
I. CARE/HARM

A. Harm to Human

1. You see a girl punch another girl in the stomach.

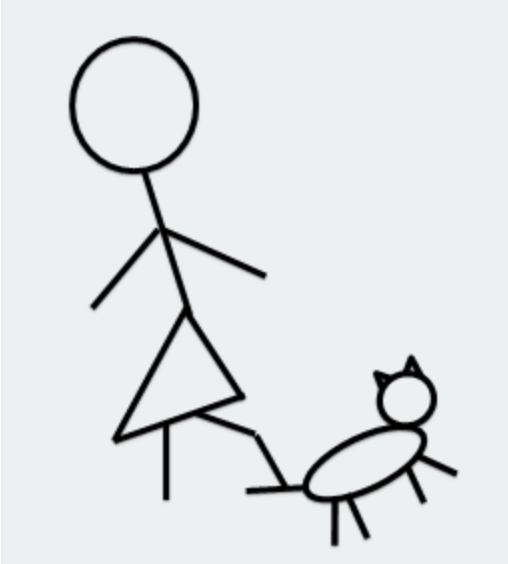


2. You see a girl calling a girl stupid.

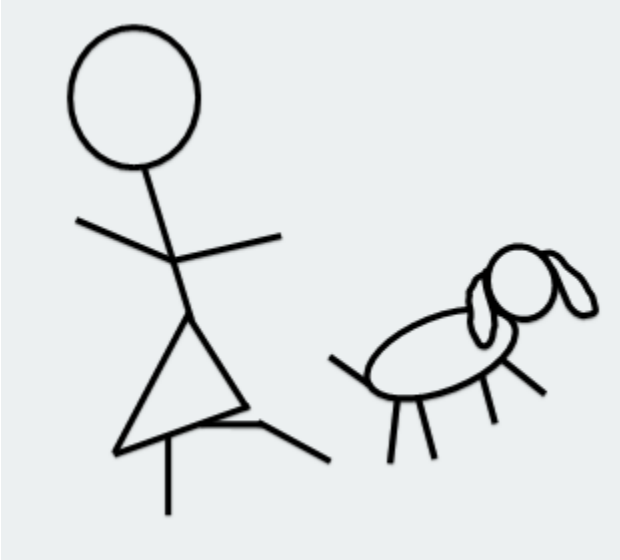


B. Harm to Animal

3. You see a girl stomp on the tail of her pet cat.



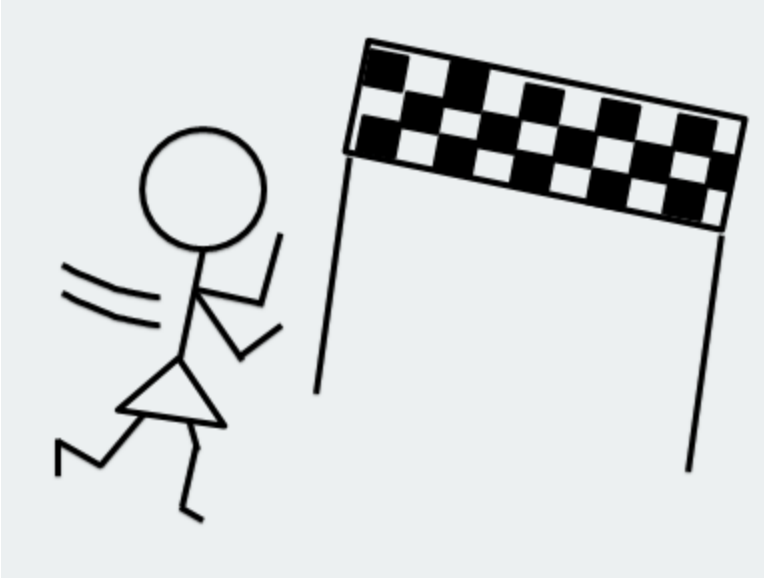
4. You see a girl kick a stray dog.



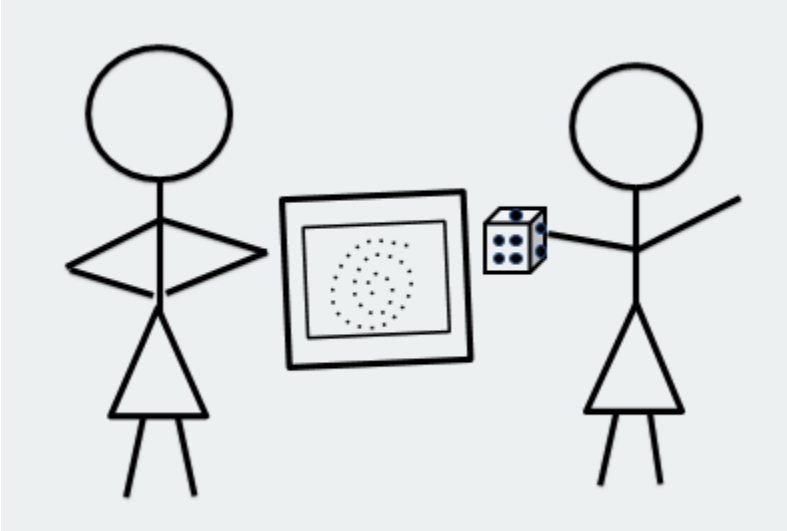
II. FAIRNESS

A. Cheating

5. You see a girl cheating in a race by taking a shortcut.

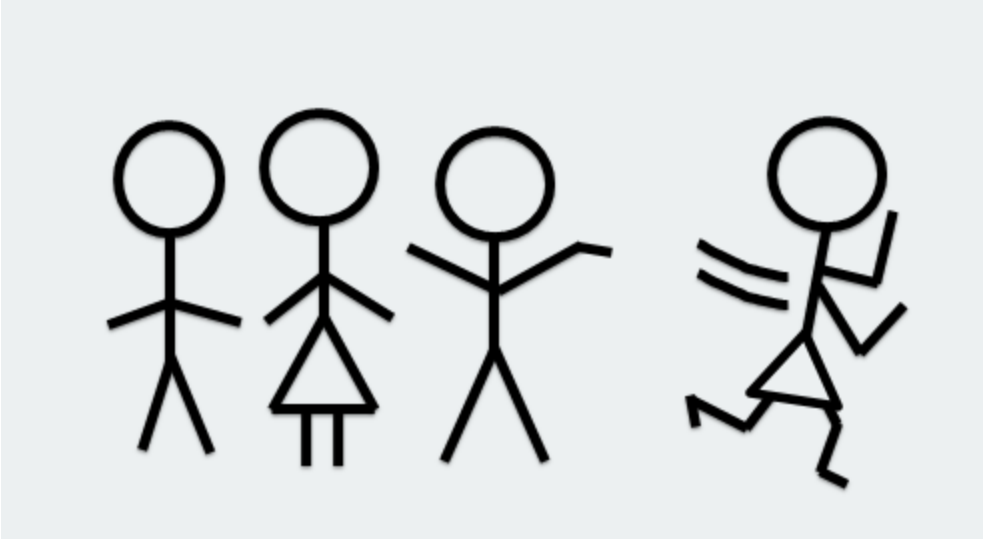


6. You see a girl cheating in a board game.

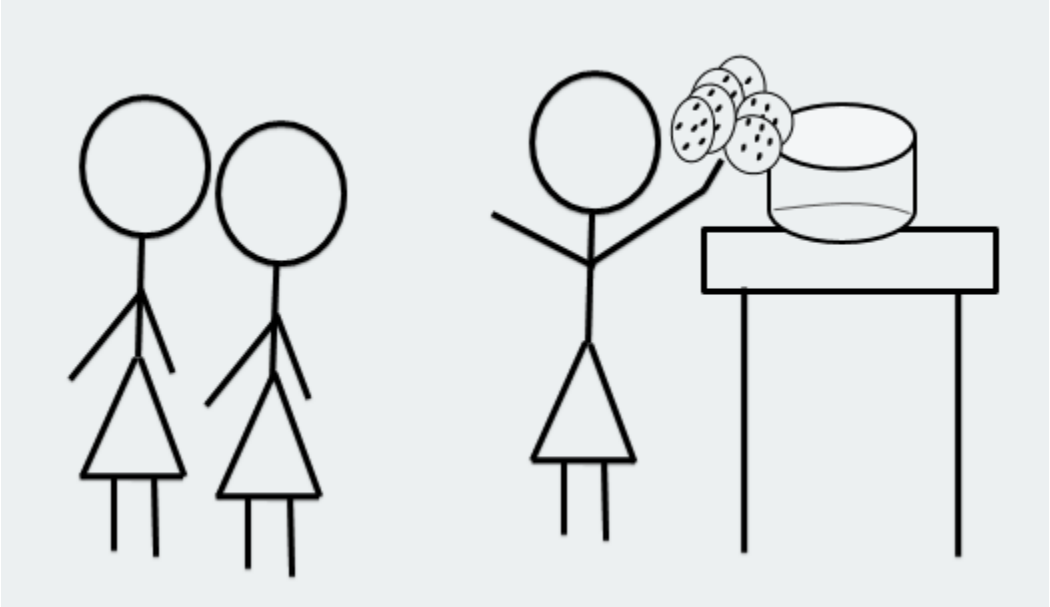


B. Inequality

7. You see a girl cut to the front of the line.



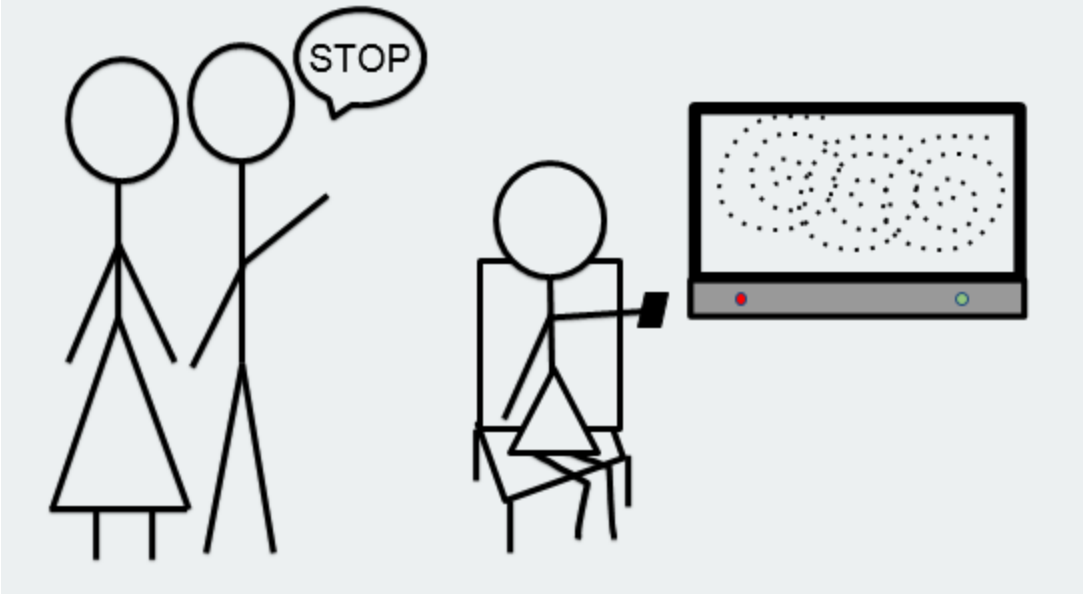
8. You see a girl taking all of the cookies, and leaving none for others.



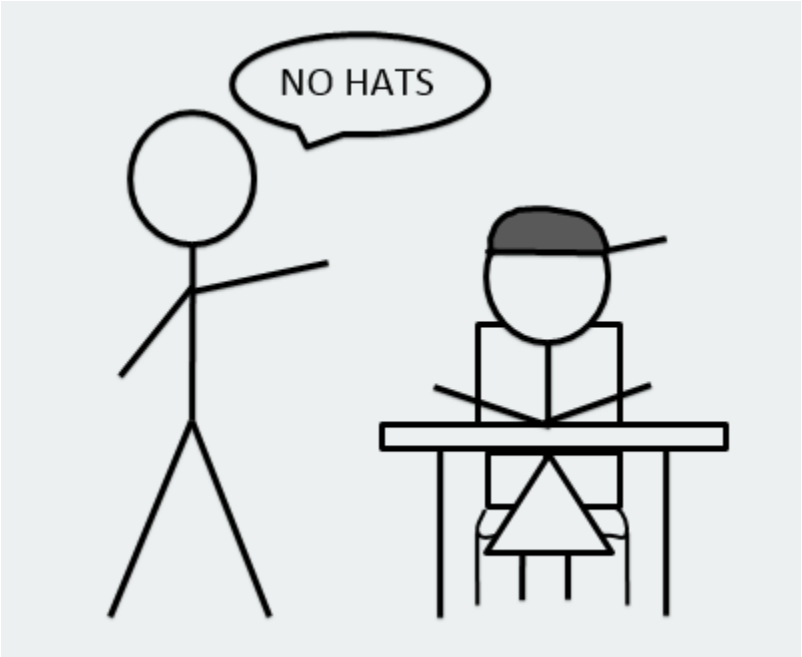
III. AUTHORITY

A. Disobedience

9. You see a girl ignore her parents when they tell her to stop watching TV.

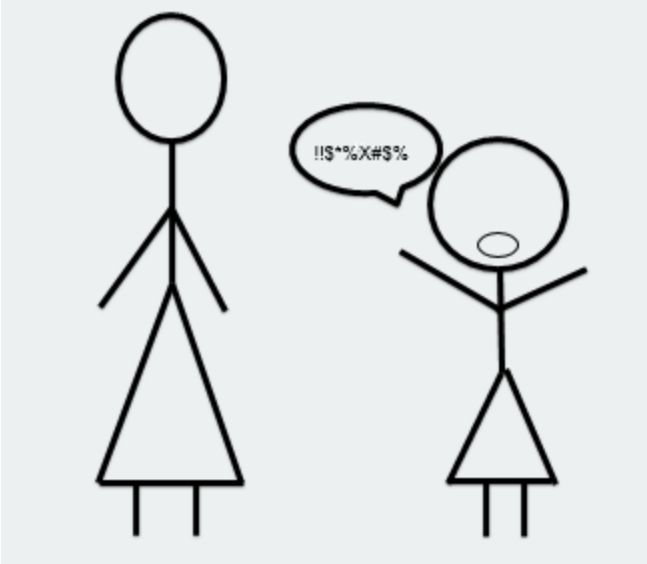


10. You see a girl wearing a hat at school, even after the teacher asks her not to.

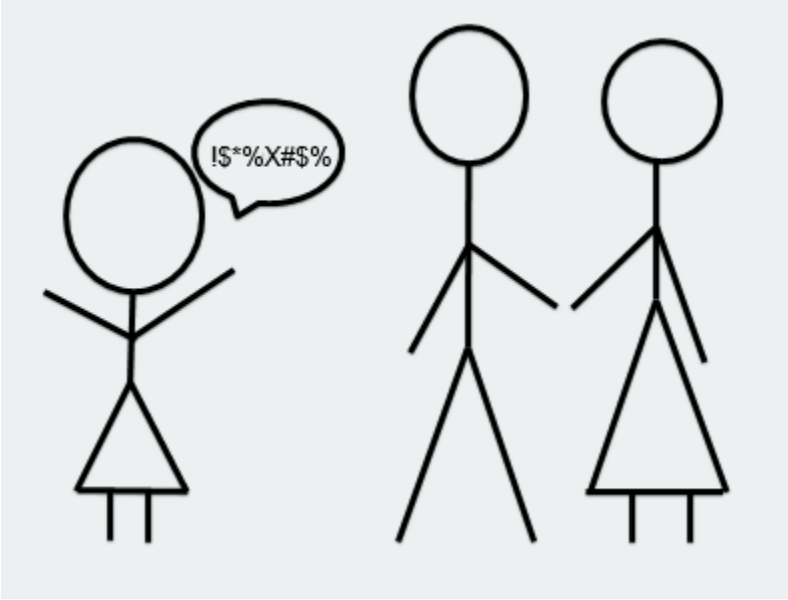


B. Disrespect

11. You see a girl calling her teacher bad words.



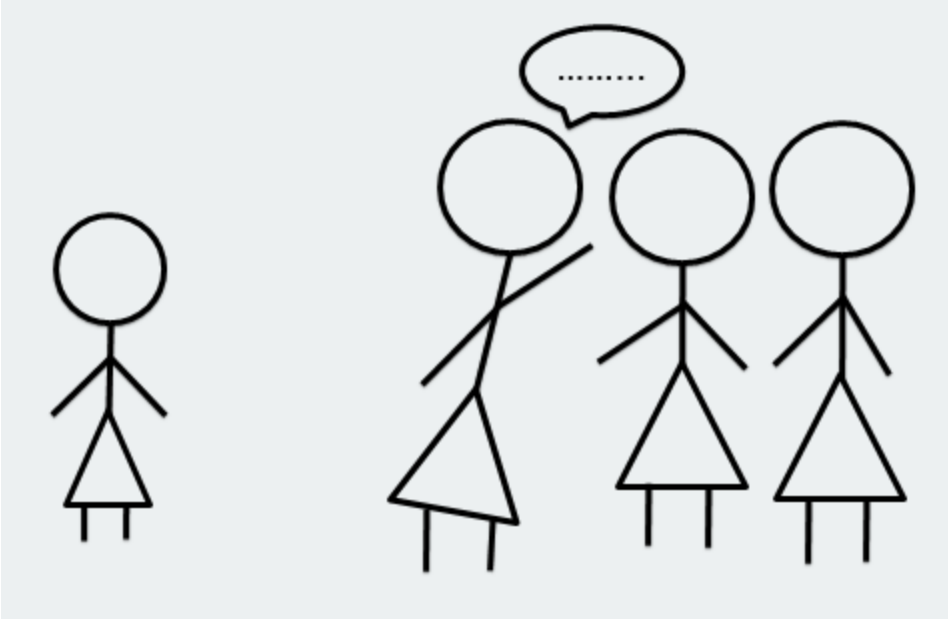
12. You see a girl calling her parents bad words.



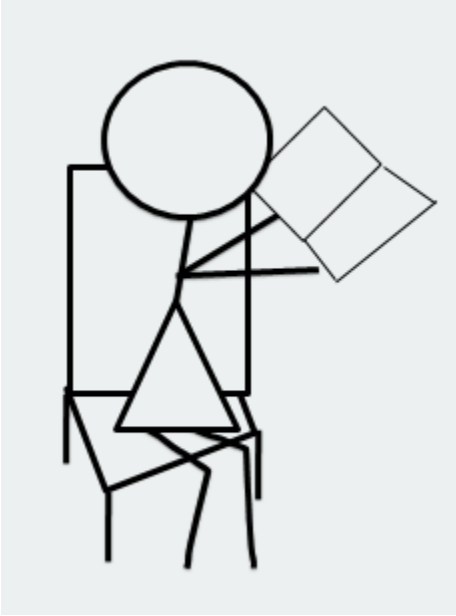
IV. LOYALTY

A. Disloyalty to Sibling

13. You see a girl telling secrets about her sister to people her sister doesn't like.

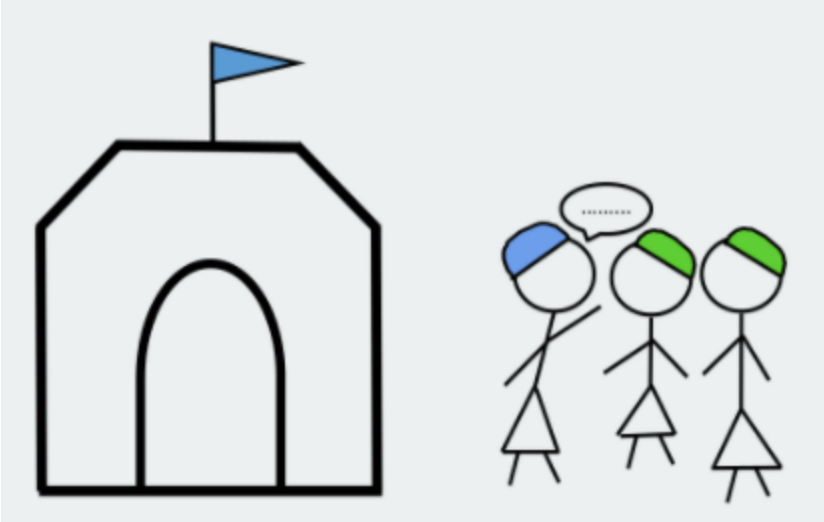


14. You see a girl reading her sister's diary.

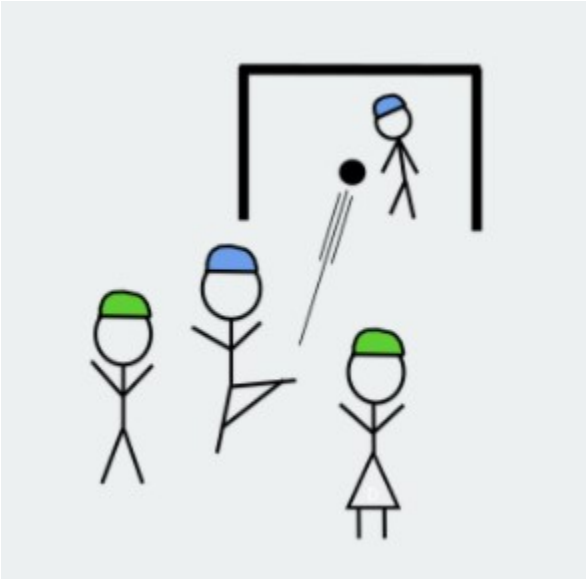


B. Disloyalty to Group

15. You see a girl teach a secret password to people who are not in her club.



16. You see a girl score a goal against her own team to help the other team win.

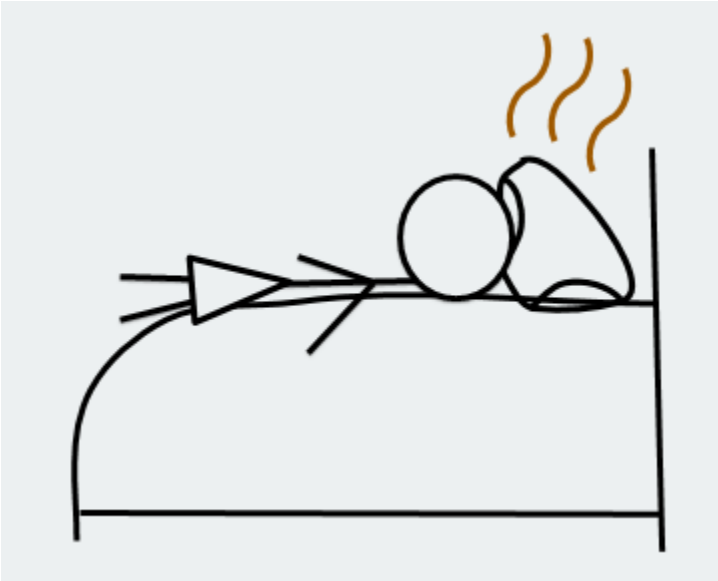


V. SANCTITY/PURITY

A. Hygiene

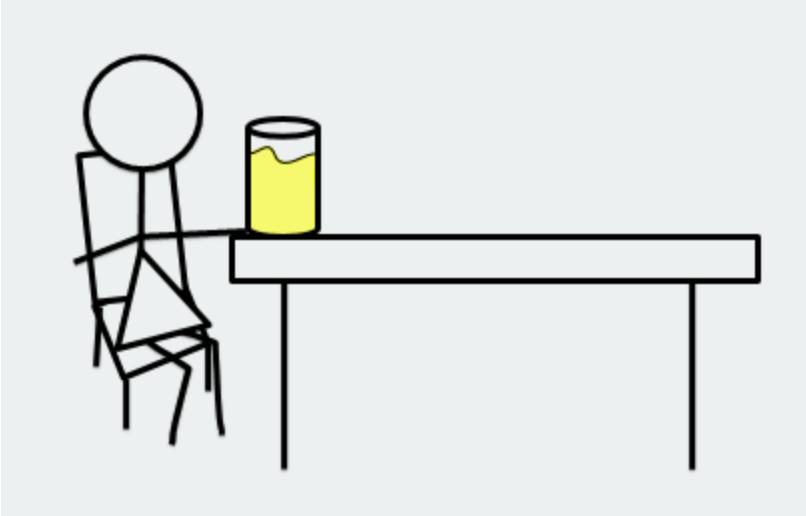
17. You see a girl rubbing poop on herself in the shower.

18. You see a girl using a dirty diaper as a pillow.

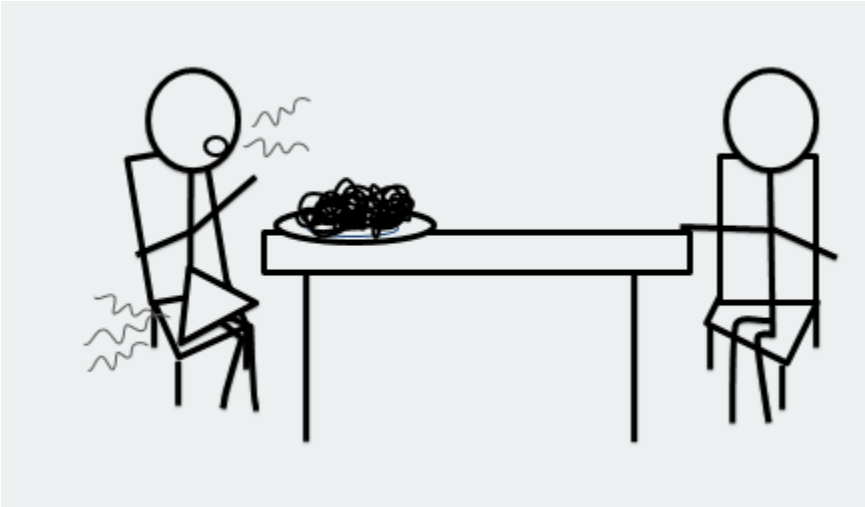


B. Food

19. You see a girl drinking pee with her dinner.



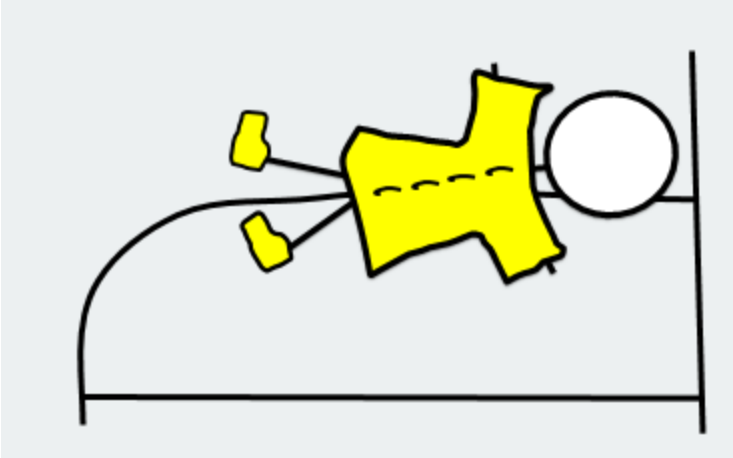
20. You see a girl loudly burping and farting while eating.



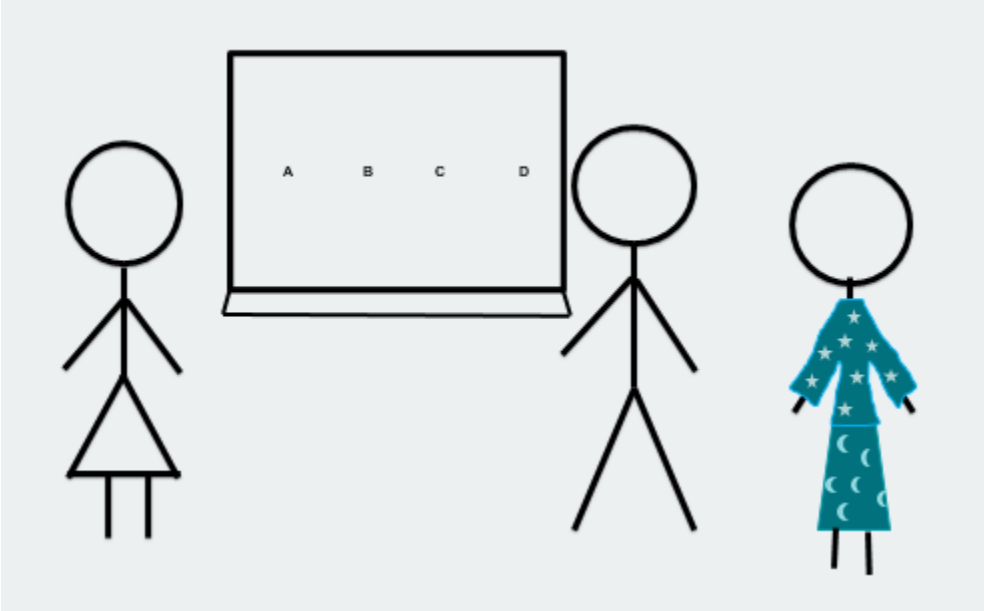
VI. ODD BEHAVIOUR

A. Clothing

21. You see a girl sleeping in her raincoat, instead of sleeping in pajamas.

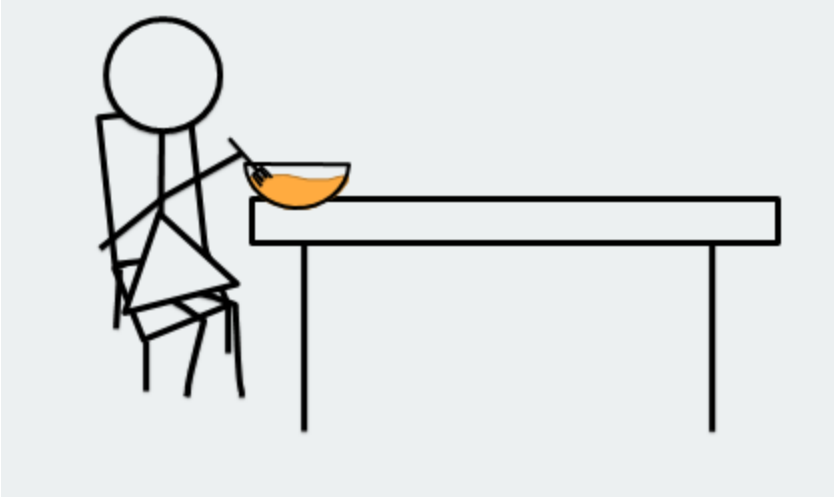


22. You see a girl wearing her pajamas to school.

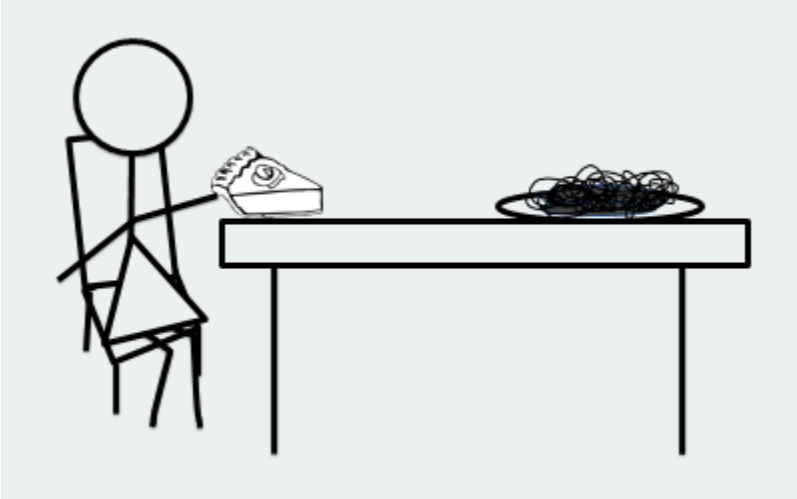


B. Food

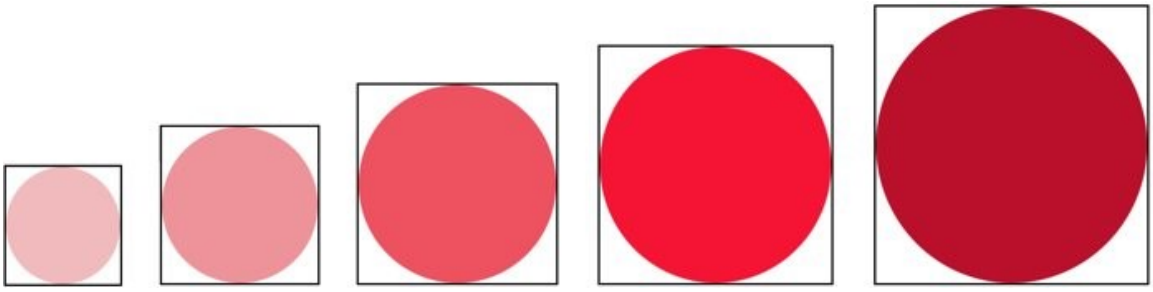
23. You see a girl eating her soup with a fork.



24. You see a girl eat dessert before dinner is served, instead of eating it afterwards.



Appendix D: MFQK Severity Scale



If children rated an MFQK item as bad, they were presented with this scale and asked, "how bad is it?". During the training phase, children were taught that circles, from biggest to smallest, represented 'not very bad', 'a little bad', 'quite bad', 'very bad', and 'very very bad'.

Appendix E: MFQK Emotion Scale

For each item of the MFQK, children were asked to indicate how the depicted action made them feel (happy, sad, angry, yucky, or neutral).

Appendix F: MFPT Reliability Coding

Percent agreement between original coding and third-party video coding of participant responses on the MFQK, by foundation and scenario

Foundation	Scenario	Question	# Code-able Videos	Percent Agreement	Average Percent Agreement
Care	1	Q1	130	99.2	92.9
		Q2	130	96.2	
		M1	131	87.8	
		M2	129	88.4	
	2	Q1	132	97.7	96.0
		Q2	133	97.0	
		M1	133	96.2	
		M2	132	93.2	
Fairness	1	Q1	131	98.5	96.7
		Q2	132	99.2	
		M1	129	94.6	
		M2	129	94.6	
	2	Q1	128	98.4	97.0
		Q2	126	97.6	
		M1	127	96.9	
		M2	126	95.2	
Loyalty	1	Q1	125	97.6	97.6
		Q2	123	97.6	
		M1	125	98.4	
		M2	126	96.8	
	2	Q1	122	99.2	96.7
		Q2	120	97.5	
		M1	121	94.2	
		M2	122	95.9	
Authority	1	Q1	123	95.1	95.7
		Q2	123	95.1	
		M1	121	95.9	
		M2	121	96.7	
	2	Q1	123	91.1	90.2
		Q2	122	90.2	
		M1	121	89.3	
		M2	122	90.2	
Sanctity	1	Q1	130	99.2	98.1
		Q2	130	99.2	
		M1	129	96.9	
		M2	129	96.9	
	2	Q1	131	97.7	96.6
		Q2	132	99.2	
		M1	133	94.7	
		M2	131	94.7	

Note. Q1 = 'Which puppet do you like best?'; Q2 = 'Which puppet did something bad?'; M1 = manipulation check 1, M2 = manipulation check