# Trust and Skepticism:
# Children's Use and Evaluations of Moral Testimony Across Two Cultures

A Dissertation
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Pearl Han Li

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Adviser:
Melissa A. Koenig

AUGUST 2022

**Acknowledgements**

I would like to express my deepest gratitude to my advisor, Melissa Koenig. I feel extremely lucky that you took me on as a student five years ago. Thank you for challenging my intellectual thinking in your own thoughtful way, for teaching me your contemplative style of interrogating research questions, for being incredibly encouraging towards me as I grappled with balancing family and work as a new parent, and for making yourself available intellectually, physically, and emotionally whenever I'm in need. I will never forget how you shared words of encouragement before my first talk at a conference, how you celebrated each of my small accomplishments, how you skipped your meeting with all the important people just to chat with me for four hours in the San Diego sun, how you threw me the best baby shower in your favorite park, and all of our writing sessions/lunches in the summer and walks along the lakes. Our relationship means the world to me.

Also thanks to my committee members, Stephanie Carlson, Gail Ferguson, Valerie Tiberius and Andrew Zieffler, for your encouragement, guidance, kindness and support. Thanks to Dr. Tiberius for introducing me to the moral philosophy world, showing a genuine interest in my work, and for offering invaluable feedback in your extremely warm and humble way. To Dr. Ferguson for always showing sheer and sincere kindness towards me, and for pushing me to have a critical eye regarding the bias in the literature towards Western and euro-centric contexts. To Dr. Carlson for greatly contributing to my growth of confidence when I coped with PhD impostor syndrome by kindly complimenting on my writing and encouraging me to speak up in your graduate

courses. I also owe it to Dr. Zieffler for choosing a minor in Quantitative Methods. Thanks for convincing me to divorce SPSS and for making stats fun and unintimidating (which I think is one of the hardest jobs in the world).

I am extremely grateful to have had the great pleasure to work with everyone in the Early Language and Experience Lab. My special thanks go to the lab managers and research assistants who have helped with these projects, this dissertation would not be possible without your help: George Budzius, Yara Ghazal, Joanne Jung, Maeve Kinderman, Kelly Nydahl, Steven Reaves, Sara Wilkerson, and Rendahui Wu. A big warm thanks also goes to the amazing team of Chinese researchers who helped with these projects in summer 2021, Luk Shu Chan, Viola Ruoyu Fei, Jun Fu, Bingru Li, Jinghui Li, Xiaomeng Liu, Chunlu Yang, Ying Wang, Chenzi Wei, Yi Zhao, and Andi Zhu. Thank you to my lab mates, Sarah Suárez, Katherine Ridge, Annelise Pesch, Erika DeAngelis and Norwood Glaspie. Special thanks to Narges Afshordi, for your friendship, kindness, wisdom and support. My thanks also go to Hannah Allen and Daiqing Zhao for helping with the stimuli in these projects.

I am extremely fortunate to have been a part of my amazing cohort – Carolyn, Faith, Isa, Sophie and Hayley. Special thanks to Faith VanMeter, for getting through all the stats assignments together, for sharing a sense of self-deprecating humor and social anxiety to speak up in class, for being there for all my successes and failures, for all the home cooked meals and presents to spoil my baby, and for being a beautiful human being you are. To my friend and other officemate Hayley Brigman, for your incredible support and kindness, for all the boba we consumed together, for being in my COVID pod at the start of the pandemic, and for all your prayers for my wellbeing and success. I look

forward to lots of visits and sleepovers after we move to North Carolina. To Sophie, for your support, brilliance and friendship.

I would also like to thank Paul Harris, for introducing me to the world of developmental research, teaching me to use the most simple and clear experimental designs, convincing me to come to ICD and offering support throughout my graduate studies. You planted the seeds of curiosity that greatly contributed to this dissertation project. My thanks also go to Samuel Ronfard for his generosity, kindness and mentorship. Thanks for helping me through my very first research project and encouraging me to apply to graduate school. Our coffee chats were always the highlights of my conference experience throughout the years. To my brilliant friend, Angela Weiyi Wang, for bringing home a job ad and encouraging me to apply to my very first research assistantship in developmental psychology, for visiting me while I was pregnant to help set up the nursery, and for always being one of my biggest advocates. To my best friends in college, Bolin Bai and Xidi Yan, for your love and support throughout the years, and to Danruo Zhong and Qiyue Cai for our support group and for all your help with local recruitment.

Thanks to the University of Minnesota for awarding me a Doctoral Dissertation Fellowship and the Developing Beliefs Network for awarding me the Predoctoral Dissertation Fellowship, providing me with the financial means to complete this project.

Finally, I'd like to thanks to my husband, Yanning Jia, for moving all the way from Beijing to Minnesota for my Ph.D., and for supporting me to do everything I put my mind to. To my family, for always offering support and love across the globe and for

going out of their ways to help me disseminate my study recruitment materials. Special thanks to my mom, Weihong Liang, who gave me my English name (Pearl) and taught me my first English words with the limited resources that she had. Thank you for coming all the way to MN against all odds to take care of your grandson so I could finish writing the dissertation, and for making me the person I am today. To Charlie, thank you for being my perfect excuse to procrastinate, and for bringing me more joy than I could ever have imagined possible.

This dissertation is dedicated to my Grandma, Guizhi Jia, for being my strong, fearless female role model, and for always watching over me and our loved ones.

## Abstract

While information from other people is an important source of factual knowledge, little is understood about the role of testimony in children's acquisition of moral knowledge (Harris et al., 2018). In fact, many psychologists argue that morality is relatively impervious to direct testimonial influence, and have privileged independent moral abilities and intuitions (Hamlin, 2013; Smetana, 1981). These accounts are consistent with recent philosophical work holding that while testimony is an acceptable source for non-moral beliefs, it is problematic to acquire moral understanding via testimony (Hopkins, 2007). Is testimony really insufficient to transmit moral understanding? In two cross-cultural studies, I assessed how children and adults balance independent thinking and reliance on testimony, and how cultural mechanisms contribute to people's moral agency and moral learning.

In Study 1, I explored children's moral judgments in the context of countervailing explanations from an adult, which are either (1) supported by reasoning about the consequences, or which (2) emphasize authority-related considerations. 379 3- to 5-year-old children and 180 adults from the U.S. and China listened to stories in which the protagonists engaged in actions that make others cry, referred to only by means of unfamiliar words such as "mibbing". Then, depending on the condition, participants were exposed to an adult speaker who provides Utilitarian explanations (e.g., "Mibbing is good because Devon can play with the new toy after mibbing."), Authoritarian explanations (e.g., "Mibbing is good because I get to decide whether mibbing is good."), or Control statements (repeating basic information in the story with no moral judgment). Afterwards, I asked for participants' own moral judgments and explanations. While adults

were not persuaded by the speaker's counter-intuitive claims, children in both countries were significantly influenced by the two types of counterintuitive explanations, and judged the novel, distress-inducing actions to be more morally permissible. With age, children from both countries became less receptive towards Authoritarian explanations in which the adult speaker exerts power and authority over children.

In Study 2, I examined individuals' own evaluations of different sources of moral and empirical knowledge across cultures. 261 4- to 6-year-old children and 163 adults in the United States and China were randomly assigned to one of two between-subjects conditions, a Moral Knowledge condition (i.e., in which agents make judgments about moral actions such as pushing or helping others) or an Empirical Knowledge condition (i.e., in which agents make judgments about the hidden contents of containers). In each condition, participants were introduced to two speakers, one who showed independence in her thinking, and the other who relied on testimony to make moral or empirical judgments. Children were then asked to indicate the agent with the best way of thinking (e.g., "Who has the best reason to know that…?") and provide justifications for their choices. I found that Chinese children were more likely to judge independent thinking as a better reason, and showed a stronger preference for the independent agent regardless of condition. By contrast, U.S. children were less selective, and considered independent deliberation and reliance on testimony as equally legitimate ways to justify one's knowledge. Unlike children, adults from both countries distinguished between the two types of judgments, and chose reliance on testimony at significantly lower than chance levels in the Moral Knowledge condition, but not in the Empirical condition.

Together, my results indicate that although moral testimony might be considered a less legitimate source of moral knowledge by adults, children depend greatly on others' testimony when optimizing their learning in moral contexts. These results also extend our understanding of the cultural specificity and generality of the role played by testimony and authority in children's moral development.

# Table of Contents

## List of Tables

## List of Figures

## CHATPER 1: Background

Children and adults get around in the world in large part by relying on the testimony of others, making it hard to imagine how the accumulation of scientific, historical, and conventional knowledge could have happened without testimony. Empirically, many of the testimonial learning studies in the developmental literature also focus on the transmission and acceptance of factual knowledge such as the name of an object, categorization of animals, or the existence of germs, even if they cannot easily verify these claims themselves (Harris & Koenig, 2006; Lane, 2018, for reviews, see Harris et al., 2018; Mills, 2013). Where does that leave us with respect to other kinds of knowledge? Are there any limits to the types of knowledge that can be transmitted by testimony?

Although testimony seems to be an undisputed source of empirical knowledge, some have argued that it is not legitimate or even possible to acquire moral knowledge on the basis of testimony alone. Rather, they argue that moral beliefs need to be acquired and justified, at least to some extent, on the basis of one's own independent reflection and reasoning (Gelfert, 2014). Metaphorically, this means that moral agency requires individuals to drive their own bus, making it epistemically and morally wrong for moral agents to be passengers who allow their own moral judgments to be fully established by someone else. Thus, it is less clear whether it is legitimate or even possible to rely on testimony in the moral domain. Indeed, many developmental and social psychologists have argued that core aspects of moral understanding are authority independent (Smetana, 2006), innate (Hamlin, 2013), or independently acquired through intuitions or reasoning (Haidt & Joseph, 2007; Paxton, & Greene, 2010; Kant, 1785/1959). If any of these positions are true, then many sources of social influence, including testimony,

might be insufficient to transmit moral knowledge or alter core aspects of moral understanding. However, none of these theoretical perspectives have discussed the role of testimony in detail (Harris, 2012).

One view in contemporary Western philosophy, *the asymmetry thesis*, does specify the role that testimony cannot take in the moral domain. This view holds that while testimony is a perfectly acceptable source for non-moral beliefs in other domains, it is problematic for a listener to base her moral beliefs by exclusively deferring to testimony (Hopkins, 2007). Unless an agent has their own understanding of the reasons and principles that underlie their obligations, the claim is that moral testimony cannot do more than help fine-tune, extend, or better specify the applications of those principles (Howell, 2014; Jones, 1999; Nickel, 2001). Does this claim hold for young children?

Below, I draw on the arguments for this philosophical view in order to examine the case for skepticism toward moral testimony as a source of moral knowledge. I also review empirical work on children's social learning and moral development, in the hopes of advancing our understanding of testimony's role in children's acquisition and use of moral knowledge. My central argument is that the theories on the autonomy of moral thinking are largely overstated. Children actually rely heavily on moral testimony as they are navigating the moral world as developing moral agents. Note that most of the research described here has been done in North America except where noted, underscoring the importance of the cross-cultural efforts I propose to undertake.

**Appreciating the Limits of Moral Testimony?**

Work in experimental philosophy shows that testimony is treated as a second-class citizen in moral thinking by adults. When adults on Amazon MTurk were asked to evaluate the legitimacy of forming a moral belief on the basis of testimony versus first-

hand experience (e.g., "Suppose that [a friend/ first-hand experience] tells you that a particular action is immoral. Is it *legitimate* to rely solely on what [the friend / firsthand experience] tells you when forming your own opinion about whether the action is immoral?"]), beliefs formed on the basis of a friend's testimony were judged to be less legitimate than beliefs about descriptive matters such as the size of a painting (Andow, 2019). Why do adults have the intuition that moral beliefs based on testimony are inferior or less legitimate compared to testimonially-based empirical or factual beliefs?

These intuitions about the suspicious nature of second-hand moral knowledge are consistent with the philosophical tradition that moral understanding must be "self-given" through the exercise of one's own reasoning (Kant, 1785/1959). In the past two decades, a line of work in Western philosophy specifically addressed the issues of testimony, arguing that dependence on testimony for moral knowledge is epistemically unattainable or morally problematic. In fact, an *asymmetry thesis* has been proposed, arguing that there are interesting principled differences between moral and non-moral testimony. Specifically, this view holds that while testimony is a perfectly acceptable source for non-moral beliefs in other domains, it is problematic for a listener to base her *moral beliefs* on testimony (e.g., Hopkins, 2007). Philosophers who argue for the asymmetry thesis are sometimes known as *moral testimony pessimists*. To a certain degree, pessimists consider relying on testimony to be inferior to judging for oneself (Crisp, 2014), and argue that moral beliefs formed on the basis of another's testimony are in poor epistemic or moral standing (e.g., Nickel, 2000; McGrath, 2011). Below, I will overview the main pessimistic arguments about dependence on moral testimony, which can be broadly classified into two categories: (1) Epistemic doubts regarding testimony's ability to

transmit moral understanding, and (2) moral reservations related to whether deferring to moral testimony threatens our ability to exercise good moral agency.

Epistemically, moral testimony pessimists hold that testimony cannot transmit true moral understanding (e.g., Hopkins, 2007; Nickel, 2000; McGrath, 2009, 2011). While it can be relatively easy to acquire empirical knowledge in various empirical domains, their position is that it is difficult to acquire true moral understanding via testimony. For example, the *understanding argument* (Nickel, 2001) contends that moral understanding is holistic and includes the ability to build networks of facts and justifications (Hills, 2009). In this respect, holding a moral belief not only involves agreeing to a proposition (e.g., one should become a vegetarian), but also requires a grasp of the connections between moral beliefs and their justifications (Hills, 2009; Nickel, 2001; McGrath, 2011). In other words, if a listener comes to believe a moral claim based on the testifier's word, or on her authority, she may still fail to appreciate the reasons behind the claim, and the connections with other moral beliefs. Relatedly, failure to comprehend reasons after reliance on moral testimony may also prevent the listener from extending their judgment to other circumstances, or from applying moral claims at appropriate times (*extension argument*, Crisp, 2014; Nickel, 2001). Given the complex and subtle nature of many situations which moral agents face, it seems essential for one to have a flexible mastery of moral reasons. For example, an individual could be told that it is wrong to steal from one's neighbor. But based on that testimony alone, he would be ill-prepared to respond to changes that arise, and unable to draw the appropriate conclusion about stealing from a store to feed his family.

Other epistemic arguments against the reliance on testimony include the difficulties in identifying moral experts (Hills, 2009; Goldman, 2001; Driver, 2006; Matheson, McElreath, & Nobis, 2018; McGrath, 2009, 2011), and the considerable amount of disagreement about moral matters even among the experts (Hopkins, 2007) — both of which can threaten the transmission of moral beliefs. Another related argument, sometimes referred to as the *calibration worry*, holds that independent access to facts is not readily available in the realm of morality (McGrath, 2009), and this in turn prevents us from verifying the track record of potentially reliable testifiers. In both respects, we can see how the usual markers of expertise – relevant skills, participation in consensus and track record – seem to be difficult to judge in the moral domain.

So far, the skeptical arguments against moral testimony that I have covered focus on the independent epistemic agency that is necessary in the moral domain. Aside from these epistemic arguments that concern moral testimony's limits, the most prominent *moral argument* against the dependence on testimony builds on the *understanding argument*, and proposes that understanding is necessary for morally worthy action (Howell, 2014; Jones, 1999; McGrath, 2011; Hills, 2009, 2013). Note that while the previous arguments primarily concern the possibility of transmitting moral understanding, this argument questions deference to testimony because it is in tension with one's moral agency. Specifically, reliance on moral testimony is problematic because it runs against the central aspects of our moral agency such as moral worth and moral virtue. Some contend that what virtue requires is that the agent makes good judgments for herself rather than doing what she is told (e.g., McGrath, 2011; Hills, 2009, 2013; Nickel, 2001).

In summary, moral testimony pessimism is concerned both with *epistemic reasons* that cast doubt on the possibility of moral testimony transmitting moral knowledge, and *moral reason*s related to the moral worthiness of deferring to testimony. Although individual's moral evaluations of deference to testimony remain largely untapped in empirical work, the epistemic reasons against relying on moral testimony are generally consistent with psychological accounts that dominate discussions of moral development and moral thinking, which imply that the psychological processes governing children's moral decisions are either innate (e.g., Hamlin et al., 2007; Mikhail, 2007) or acquired through independent learning and reasoning (Piaget, 1965, Smetana, 2006; Rhodes & Wellman, 2017) ). In the next section, I use recent developmental evidence to evaluate the role of moral testimony on children's social learning, arguing that in contrast to these skeptical accounts, children actually rely heavily on testimony to acquire moral understanding.

**Children's Use of Moral Testimony**

In light of philosophical reservations about moral testimony as a source of knowledge (Hopkins, 2007; Jones, 1999), and given evidence that adults believe that reliance on testimony in the moral domain is more problematic than in the factual domain (Andow, 2019), what do children think? Admittedly, the skeptical arguments are more focused on moral testimony being problematic for mature moral agents who have capabilities for gaining moral understanding under their own power, and moral testimony pessimists have fewer issues with children depending on testimony to form moral beliefs (Hopkins, 2007; Hills, 2009; Lewis, 2020). However, none of the philosophical accounts make it clear why children's dependence is more acceptable than adults' dependence. Moreover, as I review in this section, the uniqueness of moral testimony in the context of

the developmental literature on moral development is still relevant. Namely, the developmental traditions and mainstream theories tend to prioritize children's autonomous moral thinking - arguing that moral decision-making can be treated as a primarily intuitive and independent endeavor. As a result, it is often implied that testimony plays a limited role in children's moral thinking when research has scarcely begun to explore how children evaluate testimony in the moral domain. In this section, I issue a call for research on this question and discuss reasons to think that children treat moral testimony seriously and rely heavily on others' moral testimony to acquire moral understanding.

***Moral Development as An Independent Endeavor***

Much developmental research has taken a constructivist or nativist approach, arguing that children either (1) independently develop moral competencies by actively assessing their social environment, or (2) are born with innate, nascent moral intuitions that are independent of social influences. As a dominant moral development theory over the past two decades, Social Domain Theory argues that preschoolers understand that the wrongness of basic transgressions (e.g., hitting, stealing) is not tied to how an adult authority reacts (Turiel,1983; Smetana,1981;1985; Smetana et al., 2012). Instead, children's understanding of moral norms is related to the intrinsic consequences of the acts for others (e.g., concerns for others' welfare, Smetana, 2006). Additionally, Social Domain Theory posits that children's welfare-based reasoning is largely constructed by directly and independently observing and experiencing the negative consequences (e.g., distress) caused by harmful actions as victims, transgressors or third-party observers of moral conflicts (Smetana, 2006). More recently, Rhodes and Wellman (2017) argued that

children's moral thinking is guided by intuitive theories on mental states as well as on group membership and relations, indicating that children can acquire rule-like structure in the moral domain from rational inference with little explicit instruction.

There has also been a surge of evidence leading some to suggest that basic moral systems and representations are innate or very early emerging, indicating that core aspects of morality are independent of social learning. Developmental psychologists supporting the nativist perspective argue that infants demonstrate harm-based and fairness-based moral beliefs during the first year of life (Hamlin, Wynn, & Bloom, 2007; Kuhlmeier, Wynn, & Bloom, 2003; Geraci & Surian, 2011; Sloane, Baillargeon, & Premack, 2012, for a review, see, Woo, Tan, & Hamlin, 2022). For example, 6-10-month-olds were more likely to reach for an agent that had helped another agent compared to one that had hindered an agent (Hamlin et al., 2007). Even 3-months-olds seem to show a preference for helpers compared to hinderers (Hamlin, Wynn, & Bloom, 2010), raising the possibility that young infants' own preferences are guided by the moral characteristics of agents. There is also evidence suggesting that 18-month-olds spontaneously help others in the absence of social praise or encouragement, leading some to argue that altruistic helping is the result of a natural or evolutionary disposition (Warneken, & Tomasello, 2009). In response to these studies, nativist accounts which are advanced to explain the origins of morality have become increasingly mainstream (Bloom, 2010, but see Tasimi, 2020, who argued that infants construct moral knowledge from interacting with their caregivers).

Since these developmental accounts prioritize the innate origins and authority-independent nature of moral development, the role of direct testimonial influence is often

left unaddressed (Harris, 2012). Specifically, these dominant developmental theories are largely guided by the assumption that moral development is motivated by the hard labor of the child, either because of children's innate moral capacities (Bloom, 2010), or the active role children play in independently constructing moral systems (Smetana, 2006; Rhodes & Wellman, 2017). Thus, these approaches leave open questions about the amount of weight that children put on *testimony* supplied by adult and peer informants. Here, I argue that moral development is actually an *interdependent* endeavor. Below, I first argue that some of the arguments for the suspicious nature of moral testimony do not actually prevent children from gaining knowledge from testimony. I then discuss how moral testimony can afford children opportunities to acquire moral understanding in ways that independent deliberation cannot achieve.

### *Is Moral Testimony Suspicious to Child Learners?*

**Strong Priors.** One reason I might reasonably suspect that children find moral testimony to be less legitimate than testimony in factual domains is that children may hold innate or early emerging beliefs representing basic moral principles. As reviewed, infants are said to hold innate intuitions about harm and fairness related principles (Hamlin et al., 2007; Sloane, et al., 2012; Bloom, 2010), and they also have opportunities to experiment with their social environments to acquire moral knowledge from observation and trial and error (e.g., that hitting or pushing others causes distress, Heyes, 2019; Sterelny, 2010). In fact, the limited research on moral judgments has shown that children are sometimes resistant to social and testimonial influences when it comes to harm-related principles. For example, using non-verbal testimony, Kim, Chen, Smetana and Greenberger (2016) found that only 20% of the U.S. preschoolers conformed to their

peers' moral judgments (signaled by hand-raising) that hitting another child was good on at least one of four trials.

However, when the situation is made more ambiguous and described using novel language, adult testimony can lead children to judge actions that cause distress as permissible. Li, Harris and Koenig (2019) found that 3- to 5-year-old Chinese and U.S. American children independently judged novel, distress-inducing actions (e.g., described as "mibbing") to be morally wrong, but they often revised their original judgments after hearing counter-intuitive claims made by adults (e.g., "It is good to mib."). Moreover, Chinese children were more likely to defer to adult testimony than U.S. children, which might be related to cultural variability in socialization practices that prioritize the relative weighting of social learning versus independent judgments.

Beyond harm-related moral principles, children are receptive to testimony about other moral norms as well. A small but growing body of developmental research has shown that testimony can significantly affect moral beliefs about group norms (Chalik, & Rhodes, 2015; Lane, Conder, & Rottman, 2019), distributive strategies (Benozio, & Diesendruck, 2016; Heck, Chernyak, & Sobel, 2018; Rottman, Zizik, Minard, Young, Blake, & Kelemen, 2020), as well as influence moral behavior such as helping (Dahl, 2015; 2018; Hammond, & Carpendale, 2015; Dahl, & Brownell, 2019) and lying (e.g., Fu, Heyman, Qian, Guo, & Lee, 2016; Zhao, Heyman, Chen, & Lee, 2017 ; Zhao, Chen, Sun, Compton, Lee, & Heyman, 2018). Thus, this broad range of findings might be explained by the variable strength of children's prior beliefs: children likely have weaker priors about some moral behaviors relative to others. This is consistent with Rottman and Young's (2015) developmental model that some moral norms (e.g., norms involving

impurity) might be more heavily shaped by social communication than others (e.g., norms involving harm).

**Moral Expertise.** Another reason that makes it difficult to defer to moral testimony , according to the pessimists, is the difficulty of identifying moral experts. As reviewed, it is difficult for adults to assess the reliability of moral testimony and the track record of an informant especially in complex situations (Hills, 2009; Goldman,2001; Driver, 2006; Cholbi, 2007; McGrath, 2011). Do children have the same issues in identifying moral expertise? Previous work has shown that preschool-aged children can identify expertise in different factual and scientific domains (e,g, Aguiar, Stoess & Taylor, 2012; Boseovski & Thurman, 2014; Danovitch & Keil, 2004; Koenig, & Jaswal, 2011; Lane & Harris, 2015; Lutz & Keil, 2002), and keep track of an informant's record of competence (Corriveau, Meints & Harris, 2009; Koenig, Clément & Harris, 2004; Pasquini, Corriveau, Koenig & Harris, 2007; for a review, see Stephens, Suárez & Koenig, 2014). However, previous research has not explicitly assessed how or why children attribute moral expertise to speakers who make moral claims. One possibility is that similar to adults, children may find reliable informants about moral issues more difficult to identify, resulting in a higher tendency of self-reliance.

This being said, it is also plausible that children's trust in moral testimony is not deterred by the difficulty in identifying moral experts. After all, expertise is often understood in relative terms (Driver, 2006; Hopkins, 2007). Since moral knowledge often comes with prolonged experience and training in navigating social situations, it is not uncommon for young children to come across people who are more skilled at making moral judgments than themselves. Indeed, preschool-aged children are perfectly capable

of assessing the reliability of an informant and her empirical claims by relying on factual evidence (e.g., acquiring first-hand evidence about what activates a music box, or the location of a sticker, Hermansen et al., 2021; Ronfard, & Lane, 2019) or speaker characteristics (e.g., sincerity, familiarity, group membership, age, for a review, see Harris at al., 2018). Furthermore, moral expertise can come in many forms. For example, moral experts can be moral practitioners who act morally well. In fact, preschool-age children have been found to use people's morally good or bad actions as a basis for distinguishing between reliable and unreliable informants (e.g., Johnston, Mills, & Landrum, 2015; Lane, Wellman, & Gelman, 2013; Vanderbilt, Liu, & Heyman, 2011). There is also evidence suggesting that children selectively seek moral advice from more helpful or prosocial agents (Danovitch, & Keil, 2007; Doebel & Koenig, 2013). Moral experts can also be those who are better able to apply moral principles in given situations and make less biased decisions (Driver, 2006; Jones, & Schoroeter, 2012), or those who have more relevant moral experience than others in virtue of their social location (e.g., being the person who experienced the transgression, Jones, 1999). Indeed, young children tend to defer to adult judgments when applying harm and fairness-related norms to novel situations, even when no competence-related information about the adult informants are given (Rottman et al., 2017; Li et al., 2019). Moreover, moral experts can be those who are putatively granted a superior status (e.g., parents, teachers as in Mammen et al., 2019; Kruger & Tomasello, 1986). Thus, while I appreciate the pessimists' concerns about identifying moral experts, I also think there are reasons to expect that children selectively use and trust testimony in the moral domain.

**Moral Disagreement and Consensus.** Another pessimistic argument against the value of moral testimony concerns the considerable amount of disagreement in the moral domain (Hills, 2013; Hopkins, 2007), with the rationale that the lack of consensus would undermine our confidence in potential informants and their testimony, pushing for first-personal authority in making moral judgments. Does observed disagreement in the moral domain hinder children's own trust in testimony? For children, many moral situations they encounter are not actually subject to sustained moral disagreement. After all, it seems implausible to think that children would be exposed to disagreement about basic moral principles such as the wrongness of hitting or hurting others, or the niceness of sharing and helping. Nevertheless, children may observe a range of opinions regarding more subtle, or complex situations when adults and peers make judgments about what counts as "harm", whether there are valid justifications for inflicting harm, or whether a victim is without fault. In fact, children are sensitive that some moral issues are more controversial than others. For example, Heiphetz and colleagues (2018) found that 5- to 8-year-old children were less sure that other agents (another person, god) would judge controversial beliefs such as telling prosocial lies to be okay than uncontroversial good behaviors such as helping others, indicating that children understand that some moral issues may elicit more disagreement between individuals than others (see also Hussar & Harris, 2010).

Will the prevalence of moral disagreements inhibit children's ability to identify reliable informants and learn from them? Here, I argue that children's reliance on testimony is not deterred by moral disagreements. On the contrary, they might be even more likely to defer to testimony from authorities when navigating controversial moral

issues, possibly because they have less certainty about the given situation. Although children's judgments can be authority-independent when it comes to judging prototypical harm-related moral transgressions such as hitting and shoving (e.g., Smetana & Ball, 2017), there is evidence showing that children are especially likely to rely on adult authorities when resolving moral dilemmas. For example, even though children's acceptance of others' testimony has not been directly investigated, Noh and colleagues (2020) found that 6- to 12-year-old Korean children's judgment for the permissibility of necessary psychological harm (e.g., calling a friend a bad name to prevent him from climbing on the roof and hurting himself) is significantly affected by maternal disapproval. Mammen and colleagues (2019) also found that when faced with moral dilemmas (i.e., conflicts between keeping a promise and helping others), 4- and 6-year-old children engaged in collaborative reasoning with others, and produced fewer challenges when having discussions with their mothers compared to with peers, raising the possibility that children may readily defer to an adult authority's judgment when forming beliefs about complex moral situations.

Aside from possibly relying on sources of authority, another interesting issue concerns whether children would side with the majority in cases of moral disagreement. In the empirical domain, when there are disagreements between informants, children are sensitive to consensus when deciding whose claims to trust when making conventional judgments (e.g., Corriveau, Fusaro, & Harris, 2009). For example, when learning the name of a novel object, preschoolers were more likely to endorse information that had been approved by three consensual informants over a lone dissenter (Corriveau et al., 2009; Corriveau & Harris, 2010). Similarly, children were also more likely to imitate the

use of a novel tool demonstrated by three agents compared to that displayed by a single individual (DiYanni, Corriveau, Kurkul, Nasrini, & Nini, 2015; Haun, Rekers, & Tomasello, 2012). Thus, perhaps when a consensus or near consensus exists for a moral issue, children are similarly sensitive to the majority view of their cultural community. However, there is some evidence showing that children sometimes do not distinguish between the reliability of a single informant and a consensus when it comes to accepting a moral claim (Li, Harris, Koenig, 2019; Noh, Elenbaas, Park, Chung, & Killen, 2017), and even weigh an outside source of information more heavily than a consensus (e.g., deciding to be a vegetarian based on a source outside the family about animal welfare, despite being raised in families and communities that typically supported eating meat, Hussar & Harris, 2010). One possibility is that children have an emerging understanding that the moral realm is one in which consensus is less common, and a less valid cue to reliability. Moreover, it is possible that children do not prioritize consensus in the moral domain because the epistemic value of consensus is unclear. In the empirical domain, children often discern the reasoning or epistemic support behind an informant's claims (e.g., speaker has visual access to something, Corriveau & Harris, 2010; Koenig, Tiberius & Hamlin, 2019), so support from a consensus implies greater support for an empirical claim. By contrast, the reasoning behind a moral claim is typically more opaque and less connected to informational access, making it harder for children to determine the connection between consensus and accuracy.

In conclusion, despite the pessimistic worries related to moral testimony, I think that testimony can influence children's moral thinking considerably more than mainstream developmental theories acknowledge. Despite children's basic understanding

of certain moral principles, and even in cases of pervasive moral disagreement, children have strategies to use moral testimony, and acquire much of their knowledge from socially transmitted information. In the next section, I discuss ways in which moral testimony has the unique capacity to foster moral understanding.

### *Reasons to Think that Moral Testimony is Legitimate to Child Learners*

Despite reasons to think that moral testimony may be less legitimate than testimony in other domains, we have reasons to believe that children cannot afford to treat testimony this way — because testimonial input is essential for them to learn the moral beliefs of their communities and culture, and to guide and correct their moral decisions going forward. After discussing that concerns related to moral testimony may not hold for children, in this section, I draw from philosophical and psychological work which allows for the possibility of morality being an interdependent and collective endeavor, and how testimony can in fact be a unique and indispensable source of moral knowledge.

**Value of Testimony for Moral Reasoning.** Going beyond the dual-process model, Mercier (2011) suggests that moral reasoning can play an indispensable role in correcting one's moral intuitions (see also Mercier, 2016; Mercier, Castelain, Hamid, & Marín-Picado, 2017). In Mercier's view, reasoning, including moral reasoning, is fundamentally social, and mainly serves argumentative purposes when it comes to convincing others. While solitary reasoning often leads to confirmation biases, collaborative reasoning with others can often lead to improvements in moral thinking (Mercier et al., 2017; Köymen, & Tomasello, 2020). Specifically, being attentive to others' claims is beneficial because it can make one recognize the flaws in one's own

arguments, and help generate new arguments amongst conversational partners that would perhaps never occur if each individual were merely reasoning independently (Mercier, 2016). These arguments have received support from studies of children's discussions about moral issues with peers. For example, when 4th and 5th grade children were presented with Piaget-style problems in pairs (e.g., judging which transgression is more serious: breaking six cups accidentally or breaking one cup when trying to get candy), children were more likely to focus on intentions and moral values instead of external consequences (Leman & Duveen, 1999; see also, Blatt & Kohlberg, 1975; Zhang et al., 2013). There is also evidence suggesting that having conversations with peers can improve children's moral reasoning skills. For example, Damon and Killen (1982) found that having a debate with peers about how much to share in a Dictator Game led to a modest increase in 5- to 9-year-old children's level of moral reasoning after 2 months (see also Walker, Hennig, & Krettenauer, 2000). Thus, these studies have shown that children's moral reasoning can be improved under testimonial influence.

**Transcending Limits of Experience.** On the view that dependence on moral testimony can be valuable as a way to overcome our own flaws in thinking, it may also be reasonable to speculate that children may have an emerging understanding that their own capabilities and experience in moral thinking are limited. After all, it becomes worthwhile to defer to another person who has reflected more carefully on a moral issue, or who has more relevant experience than we have (Sliwa, 2012; Wiland, 2017). For example, grasping exactly when a basic moral principle applies can be a matter of experience or ability — a man might need to defer to his female roommate to identify sexist microaggressions since she has more relevant experience (Jones, 1999). Indeed,

learning when to apply a moral principle, and identifying the subtleties of specific cases takes experience and practice. Although there hasn't been any work on children's metacognitive evaluations of their own moral thinking, a couple of empirical studies have shown that children are receptive to adult testimony when applying harm-based norms and fairness norms to novel situations. For example, Rottman and colleagues (2017) found that neutral behavior (e.g., aliens covering their heads with sticks, or painting their faces white) could be perceived as morally impermissible by 7-year-olds after they heard counter-intuitive testimony invoking moral principles such as harmfulness and fairness (e.g., "It really hurts others when Kulvaws paint their faces white."). This finding also implies that children likely rely on testimony to appreciate the wrongness of victimless moral violations (e.g., Rozin, 1999). When new moral situations cannot be directly accessed through first-hand experience or socialization practices from children's own family, children can also acquire information from other people who are outside their direct social circle, independently learning that, for example, meat eating is wrong because of animal welfare-related concerns (Hussar & Harris, 2010).

Because of testimony's ability to transcend the limits of individual experience, it can also play an essential role in the intergenerational transmission of culturally specific moral knowledge. Although psychological and anthropological research have documented some universal moral beliefs across cultures (e.g., Curry et al., 2019), much empirical evidence has highlighted the important cultural and historical differences in moral norms, beliefs and practices, as well as the different patterns of emphasis on some moral principles over others. For example, robust cultural variations have been found in the three aspects of ethics (autonomy, community, and divinity) across populations with

diverse social statuses and cultures (e.g., Haidt, Koller, & Dias, 1993; Jenson, 2004). More direct developmental evidence has shown cross-cultural differences in children's moral development such as prosocial sharing (Cowell et al., 2017; Schaefer et al., 2015) and helping (Callaghan et al., 2011). Therefore, it is plausible to assume that children acquire moral beliefs and practices through cultural learning from their own communities. Although observational learning from the environment can play an important role in moral socialization (Bandura, 1986), cultural testimony would be essential and irreplaceable in transmitting virtues such as what it means to be kind, or moral beliefs such as group-related obligations or concerns about purity and sanctity (Rottman, & Young, 2015; Haidt, & Bjorklund, 2008).

**Correction of Biases.** Another reason for children to trust testimony is that children may have the emerging understanding that our judgments may sometimes be compromised or biased by our own self-interest (Mills & Keil, 2005; Sliwa, 2012). Indeed, it can be difficult to have an impartial perspective in our moral thinking, since even children tend to reify norms as marking the way things ought to be rather than as variable, contingent and due to external forces (Peretz-Lange, Perry & Meuntener, 2021; Roberts, Ho & Gelman, 2020; Vasilyeva, Gopnik & Lombrozo, 2018). This argument is consistent with Mercier's (2016) point that independent reasoning often results in confirmation biases, leading young moral agents to favor arguments that already support their thinking. Although this particular evaluation of testimony has not been assessed, there is some evidence showing that from a young age, children realize that people's moral judgments can be biased towards themselves, towards friends or family members and toward in-group members. For example, preschool children treated moral

transgressions, such as taking a toy, as more permissible when the transgressor was a friend of theirs (Slomkowski & Killen, 1992). Children were also more sympathetic toward a friend in need compared to an acquaintance (Costin, & Jones, 1992). In addition, with age, children increasingly prioritize social relationships when allocating resources, sharing more with friends who already have resources compared to nonfriends or strangers who are needy (Paulus, & Moore, 2014). Children's judgments are also biased towards in-group members. They judge between-group harm to be more permissible (Rhodes, & Chalik, 2013), and are less likely to tattle on an in-group member's moral transgression compared to an out-group member (Misch, Over, & Carpenter, 2018). Although endorsing biases is not equivalent to being aware of one's own biases, these results raise the possibility that children may rely on testimony in acquiring and changing moral beliefs when others serve as better judges of a given situation.

**Deference as the first step to achieve moral understanding.** Lastly, children cannot afford to treat testimonial input as secondary because dependence on moral testimony can give rise to moral understanding (Gelfert, 2014; Jones, 1999). Indeed, deferring to moral testimony may be the first step on the way to acquiring moral understanding. Even if a child decides to make a moral judgment just based on someone's say-so without fully understanding the nuances (e.g., judging that an action is harmful based on someone's claim), he may still be able to gradually improve his moral sensibility and hone his own capacities of moral perception over time, and eventually become better at making relevant moral judgments and correct distinctions between cases himself (Jones, 1999; Sliwa, 2012; Matheson et al., 2018; Paddy Jane, 2018).

Specifically, since our moral commitments often shape our moral perception, it is plausible that the bulk of learning cannot happen without having accepted certain beliefs as truths by deferring to moral testimony. Imagine a child who understands the principle of fairness but cannot always tell what is considered fair or unfair. When she was unsure about a situation, her parents would tell her that she was being unfair (e.g., to choose a fun toy for herself and a boring toy for her sister), and, without fully understanding the reasoning behind her parents' judgments, the child defers. But over time, with more and more practice, the child develops more sensitivity to how the fairness norm applies to different contexts, and gradually develops the capacities for moral understanding. Thus, dependence on testimony could in fact promote understanding.

In summary, while testimony is an indisputable source of empirical knowledge, the role of testimony in the moral domain remains an open question. Here, I have reviewed the controversy in the moral philosophy literature on whether testimony can effectively and legitimately transmit moral understanding. I have also reviewed psychological work examining the long-standing issue of how moral competence develops, and argued that moral development is a socially attuned, interdependent endeavor, whereby testimony from others plays an essential role. By reviewing evidence that moral development can be profoundly influenced by testimony, I argue that efforts to clarify the nature of children's reliance on moral testimony in a more systematic and careful manner can be extremely informative for understanding the learning mechanisms and developmental processes that underlie the acquisition of moral beliefs.

**Aims of the Current Studies**

To better document and understand how moral decisions are supported by testimony, my dissertation work was guided by the following main questions: First, can

testimony change moral judgments, and if so, what types of testimony are most powerful in moving children's moral judgments? Second, compared to other domains of empirical knowledge, do children find it less legitimate to justify moral beliefs on the basis of testimony? Lastly, are there cross-cultural differences in children's use and evaluations of testimony in the moral domain?

**Aim 1: To investigate whether testimony changes children's moral judgments, and if so, identify the types of testimony that are most powerful in moving moral judgments.** My own research has found that 3- to 5-year-old Chinese and U.S. American children independently judged novel, distress-inducing actions to be morally wrong, but revised their original judgments after hearing contrary claims made by adult experimenters (Li, & Harris, Koenig, 2019). However, there is still a concerning lack of research on how testimony from other people affects children's moral judgments, and no experimental work has tested whether some types of testimony exert a more powerful influence on children's moral judgments than others. Study 1 explored children's moral judgments in the context of countervailing claims from an adult, which were either (1) supported by reasoning about the consequences, or which (2) exerted power and authority over children.

**Aim 2: To examine children's own evaluations of different sources of moral and empirical knowledge.** Although a major line of research has explored children's ability to identify and to selectively learn from competent or knowledgeable sources over less knowledgeable ones (Koenig, Clément, & Harris, 2004), no studies to date have explored *children's own perspectives* on how best to acquire knowledge. But children's own views on different sources of knowledge can carry important implications for their decisions to learn from and about other people, and it supports the exercise of moral agency.

To fill the gap in our understanding of children's own view on reliance on testimony in the moral realm, Study 2 explored whether children judge reliance on testimony as less legitimate than coming to one's own conclusions, especially in the moral domain. I also assessed whether children prefer to learn from a speaker who has independent knowledge compared to one who acquires second-hand knowledge from others.

**Aim 3: To study cross-cultural differences in children's acquisition of moral knowledge.** Children's learning is situated in sociocultural contexts, making cross-cultural comparisons particularly relevant in exploring the universality and cultural specificity of the transmission of moral beliefs and norms. Previous research shows that children's reliance on others appears to be influenced by *culturally-shaped experiences and beliefs*. Specifically, cultural expectations for conformity and intellectual modesty have been found to be weaker in Western individualistic societies than in non-Western communities that are more collectivistic (Bond & Smith, 1996; Clegg, Wen & Legare, 2017; Wen, Clegg & Legare, 2019). Across Studies 1 and 2, I addressed the pressing question of cultural mechanisms underlying children's reliance on testimony for moral knowledge by including samples of Chinese children and U.S. American children.

## CHAPTER 2

### Study 1: The Influence of Explanations on Children's Moral Judgments

#### Introduction

Children acquire much of their knowledge from the testimony of other people. As reviewed, much of the work on testimonial learning has focused on the transmission and acceptance of non-social knowledge such as the name of an object, geography of the world, historical events, or the existence of germs (Harris & Koenig, 2006; Lane, 2018,

for reviews, see Harris et al., 2018; Mills, 2013). We know much less about the role played by testimony in children's moral judgments. Do children rely as heavily on others for their moral beliefs? If so, are children more influenced by some types of testimony than others?

**Testimony and Moral Development.** Testimony has long been associated with the transmission of factual knowledge (Gelfert, 2006). As reviewed in Chapter 1, much developmental research has suggested that young children's moral judgments might be relatively impervious to social influences. For example, the Social Domain Theory suggests that children's understanding of moral norms is linked to the nature of an act and its consequences for a victim, independent of any prescription by local authorities (Smetana, 1981; Smetana et al., 2012).

Despite long-standing views that take moral thinking to be "self-given" (Kant, 1785/1959) and authority independent (Smetana, 1981), children's moral learning is still deeply embedded in social, conversational and cultural contexts. Recent developmental evidence suggests that testimony can influence children's moral judgments, even when it comes to more fundamental harm-based norms and fairness norms. For example, Rottman and colleagues (2017) found that neutral behavior (e.g., aliens covering their heads with sticks, or painting their face white) could be perceived as morally impermissible by 7-year-olds after they heard counter-intuitive testimony invoking moral principles such as harmfulness and fairness (e.g., "It really hurts others when Kulvaws paint their faces white."). Conversely, adult testimony can also lead children to judge harmful actions as permissible. Li, Harris and Koenig (2019) found that 3- to 5-year-old Chinese and U.S. American children judged novel, distress-inducing actions (e.g.,

"mibbing") to be morally wrong independently, but revised their original judgments after hearing counter-intuitive claims made by adults (e.g., "It is good to mib."). However, to our knowledge, no experimental work to date have examined whether some types of testimony exert a more powerful influence on children's moral beliefs than others. Specifically, it remains unknown which types of explanations would be most effective in changing children's moral judgments.

**Children's appreciation of reason-based and authority-based arguments.** Previous experimental studies have revealed that children start to show a preference for higher-quality arguments concerning descriptive knowledge from 3- to 5-years of age (e.g., Corriveau & Kurkul, 2014; Koenig, 2012; Mercier, Bernard, & Clément, 2014). For example, 4- to 5-year-old children judged explanations involving authority endorsement, inference, and perception as better reasons for belief about the content of a container than reasons such as pretense, guessing, and desiring (Koenig, 2012). There is also a growing literature on children's appreciation of noncircular explanations over circular explanations. For example, when the task involved finding a lost animal, children were more likely to accept a perceptual argument (e.g., "because I've seen him go in this direction") compared to a circular argument (e.g., "because he went in this direction", Mercier, Bernard, & Clément, 2014). Children also preferred to learn from a speaker who had previously used noncircular explanations (e.g., "It rains because the clouds fill with water and get too heavy.") compared to one who had used circular explanations (e.g., "It rains because water falls from the sky and gets us wet.", Corriveau, & Kurkul, 2014).

Aside from trusting informants who engaged in higher quality argumentation and reasoning, children's endorsement of testimony can also be moved by a speaker's power

and authority (Bernard, Castelain, Mercier, Kaufmann, Van der Henst et al., 2016; Fusaro, Corriveau & Harris, 2011). As third-party observers, children can mentally represent and infer social dominance from infancy (Thomsen et al., 2011) to early childhood (Charafeddine et al., 2015; Kajanus, Afshordi, & Warneken, 2019), and prefer individuals with higher ranking and status (Thomas et al., 2018). High authority or dominance can be a salient cue for children to trust a speaker in testimonial learning contexts. For example, 3- and 5-year-old children preferred to trust the testimony of a dominant speaker who had more physical or decisional power (e.g., physically contesting a toy and winning) over a subordinate speaker when learning about the localization of an animal and the names of objects (Bernard et al., 2016).

To my knowledge, only a few studies have directly pitted authority against strong arguments in the empirical domain, but results on the relative influences of the two types of considerations are mixed. In one study, 4- to 6-year-old children were more likely to endorse the testimony of a subordinate speaker who provided a strong argument over that of a dominant speaker who provided a circular argument or no argument (Castelain, Bernard, Van der Henst, & Mercier, 2016). However, in another study, children were more likely to trust an expert (presented as "a scientist who knows all about living things") who provided lower-quality explanations about scientific facts compared to a non-expert who provided higher-quality explanations (Clegg, Kurkul, & Corriveau, 2019).

Moreover, none of the studies to date have directly compared the effects of reason-based arguments and authority-based arguments in the moral domain. For social psychologists, independent moral reasoning, operated through rational and controlled

cognitive processes, is an important component of moral thinking (Paxton, & Greene, 2010; Pizarro, Uhlmann, & Bloom, 2003). Although it is also deemed possible for an adult's moral judgments to be moved by social persuasion for affiliation purposes (Haidt, & Bjorklund, 2008), the relative effectiveness of reasoned persuasion and social persuasion have not been empirically explored. Observational studies with parental discipline techniques have shown that when it comes to moral socialization, it is common for parents to utilize both reasoning (e.g., discussion of the consequences of the transgression for the self) and power assertion (e.g., threaten of punishment, Grusec & Goodnow, 1994; Leman, 2005; Zahn-Waxler, Radke-Yarrow, & King, 1979). However, evidence on whether one parenting strategy is superior to the other remains inconclusive (Grusec, 2006). Here, I assessed these two types of explanations that may play an indispensable role in changing children's moral understanding. Specifically, in Study 1, I investigated children's moral judgments in the context of countervailing claims from an adult, which are either focused on reasoning about consequences (Utilitarian explanations) or exerting authority over children (Authoritarian explanations).

**Cross-cultural Differences in Testimonial and Moral Learning.** Children's social learning is situated in sociocultural contexts, making cross-cultural comparisons particularly relevant in exploring children's acceptance of moral testimony. Specifically, there might be cultural variability in socialization practices that prioritize the relative weighting of adult testimony versus the child's own independent judgment. Several studies have found that individuals from non-Western communities that are more collectivistic are more likely to emphasize conformity, interdependence, and social learning than Western individualistic societies (Bond, & Smith, 1996; Chang et al., 2011;

Clegg, Wen & Legare, 2017; Toelch, Bruce, Newson, Richerson, & Reader, 2014). For example, Clegg and colleagues (2017) found that Ni-Vanuatu adults were more likely to judge a high-conformity child who imitated at higher fidelity as intelligent, whereas U.S. adults were more likely to credit a low-conformity child for being innovative. Relatedly, research has also revealed that early socialization practices in East Asian countries like China are more likely to emphasize the importance of learning from adult authorities, conformity and intellectual humility (Li, 2005; Suizzo & Cheng, 2007; Wang, Elicker, McMullen, & Mao, 2008; Wu et al., 2002). For example, when Chinese and U.S. mothers of preschoolers were asked to describe the most important parenting practices, Chinese mothers were more likely to emphasize the encouragement of modesty and cooperation compared to U.S. mothers (Wu et al., 2002). This cross-cultural variation is manifested in children's susceptibility to counter-intuitive claims, with Asian-American and Chinese children being more likely to defer to adult testimony than U.S. children when making perceptual judgments (Corriveau & Harris, 2010; Corriveau, Kim, Song, & Harris, 2013) and moral judgments (Li et al., 2019). Asian-American children were also more likely to conform to an inefficient behavior demonstrated by a consensus of adults compared to European American children (Corriveau et al., 2017, DiYanni et al., 2015). However, little is known about how variability in cultural expectations might result in differences in children's rates of acceptance of moral explanations, nor have any studies explored how children weigh counterintuitive reasoning-based and authority-based explanations. The current study explored the role of cultural mechanisms underlying children's reliance on testimony by including samples of Chinese children and U.S. American children. Based on previous research suggesting that children in China may place more emphasis on

authority and intellectual modesty (e.g., Li et al., 2019), I predicted that Chinese children would be more deferential to adult explanations than their U.S. American counterparts, and that they might be especially sensitive to Authoritarian explanations in which the adult speaker exerted power over children.

In summary, despite the limited attention on the role of testimony on moral development, recent findings suggest that testimony plays an important role in stimulating and changing children's moral thinking. Moreover, although children prefer to learn from informants who could either provide good reasoning or have higher authority in the empirical domain, it remains unknown whether reason-based and authority-based testimony can exert a powerful influence on children's moral judgments. The primary aim of Study 1 was to assess the influences of these two types of explanations on participants' moral judgments. In three between-subjects conditions, 3- to 5-year-old Chinese and U.S. children and adults listened to stories with illustrations in which the protagonists engaged in actions that made others cry, referred to only by means of unfamiliar words such as "mibbing". Then, depending on the condition, children were exposed to an adult speaker who provided *Utilitarian* explanations (e.g., "Mibbing is good because Devon can play with the new toy after mibbing."), *Authoritarian* explanations (e.g.,"Mibbing is good because I get to decide whether mibbing is good."), or *Control* statements (repeating basic information in the story). Afterwards, children were asked a series of questions to evaluate the novel action as well as the transgressor.

## Method

**Ethics Statement.** This study was approved by the Institutional Review Board at

the University of Minnesota ("Moral Explanations", IRB ID: STUDY00007238). A parent or guardian provided written consent. The data in the U.S. was collected between September 2019 and August 2020; the data in China was collected between June 2021 to August 2021.

**Participants.** I recruited a total of 379 typically developing children (198 U.S. children: $M_{age}$ = 54.7 months; range = 40.2 – 71.3 months; 181 Chinese children: $M_{age}$ = 55.7 months; range = 36.1 – 72 months), including 95 3-year-olds (40 female, 55 male, $M_{age}$ = 44.3 months; range = 36.1 – 47.9 months), 163 4-year-olds (91 female, 72 male, $M_{age}$ = 53.9 months; range = 48.2 – 59.3 months), and 121 5-year-olds (61 female, 60 male, $M_{age}$ = 65.2 months; range = 60.1 – 72 months). An additional 7 children were excluded due to failure to complete the study ($n$ = 5) or experimenter error ($n$ = 2). U.S. participants were recruited through a university-managed database of families from the greater Twin Cities areas, who have indicated interest in participating in research. Chinese participants were recruited from a private preschool in Changping district, Beijing.

All parents were invited to complete a voluntary questionnaire with demographic questions. In the U.S. sample, 89.9% identified their child as White, 1.5% Black, 1% Asian, 1% Native Hawaiian or other Pacific Islander, 0.5% American Indian or Alaska Native, and 5.1% identified their child with multiple races or ethnicities; 4% were Hispanic (of any race). Self-reported education was fairly high: 94.95% of the mothers and 83.33% of the fathers had a Bachelor's degrees or above. Parent reported income was primarily middle class, with 63.64% reporting an annual income greater than $10,000 per year. In the Chinese sample, all children were identified as ethically Chinese, with

90.06% identified as the Han ethnicity. Bachelor's degrees or above were held by 77.9% of mothers (169 out of 181, or 93.37% of parents answered this question) and 83.53% of fathers (170 out of 181, or 93.92% of parents answered this question), which was higher than 41.98% of Beijing's population in 2021 (Beijing Statistics Bureau, 2021). The sample in this study also had a higher average annual income than that of the Beijing population. Out of the 163 families that answered this question, approximately 79.75% of the families earned an annual income of 25,2000 RMB or more, which is higher than the average for Beijing urban families in 2021 (248,286 RMB based on a family of three; Beijing Statistics Bureau, 2021).

All Chinese children and 57.07% of the U.S. children were tested in in-person sessions. After in-person data collection was cut short due to COVID-19, I tested the rest of the U.S. participants via video-conferencing. The stimuli that children saw (i.e., a series of PowerPoint slides with pictures and videos) were identical across in-person and online sessions, and there were no significant differences between the settings in children's responses across all the tasks (all *p*s > .38*)*.

In order to gauge the degree of change from childhood to adulthood, I also recruited and tested adults in the same procedure online using Qualtrics. 90 U.S. participants ($M_{age}$ = 29.13; range = 18-56; 69 female, 19 male and 2 non-binary; 78.9% White, 12.2% Asian, 2.2% Black, 1.1% American Indian or Alaska Native, 1.1% more than one race, 4.4% another race, 12.2% Hispanic of any race) were recruited online through the online participant recruitment platform Prolific. A comparable sample of 90 Chinese participants ($M_{age}$ = 30.64; range = 18-59; 65 female and 25 male; all ethnically Chinese, 92.2% of the Han ethnicity) were recruited online through social media.

**Procedure.** Children were randomly assigned to one of three between-subject conditions: the *Utilitarian* condition, the *Authoritarian* condition, or the *Control* condition. During the study, children went through three trials in which they listened to stories about children who engaged in novel actions that made others cry. Then, depending on the condition, children were exposed to an adult speaker who either provided explanations on why the novel actions were permissible (in the *Authoritarian* and *Utilitarian* conditions), or provided control statements (in the *Control* condition). Afterwards, children answered a series of questions about the novel action and the transgressor (i.e., *moral judgment questions, transgressor evaluation task)*. In order to assess whether children were merely complying with the informant or had genuinely revised their beliefs, I also included a *true belief* task, in which children judged the permissibility of the actions in a new context (*context transfer question*s), and decided whether to transmit their judgment to a puppet who wanted to learn about the novel actions (*Information transmissions)*. Each task is described below in the order that it was presented to children (please see Figure 1 for a visual depiction of the procedure).

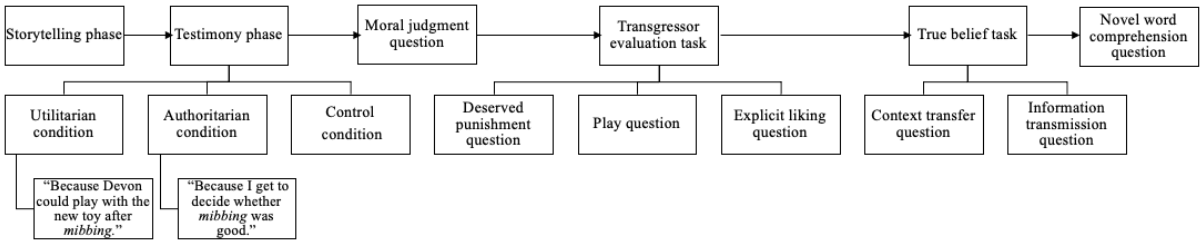


*Figure 1.* Flow Chart of the Procedure in Study 1.

***Storytelling phase.*** The experimenter always began the study session by saying: "I am going to tell you some stories today. In each of the stories, I might hear a word that we have never heard before, and I might need your help figuring out what it means.

Okay, ready?" Then, in each trial, the experimenter told children a story with the aid of three drawings. In each story, the protagonist engaged in an action that were referred to only by means of unfamiliar words such as "mibbing", and it was made clear to children that the novel actions caused harm and distress to another child (e.g., "Devon *mibbed* Casey and made Casey very upset and he started crying.", see story scripts in Table 1 in Appendix). The reason for using novel actions was to avoid any uncontrolled bias from adult testimony associated with familiar actions. Previous research has shown that children could make a moral evaluation of these unfamiliar actions because each story recounted their distress-inducing consequences (Li et al., 2019; Smetana, 1985).

***Testimony phase.*** After narrating the first story, the experimenter told the child: "Alright! We have heard the first story. Now, I'm going to show you a video of a lady who heard the same story about *mibbing* yesterday. Let's pay close attention to what she says, okay, ready?" Each video featured a confederate sitting at the table with an adult female informant. In the two testimony conditions (*Utilitarian* condition *and Authoritarian* condition), the confederate in the video asked the informant, for example: "What do you think? Was it good or bad for [Devon to *mib*]?" The informant in both conditions replied by stating that it was good to "mib". Then, the experimenter in the video continued to ask: "Why was it good to *mib*?" and the informant either provided a *Utilitarian* explanation that reasoned about positive consequences for the transgressor (e.g., "Because Devon could play with the new toy after *mibbing.*") or an *Authoritarian* explanation that exerted power and authority over children (e.g., "Because I get to decide whether *mibbing* was good."). By contrast, in the *Control* condition, the same confederate also asked the informant two questions, but the questions were about basic information in

the story (e.g., “Who did the *mibbing*, Devon or Casey?”), and the informant did not give any judgments of the novel actions.

***Moral judgment questions.*** After hearing each story and receiving adult testimony, the experimenter first asked comprehension check questions to make sure the child could remember that the novel actions caused distress to the victim, and what the informant had said in the video. If the child could not remember the story outcome, the experimenter showed them the last drawing with the victim in tears to remind them that the victim was upset after the protagonist engaged in the novel action. If the child could not remember what the informant said, the video was shown again to the child.

After the comprehension checks, the experimenter asked for children’s own *moral judgment,* for example: “What do you think? Was it good or bad for [Devon to *mib* Casey] ?” If the child indicated that it was bad, the experimenter said: “Do you think it’s a little bad or very bad?” If the child indicated that it was good, the experimenter followed up by asking if it was “a little good” or “very good”. Children’s responses were then converted to a 4-point scale (very bad =3, a little bad =2, a little good = 1, very good = 0).

Following each *moral judgment question,* the child was also asked a *justification* question: “Why do you think it was good/bad to [*mib*]?” Children’s answers were later transcribed via video recordings of the sessions. For each country, a second coder, who was blind to condition and hypotheses, coded a third of the justifications. Reliability coding was then calculated on the basis of these justifications. Interrater reliability was high for both samples (Chinese sample $\boldsymbol{\kappa}$ = 97.7%; U.S. sample $\boldsymbol{\kappa}$ = 96.6%). Discrepancies were resolved through discussion.

***Transgressor evaluation task.*** Afterwards, the experimenter asked the child a few questions about the transgressor in each trial. First, the experimenter asked a *deserved punishment question* (adapted from Smetana et al., 2012), for example: "What do you think should happen to [Devon]? Should [Devon] get in trouble?" and if yes, "Should s/he get in a little trouble or a lot or trouble?" Children's responses were coded into a 3-point scale (a lot of trouble =2, a little trouble = 1, no trouble = 0).

The experimenter then asked *play questions*: "If [Devon] is here, would you want to play with him/her? Are you very sure or only a bit sure?" (scoring: very sure will play = 0, a little sure will play = 1, a little sure won't play = 2, very sure won't play = 3). These questions were followed by the *explicit liking questions*: "What do you think of [Devon]? Do you like him/her or not like him/her? Do you kind of (don't) like or really (don't) like [Devon]?" Responses were scored from 0 (really like) to 3 (really don't like).

***True belief task***. To assess whether children who accepted adult testimony were merely complying with the informant or had genuinely revised their beliefs, a *true belief* task was included to probe whether children judged the permissibility of the same actions in a new context, and whether to transmit their judgment to others.

In this task, the experimenter first asked children *context transfer questions* to examine whether children's moral judgments about the actions could be transferred to other contexts, for example: "Now let's think about a different situation. Let's say [Devon] was at [the playground] and wanted to [*mib* another classmate who was playing in the sandpit]. Would it be good or bad to [*mib* at the playground]? Would it be a little good/bad or very good/bad?" (scoring: very bad =3, very good = 0).

Then, to assess whether children were willing to transmit their moral judgments, the experimenter introduced children to a child puppet who matched the child's gender named "Sam". In in-person sessions in the U.S., the puppet was operated by the experimenter. In all other sessions, children saw pictures of the same puppet talking in a puppet voice on the computer. In each trial, the experimenter told the child that neither Sam nor herself knew anything about the novel action and asked the child an *information transmission question*: "Sam would like to learn about *mibbing*. Can you tell him whether *mibbing* was good or bad?" and then "Was it a little good/bad or very good/bad?" (scoring: very bad =3, very good = 0). Whatever the child said, the puppet then thanked the child for telling him/her about the novel action.

***Novel word comprehension question.*** At the end of each trial, to assess how children conceptualized the novel actions, the experimenter also presented children with three pictures and asked, for example: "What do you think *mibbing* means in this story? What do you think happened in this story?" The experimenter then described the three pictures for children to choose from in turn, with one picture depicting a familiar moral transgression (e.g., Devon taking the toy away from Casey), one depicting a familiar prosocial action (e.g., Devon giving Casey a gift) and the other depicting a neutral action (e.g., Devon reading a book next to Casey, see Table 4 in Appendix for the full list of actions). Children then chose the picture that they think looked most like the novel action.

***Parent questionnaires.*** To explore how parenting values might affect the extent to which children deferred to the counter-intuitive testimony, parents and adult participants filled out two questionnaires: parent authoritarianism ( e.g.,"obedience" versus "self-reliance", Feldman & Stenner, 1997) and parent social conformity (e.g.,

"Society should aim to protect citizens' right to live any way they choose" versus "It is important to enforce the community's standards of right and wrong", Stenner, 2005). For the Chinese parents, the questionnaires were translated from English to Chinese by a native Chinese speaker who was proficient in English, and then back-translated by a second bilingual research assistant.

**Results**

Because participants' responses across the *moral judgment* questions, *transgression evaluation* task and *true belief* task were all categorical and ordered, I fitted cumulative link mixed models (CLMM) with Laplace approximation. Unlike linear regression, ordinal regression models could allow for the retention of the order information in children's ratings of the novel behaviors and transgressors without assuming that they were equally distant (Bauer, & Sterba, 2011; Finch, Bolin, & Kelley, 2014). For CLMM fitting, I used the *Ordinal* package (Christensen, 2015) in R (version 1.2.5019; R Core Team, 2019). Unless otherwise stated, for each task, I included the fixed effects of Condition (Utilitarian vs. Authoritarian vs. Control, Control was used as the reference group), children's Age (continuous), and Culture (Chinese vs. U.S., Chinese was used as the reference group). Since each participant received a total of three trials, participant ID was included as a random effect. I then included interaction terms for the fixed effects sequentially, and checked whether each interaction term significantly improved model fit. The best CLMM models were chosen by performing likelihood ratio tests, which sequentially compared whether adding the interaction terms changed the amount of explained variance in participants' judgments (using R function *anova*). I also compared the overall model fit by using adjusted $R^2$ (Nagelkerke). Separate models were

also fitted for adults. Since the data was ordinal, Mann-Whitney tests were used for follow-up pairwise comparisons. Preliminary analyses revealed no effects of gender or trial number (first to third trial). Thus, reported analyses collapse across these factors.

**Moral judgment questions.** Before examining the full scale of participants' moral judgment responses, I analyzed participants' binary judgment (i.e., the action is bad = 1, the action is good = 0) with Generalized Linear Mixed-effect Models (GLMMs). The GLMM with the adult sample revealed no main effect of Condition or Culture, indicating that adults from both countries judged the novel actions as bad, regardless of the Condition they were assigned to (Utilitarian: 98.9%, Authoritarian: 94.4%, Control: 99.4%). For children, the best model revealed significant main effects of Condition (Authoritarian: OR = 0.05, 95% CI = [0.02, 0.12]; $p < .001$; Utilitarian: OR = 0.02, 95% CI = [0.009, 0.05]; $p < .001$) and Age (OR = 5.2, 95% CI = [3.12, 8.67]; $p < .001$), and no significant effect of Culture or interaction effects. These results suggest that children across two cultures were more likely to judge the novel actions as bad with age. Moreover, children from both countries were less likely to judge the novel actions as bad in the Authoritarian (57.4%) and Utilitarian conditions (48.2%) compared to the Control condition (88.4%). Follow-up Pearson chi-square test also revealed that the proportion of responses judging the action to be "bad" was higher in the Authoritarian condition compared to the Utilitarian condition, $\chi^2$ (1, $N = 764$) = 6.47, $p = 0.01$, Cramér's $V = 0.09$.

When examining the full-scale *moral judgment* scores (very bad = 3, very good = 0), I found that for the adult participants, only 2.41% of the responses indicated that the novel action was "a little good" (13 out of 540; 10 responses were from the Authoritarian

condition, 2 responses were from the Utilitarian condition, and 1 response was from the Control condition), and there were no responses indicating that the action was "very good". The CLMM analyses with the adult participants revealed significant interactions between Condition and Culture (Authoritarian × U.S., OR = 0.08, 95% CI = [0, 0.82], $p$ = 0.03; Utilitarian × U.S, OR = 0.002, 95% CI = [0, 0.002], $p$ <.001). Follow-up Mann-Whitney tests with Benjamini-Hochberg adjustment revealed that Chinese adults rated the novel actions less negatively then U.S. adults in the Authoritarian ($M_{\text{Chinese}}$ = 2.23, $M_{\text{U.S.}}$= 2.6, $W$ = 3210.5, $p$ = 0.018) and Control ($M_{\text{Chinese}}$ = 2.29, $M_{\text{U.S.}}$= 2.86, $W$ = 3306.5 , $p$ = 0.018) conditions, but not in the Utilitarian condition ($M_{\text{Chinese}}$ = 2.48, $M_{\text{U.S.}}$= 2.53, *ns*) .

For the main analysis of children's *moral judgment* score, a mixed-effects ordinal logistic regression model which included two interaction terms yielded the best fit for the data (compared to the null model, $LR$ (8) = 118.7, $p$ < 0.001, $pseudoR^2$ = 0.11(Nagelkerke)). This model revealed a significant two-way interaction between Age and Condition (Authoritarian), OR = 2.76 , 95% CI = [1.37, 5.57], $p$ = .005; a significant interaction between Culture and Condition (Utilitarian), OR = 3.01 , 95% CI = [1.11, 8.18], $p$ = .03; as well as main effects of Age (OR = 1.79, 95% CI = [1.08, 2.98], $p$ = .024) , Culture (OR = 0.44, 95% CI = [0.21, 0.90], $p$ = .025) and Condition (Authoritarian: OR = 0.12, 95% CI = [0.06, 0.26]; $p$ < .001; Utilitarian: OR = 0.08, 95% CI = [0.04, 0.16]; $p$ < .001).

To follow up on these interactions, I ran separate CLMMs on children's judgments in each culture, with Condition, Age and their interaction terms as fixed effects and participant ID as a random effect. For the Chinese children, I found a

significant main effect of Condition (Authoritarian: OR = 0.10, 95% CI = [0.05, 0.24]; $p$ < .001; Utilitarian: OR = 0.06, 95% CI = [0.03, 0.14]; $p$ < .001). There was also a significant two-way interaction between Age and Condition (Authoritarian), OR = 3.67, 95% CI = [1.17, 11.6]; $p$ = .03. To better understand age-related effects, I split the sample into three age groups. Mann-Whitney tests with BH adjustment revealed that 3- and 4-year-old Chinese children rated the novel transgressions as significantly more permissible both in the Utilitarian condition ($M_{3\text{-year-olds}}$ = 0.846 , $W$ = 1000.5, $p$ <0.001 ; $M_{4\text{-year-olds}}$ = 1.54 , $W$ = 5539.5 , $p$ <0.001) and the Authoritarian condition ($M_{3\text{-year-olds}}$ = 1.19, $W$ = 382.5, $p$ = 0.01 ; $M_{4\text{-year-olds}}$ = 1.44 , $W$ = 1648.5 , $p$ <0.001) compared to the Control condition ($M_{3\text{-year-olds}}$ = 1.94 ; $M_{4\text{-year-olds}}$ = 2.49), but their ratings in the two testimony conditions did not differ. For the 5-year-old Chinese children, children's moral judgment ratings in both the testimony conditions (Authoritarian: $M$ = 2.28, $W$ = 1395.5, $p$ = 0.02; Utilitarian: $M$ = 1.75, $W$ = 2625, $p$ <0.001) still significantly differed from the Control condition ($M$ = 2.49), but children's moral judgments in Utilitarian explanations had a significantly stronger effect compared to the Authoritarian condition ($W$ = 2164.5, $p$ = 0.002).

The CLMM with the U.S. children also revealed a significant main effect of Condition (Authoritarian: OR = 0.33, 95% CI = [0.17, 0.62]; $p$ < .001; Utilitarian: OR = 0.26, 95% CI = [0.13, 0.49]; $p$ < .001) and a marginally significant Age × Condition interaction, OR = 2.24, 95% CI = [0.93, 5.41]; $p$ = .07. Post-hoc adjusted Mann-Whitney tests revealed that that similar with their Chinese peers, 3-year-old U.S. children also rated the novel transgressions more positively both in the Utilitarian condition ($M_{3\text{-year-olds}}$ = 1.12, $W$ = 2413 , $p$ =.0001) and the Authoritarian condition ($M_{3\text{-year-olds}}$ = 1.13 , $W$ =

1039.5, $p$ =.0001 ; $M_{4\text{-year-olds}}$ = 1.68 , $W$ = 2472.5 , $p$ = 0.046 ) compared to the Control condition ($M_{3\text{-year-olds}}$ = 1.95; $M_{4\text{-year-olds}}$ = 2.03 ), and children's judgments in the two testimony conditions were not significantly different. The two types of explanations also affected 4-year-old children's moral judgments ($M_{4\text{-year-olds Utilitarian}}$ = 1.63, $M_{4\text{-year-olds Authoritarian}}$ = 1.68, $M_{4\text{-year-olds Control}}$ = 2.03) but the effects were marginal after the BH adjustment (Utilitarian vs. Control, $W$ = 3343.5, $p$ = 0.069; Authoritarian vs. Control, $W$ = 2472.5 , $p$ = 0.069). For the 5-year-old U.S. children, Utilitarian explanations still had an powerful effect on children's moral judgments compared to the Control condition (Utilitarian: $M$ = 1.78, Control: $M$ = 2.38, $W$ = 2625 , $p$ =.0005), but children's judgments in the Authoritarian condition did not significantly differ from their judgments in the Control condition ($M$ = 2.02, $p$ = 0.12), indicating that Authoritarian explanations were no longer effective in altering children's moral judgments.

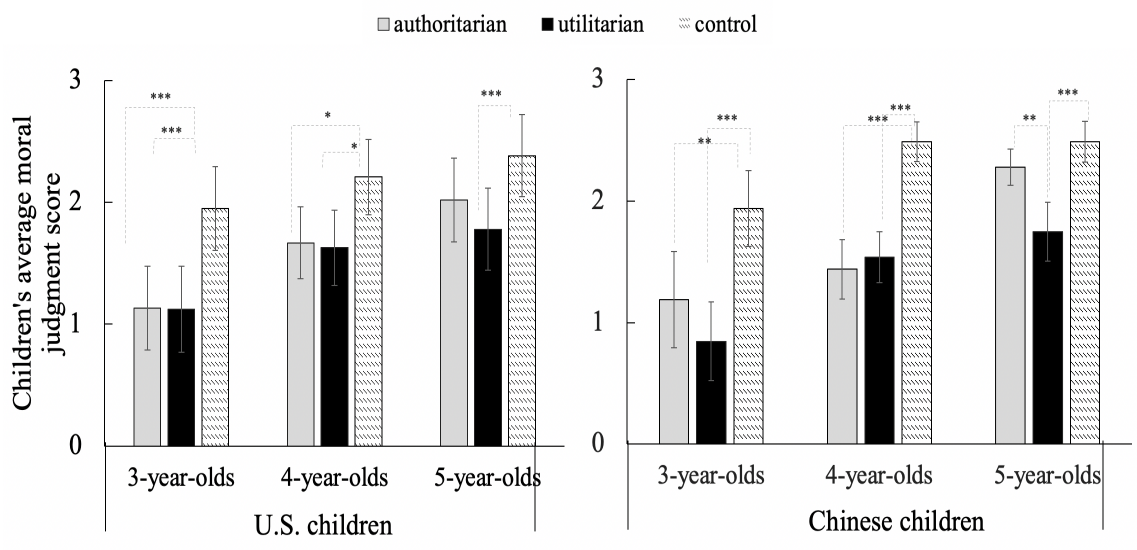


*Figure 2.* Children's Average Moral Rating Scores (very bad = 3, very good = 0) as a Function of Condition and Age Group in Study 1.

**Children's justifications of their moral judgments.** When asked to offer an explanation for their moral judgments, child participants gave an explanation in 72.14% of the trials in the Utilitarian condition, 67.19% of the trials in the Authoritarian condition and 69.62% of the trials in the Control condition. The children who did not give a justification either said "I don't know" or remained silent.

All children's justifications for negative judgments could be assigned to one of six categories: (1) Harm to victim; (2) Familiar moral transgressions; (3) Suggestion of positive actions; (4) Transgressor punishment; (5) Restatement; and (6) Irrelevant. Children's justifications of positive judgments could also be assigned to one of six categories: (1) Fairness; (2) Utilitarian explanations; (3) Testimony or authority; (4) Transgressor neutral/good behavior; (5) Restatement; and (6) Irrelevant (see definitions and examples of the categories in Table 1 and 2).

Table 1. *Children' Justification Categories for Negative Moral Judgments in Study 1.*

| Category Name | Description | Example |
|---|---|---|
| Harm to victim | Stating that the action would hurt victim's feelings, or get the victim physically hurt. | "Because her classmate was crying."; "Because it kinds of hurts her feelings." |
| Familiar moral transgressions | Transforming the story into a familiar moral transgression such as taking things that belong to the victim, bullying, or relating to one's own experience that involves a familiar moral transgression. | "Because she took something away"; "Because it's kind of like hitting or hurting people" |
| Suggestion of positive actions | Proposing alternative prosocial actions that the transgressor should have done such as asking for permission first. | "Because she should've asked before mibbing for the toy" "Because you should negotiate with he/her. If he/she agrees that you could play, then you can play." "Because as an older sister, you should let your younger sister have it." |

| | | |
|---|---|---|
| Transgressor punishment | Stressing the negative consequences facing the transgressor. | "Because if you take other people's toys you might get in trouble and get a time out."; "If Bubu takes her sister's ice cream her mom will never buy her ice cream again." |
| Restatement | Restating their negative evaluation of the action, or commenting on the transgressor's bad moral characters | "Because it's mean."; "Because it's not nice." |
| Irrelevant | Giving other justifications that are not listed above or not related to the vignette. | "Because my new dog mibbed" |

Table 2. *Children' Justification Categories for Positive Moral Judgments in Study 1.*

| Category Name | Description | Example |
|---|---|---|
| Fairness | Appealing to fairness concerns related to how the victim needs to share or how the transgressor should be able to take a turn | "Because you need to share your food." "Because they chose their ice creams without discussing it first." |
| Utilitarian explanations | Emphasizing good consequences for the transgressor or how they would have positive feelings | "Because then both children can play, they will be so happy!" "Because he could eat the ice cream himself after taking it." |
| Testimony or authority | Referring to the testimony they heard, or that the informant has the authority to decide whether the action was good. | "Because she said it's good." "Because the lady in it gets to decide." 'Because I had to listen to that lady." |

| | | |
|---|---|---|
| Transgressor Neutral/Good Behavior | Explaining transgressor's behavior in a neutral or a good way and downplaying the harm, or adding details that would make the action acceptable. | "Because he will return it soon after playing with it."; "Because he has asked her permission."; "Because he didn't mean it"; "Because (victim) actually didn't cry, she was happy"; "Because she is doing that for her good friend.";"It is just a tiny thing." |
| Restatement | Restating that their positive evaluation of the action, or commenting on the transgressor's good moral characters | "Because tamming is good for you."; "Because he is a good boy." |
| Irrelevant | Giving other justifications that are not listed above or not related to the vignette. | "Because he didn't hit me." |

Importantly, one exploratory question is whether children in the Utilitarian condition would be more likely to give *utilitarian* explanations that are similar to the adult informants', and whether children in the Authoritarian condition would be more likely to use *testimony or authority* explanations that cite the informant's authority. Indeed, I found that a sizable proportion of responses in the Utilitarian condition that used *utilitarian* explanations that focused on the good consequences for the transgressor (51% of the responses for the Chinese children and 31% of responses for the U.S. children in the Utilitarian condition, see Table 3). By contrast, children in the Authoritarian condition did not overwhelmingly cite the informant or her authority (15.3% of the responses for the Chinese children and 16.7% of responses for the U.S. children in the Authoritarian condition).

Table 3. *Participants' Justifications for Positive Moral Judgments Across Countries and Conditions in Study 1.*

| | Chinese children | | | U.S. children | | |
|---|---|---|---|---|---|---|
| | Authoritarian | Utilitarian | Control | Authoritarian | Utilitarian | Control |
| Fairness | 4.2% | 3.9% | 0.0% | 0.0% | 8.0% | 6.5% |
| Utilitarian explanation | 2.8% | 51.0% | 8.3% | 4.4% | 31.0% | 0.0% |
| Testimony or authority | 15.3% | 2.0% | 0.0% | 16.7% | 0.0% | 0.0% |
| Transgressor neutral/positive behavior | 15.3% | 8.8% | 0.0% | 1.1% | 6.0% | 9.7% |
| Restatement | 11.1% | 4.9% | 8.3% | 21.1% | 16.0% | 19.4% |
| Irrelevant | 15.3% | 2.0% | 8.3% | 4.4% | 2.0% | 16.1% |
| No response | 36.1% | 27.5% | 75.0% | 52.2% | 37.0% | 48.4% |

**Transgressor evaluation task.** For this task, I analyzed children's responses to the *deserved punishment questions* (e.g.,"Should [Devon] get in trouble?" ), *play questions* (e.g., "If [Devon] is here, would you want to play with her?") and *explicit liking questions* (e.g.,"What do you think of [Devon]? Do you like her or not like her?") separately with CLMMs.

***Deserved punishment.*** I coded participants' responses to the *deserved punishment questions* into a three-point scale (a lot of trouble =2, a little trouble = 1, no trouble = 0). For adults, I found a moderate effect of Culture, OR = 7.87, 95% CI = [1.3, 47.8]; $p$ = .025, indicating that Chinese adults were less likely to judge that the transgressor should get into trouble compared to US adults ($M_{\text{Chinese}}$ = 0.967, $M_{\text{U.S.}}$ = 1.3). There were no effects of Condition, indicating that neither type of explanation affected adults' judgment about the deserved punishment for the transgressor. For children, the best model revealed a main effect of Condition (Utilitarian), OR = 0.00001, 95% CI = [0, 0.134]; $p$ = .02, which was subsumed by a significant Age × Condition (Utilitarian) interaction, OR = 0.255, 95% CI = [0.07, 0.82]; $p$ = 0.02. With age, children judged the

transgressor to be more deserving of punishment in the Control and Authoritarian Conditions, but they suggested milder punishment with age in the Utilitarian Condition (see Figure 3), indicating that the Utilitarian explanations might be more effective for older children. Post-hoc adjusted Mann-Whitney tests for each age group found that 3- and 4-year-old children's responses did not significantly differ across the conditions (3-year-olds: $M_{\text{Utilitarian}} = 1.03$ , $M_{\text{Authoritarian}} = 0.771$ , $M_{\text{Control}} = 0.828$ ; 4-year-olds: $M_{\text{Utilitarian}} = 0.893$ , $M_{\text{Authoritarian}} = 1.18$ , $M_{\text{Control}} = 1.24$ ); but 5-year-old children were less likely to punish the transgressor in the Utilitarian condition than the Control condition ($M_{\text{Utilitarian}} = 0.85$ , $M_{\text{Authoritarian}} = 1.18$ , $M_{\text{Control}} = 1.24$), $p = 0.001$, again indicating that the Utilitarian explanation (but not Authoritarian explanation) affected 5-year-olds' judgment about deserved punishment for the transgressor.

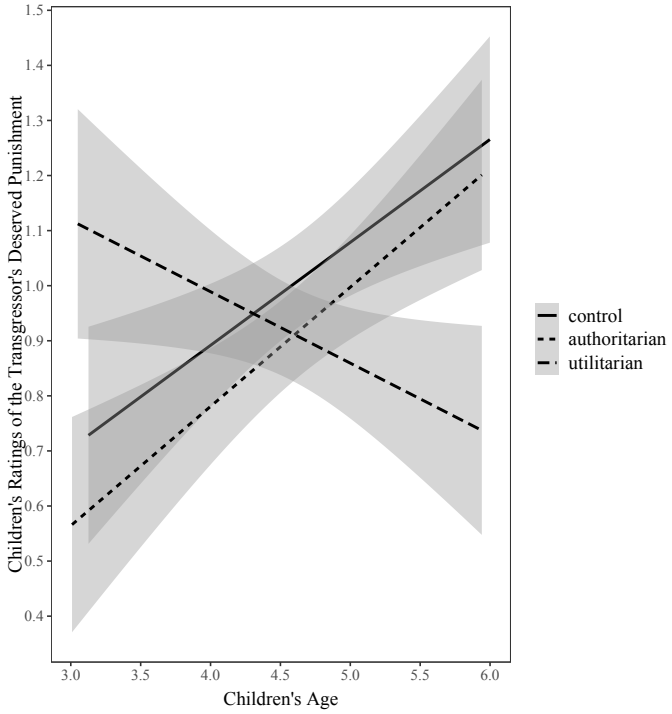


*Figure 3.* Children's Ratings of the Transgressors' Deserved Punishment as a Function of Age in Study 1. Separate Line Types Indicate Children's Deserved Punishment Responses in Each Condition (Utilitarian, Authoritarian, Control). Gray Shading Indicates Standard Errors.

***Play question.*** Participants' responses to the *play questions* were scored on a 4-point scale (very sure about not playing with transgressor = 3, very sure about playing with transgressor = 0). For adults, I found a significant main effect of Culture, OR = 15.7, 95% CI = [3.36, 773.3]; $p = 0.0005$, and no other main effects or interactions. Interestingly, Chinese adults were more likely to indicate that they wanted to affiliate with the transgressor compared to U.S. adults ($M_{\text{Chinese}} = 2.03$, $M_{\text{U.S.}} = 2.38$).

For children, I found a significant main effect of Age, OR = 3.43, 95% CI = [2.23, 5.26]; $p < 0.001$, indicating that with age, children were increasingly less likely to say that they wanted to play with the transgressor. I also found a moderate Condition (Utilitarian) × Culture interaction, OR = 4.88, 95% CI = [1.14, 20.9]; $p = 0.03$. Follow-up BH corrected Mann-Whitney tests revealed that Chinese children had lower play scores in the Utilitarian condition ($M = 1.42$) than in the Control condition ($M = 1.75$), $p = 0.01$. No significant differences were found in U.S. American children's responses to this question across the conditions.

***Liking question.*** Similar to children's responses to the *play questions,* children's responses to how much they liked the transgressors were also scored on a 4-point scale (really don't like =3, really like = 0). For adults, I did not find any effects of Condition (OR = 0.75, 95% CI = [0.12, 4.84], *ns*) or Culture (OR = 1.47, 95% CI = [0.32, 6.71], *ns*) , indicating that neither types of testimony affected adults' explicit liking judgments.

For children, the best CLMM revealed a significant main effect of Condition (Utilitarian), OR = 0.20, 95% CI = [0.07, 0.59], $p = 0.003$. There was also a significant main effect of Age, OR = 1.95, 95% CI = [1.28, 2.97], $p = 0.002$, indicating that in general, children reported that they liked the transgressors less with age. This model also

revealed a significant interaction between Condition (Utilitarian) and Culture, OR = 11.3, 95% CI = [2.6, 49.3], $p = 0.001$. Follow-up pairwise Mann-Whitney tests exploring the effects of Condition separately for each country revealed that Chinese children reported liking the transgressor more in the Utilitarian condition ($M = 1.34$) than in the Control condition ($M = 1.86$), $p < 0.001$. By contrast, the difference between the Utilitarian ($M = 1.56$) and Control ($M = 1.3$) conditions was only marginal for the U.S. children, $p = 0.073$.

**True belief task.** The *true belief task* was included to explore whether children's initial moral judgments would transfer to a new context, and whether children would transmit their initial judgment to others.

***Context transfer.*** In this task, participants evaluated the permissibility of the novel actions in new contexts (very bad =3, very good = 0). When pulling adults' moral judgment responses across all three trials, only 5.93% of the responses indicated that the novel action was "a little good" (32 out of 540; 17 responses were from the Authoritarian condition, 3 responses were from the Utilitarian condition, and 12 response was from the Control condition), and 0.7% of the responses indicated that the action was "very good" (4 out of 540; 2 responses were from the Authoritarian condition, 1 response was from the Utilitarian condition, and 1 response was from the Control condition). The CLMM with adult participants revealed a significant Condition (Utilitarian) × Culture interaction, OR = 0.02, 95% CI = [0.002, 0.225], $p = 0.001$. When separate models were computed for each country, I found that U.S. adults had similar ratings across the three conditions, but Chinese adults rated the novel actions as less permissible in the Utilitarian condition ($M = 2.38$) than in the Control condition ($M = 2.07$), $p = 0.001$.

For children, there were significant effects of Age (OR = 3.3, 95% CI = [3.29, 3.32], $p$ <.001) and Condition (Authoritarian: OR = 0.42, 95% CI = [0.27, 0.65], $p$ <.001; Utilitarian: OR = 0.36, 95% CI = [0.24,0.55], $p$ <.001), but no effect of Culture and no significant interactions. Compared to the Control condition ($M$ = 2.12), children rated the novel actions less negatively in both the Authoritarian condition ($M$ = 1.777) and Utilitarian condition ($M$ = 1.72).

***Info-transmission.*** In this task, participants were asked to transmit their moral judgments to others. For adults, 5.37% of the responses indicated that the novel action was "a little good" (29 out of 540; 16 responses were from the Authoritarian condition, 4 responses were from the Utilitarian condition, and 9 response was from the Control condition), and 0.2% of the responses indicated that the action was "very good" (1 out of 540; the response was from the Authoritarian condition). Again, I found a significant Condition (Utilitarian) × Culture interaction, OR = 0.002, 95% CI = [0, 0.03], $p$ <0.001. Post-hoc tests for each country revealed that Chinese adults rated the novel actions as less permissible in the Utilitarian condition ($M$ = 2.38) than in the Control condition ($M$ = 2.07), $p$ = 0.02, but U.S. adults rated the novel actions less negatively in the Utilitarian condition ($M$ = 2.46) than in the Control condition ($M$ = 2.77), $p$ = 0.004. However, these results were likely driven by the small percentage of adults who rated the actions as good.

For children, the pattern of results in the *info-transmission* task aligned with the *context transfer* task. There were significant effects of Age (OR = 3.46, 95% CI = [2.42, 4.96], $p$ <0.001) and Condition (Authoritarian: OR = 0.17, 95% CI = [0.09, 0.32], $p$ <0.001; Utilitarian: OR = 0.21, 95% CI = [0.12, 0.4], $p$ <0.001), but there were no effects of Culture and no significant interactions. Specifically, children in both Authoritarian ($M$

= 1.6) and Utilitarian ($M$ = 1.65) conditions judged the novel moral actions to be more permissible than children in the Control condition ($M$ = 2.22).

**Novel word comprehension question.** At the end of each trial, I asked participants to choose whether they think the novel action was more like a familiar negative moral action, neutral action, or positive moral action. For the adult sample, participants' responses did not differ significantly between the three conditions (all $p$ = .74), and the majority of the participants chose the negative moral action across the three conditions (see Table 4).

When pooling children's responses across three trials, Pearson chi-square analysis revealed that the number of choices of positive actions was moderately affected by condition, $\chi^2$ (2, $N$ = 1137) = 7.26, $p$ = 0.027, Cramér's $V$ = 0.08. More specifically, there was a slightly larger percentage of positive choices in the Utilitarian condition (28.08%) and Authoritarian condition (28.20%) than in the control condition (20.7%). When children's choices were coded as binary (i.e., choosing the positive action = 1, not choosing the positive action = 0), a Generalized Linear Mixed- Effect Models (GLMM) with Age, Condition, and Culture as predictors revealed a main effect of Age (OR = 0.19, 95% CI = [0.11,0.33], $p$ = <0.001) and a moderate effect of Culture (OR = 2.27, 95% CI = [1.10, 4.69], $p$ = 0.027), indicating that children were less likely to choose the positive action with age, and that the likelihood of U.S. children choosing the positive action was higher than Chinese children. Here, there was no significant main effect of the Authoritarian Condition ($p$ = 0.10), and the main effect of the Utilitarian Condition was marginal(OR = 2.38, 95% CI = [0.99, 5.71], $p$ = 0.05). These findings indicate that adult

testimony may have affected children's interpretation of the actions only moderately in the *Utilitarian* condition.

Table 4. *Participants' Responses to the Novel Word Comprehension Questions Across Age Groups and Conditions in Study 1.*

| | Adults | | | Children | | |
|---|---|---|---|---|---|---|
| | Authoritarian | Utilitarian | Control | Authoritarian | Utilitarian | Control |
| Negative | 87.22% | 91.67% | 92.22% | 64.83% | 62.66% | 76.08% |
| Neutral | 2.78% | 1.66% | 1.11% | 7.09% | 7.14% | 3.22% |
| Positive | 10% | 6.67% | 6.67% | 28.08% | 28.20% | 20.70% |

**Parent Characteristics.** Parent surveys on authoritarianism and social conformity were collected as I was interested in whether it would predict children's level of acceptance in moral explanations. On average, Chinese adults ($M$= 1.67 out of 4, $SD$= 0.98) and parents ($M$= 1.43 out of 4, $SD$= 0.88) scored higher on authoritarianism compared to U.S. American adults ($M$= 1.09, $SD$= 1.20) and parents ($M$= 0.86, $SD$= 1.06). A linear regression using Group (adult participants, parents) and Culture (China, U.S., China as the reference group) as predictors and authoritarianism score as the response found a significant main effect of Culture, $B$ = -0.57, $SE(B)$ = 0.09, $p < .0001$, indicating that Chinese parents and adult participants had higher authoritarian scores than U.S. participants. I also found an effect of group ($B$ = -0.23, $SE(B)$ = 0.09, $p$ = .01), indicating that adult participants had higher authoritarianism scores than parents of the child participants.

I found similar patterns of results for the average social conformity scores, with Chinese participants (Chinese adults: $M$= 0.38 out of 1, $SD$= 0.19; Chinese parents: $M$=

0.34 out of 1 , *SD*= 0.17) having higher average social conformity scores than U.S. participants (U.S. adults: *M*= 0.25 , *SD*= 0.19; U.S. parents: *M*= 0.27, *SD*= 0.19), *B* = -0.09*, SE(B)* = 0.02*, p* < .001. As expected, social conformity was highly positively correlated with authoritarianism (*r* = 0.54, *p* < .001).

However, parental authoritarianism and social conformity scores were not significantly correlated with children's moral judgment responses. When authoritarianism and social conformity were added as predictors and compared with the main models, neither of them significantly improved the fit of the models, indicating that parental authoritarianism and social conformity values did not affect children's acceptance of counterintuitive moral explanations.

**Discussion**

The current study assessed whether 3- to 5-year-old Chinese and U.S. American children would judge novel, distress-inducing actions to be more morally permissible after an adult informant had given two types of counter-intuitive explanations: *Utilitarian* explanations that reason about consequences and *Authoritarian* explanations that emphasize the speaker's authority. I found that both types of explanations moved children's (but not adults') moral judgments of the permissibility of harm-related transgressions. Children's context-transfer judgments and transmission of the judgments in the *true belief* task were similarly affected by the two types of explanations, indicating that children's responses may not have merely signaled compliance to the social demands of the task, but likely reflected genuine changes in their moral judgment. Furthermore, with age, children were more selective in their trust in adult explanations, such that they were more likely to defer to the counter-intuitive claims when provided with Utilitarian explanations compared to Authoritarian explanations.

The main finding that children's moral judgments were significantly more positive in the two testimony conditions compared to the control condition extended previous work on the influences of testimony on children's moral judgments (Li et al., 2019; Rottman et al., 2017). Specifically, the finding that both *Utilitarian* and *Authoritarian* explanations moved children's moral judgments on harm-related transgressions fills an important gap in the literature on children's testimonial learning, which is primarily focused on children's selective trust in others in the non-social domains (e.g., word learning, science and religion) instead of on the transmission of moral beliefs, attitudes and norms. On the surface, this finding seems to stand in contrast with the mainstream views in the developmental literature that preschool-aged children's understanding of moral norms is linked to the nature of an act and its consequences for a victim, independent of prescription by local authorities (Smetana, 1981; Smetana et al., 2012). Thus, an important question concerns why children judged the novel actions as more morally permissible upon hearing the adult's explanations. Based on the finding that children did not overwhelmingly interpret the novel action as positive when answering the *novel word comprehensio*n questions, it is unlikely that children have reinterpreted the action amorally or even positively. Rather, the explicit causal link between the novel actions and the apparent distress of the victim should be sufficient to indicate that harm-related moral transgressions were up for evaluation.

Infant studies (Hamlin et al., 2007; Kuhlmeier et al.,2003) as well as work conducted by Social Domain Theorists (Smetana, 2006) suggest that children understand basic, harm-based principles from an early age. Since this pattern is robust, children's commitment to the anti-harm principle was unlikely to be abandoned upon hearing

counterintuitive adult testimony. Rather, it is possible that children incorporated the explanations they heard in their moral evaluations of the novel scenarios. In the *Utilitarian* condition, testimony may have guided children to appreciate and weigh the different considerations at stake, leading them to minimize the distress caused to the victim and prioritize the positive outcomes for the transgressor. This possibility is supported by children's explanations in this condition, which were consisted of either *utilitarian* explanations stating how the transgressors got what they wanted (e.g., "Because he could eat the ice cream."), or focused on *transgressor neutral/positive behavior*, which downplayed the transgression or the harm caused (e.g., "It is just a tiny thing."). Relatedly, in the *Authoritarian* condition, children may have trusted adult authority on the interpretations of ambiguous harm-related actions, and reasoned that inducing distress in these novel contexts were more permissible than familiar moral transgressions such as pushing or stealing.

Another question I explored was whether the two types of explanations posed the same level of influence on children's moral judgments and evaluations of the transgressors. The results revealed an interesting developmental finding that 3- and 4-year-old Chinese children and 3-year-old U.S. children were equally affected by the two types of explanations, but 5-year-olds from both countries were more influenced by *Utilitarian* explanations (which emphasized the positive consequences for the transgressor), and they were less likely to be convinced by *Authoritarian* explanations (in which they were asked to believe on an adult's authority). This pattern of results was evident in children's moral judgment responses, as well as their judgments about the deserved punishment of the transgressor. This developmental trajectory is generally

consistent with children's preferences for high-quality arguments about contingent facts from 3- to 5-years of age (e.g., Corriveau & Kurkul, 2014; Koenig, 2012; Mercier, Bernard, & Clément, 2014), which revealed that older preschoolers (4- and 5-year-olds) were more attuned to argument circularity, whereas younger children (3-year-olds) are only selective under certain circumstances (Corriveau, & Kurkul, 2014).

Why would *Utilitarian*, but not *Authoritarian* explanations appeal to older children? One possible explanation is that 5-year-old children have started to appreciate that moral norms were obligatory and governed by social and cultural groups (Rhodes & Wellman, 2017; Tomasello, 2020), and unalterable by individual beliefs and desires (Chernyak, Kang, & Kushnir, 2019; Heiphetz, & Young, 2017). Thus, it may make less sense for a single individual with unspecified credibility to be able to "decide" whether an action was good or bad in the *Authoritarian* condition. In addition, it is possible that older children may have started to appreciate that a certain level of autonomy was needed in making moral judgments, so merely believing something on someone's authority may not be enough to move their moral judgments. Although from the perspective of a mature moral agent, the *Utilitarian* explanation may be similarly morally illegitimate, it nevertheless guided children to weigh the possible benefits of a moral transgression heavier than the costs to the victim. A third and not mutually exclusive possibility is that 5-year-old children may have treated the explanations in the *Authoritaria*n condition as circular because the informant did not actually provide a reason beyond exerting her power over the issue, leading children to treat the reason as lower quality than the *Utilitarian* explanations. To explore these possibilities, additional research is needed to explore other types of explanations that might be more or less effective in changing

children's moral judgments over the course of development. For example, future research could explore whether compared to younger children, older children are more convinced by explanations that emphasize collective authority (e.g., "Because that's what we decided." Tomasello, 2019) compared to explanations that exert individual authority. Future studies can also provide children with more complex situations (e.g., moral dilemmas, Levine, Mikhail, & Leslie, 2018), and assess whether other types of utilitarian arguments which stressed that the transgression could be for the greater good, or the victim's own good (e.g., "It is good because it could save more people.") would be prioritized by older children and adults.

While there were no significant cross-cultural differences between adults' judgments, I found that unlike their U.S. counterparts, 4- and 5-year-old Chinese children were more receptive towards Authoritarian explanations when making moral judgments. Given that Chinese parents scored higher on authoritarianism, it is possible that children in China placed more value on authority and conformity-related considerations, and that early socialization practices that emphasize the importance of conformity and humility are key mechanisms by which young Chinese children come to display such differences (e.g., Corriveau, Min, & Kurkul, 2014; Li, 2005; Suizzo & Cheng, 2007; Wang et al., 2008). However, because of the post-hoc nature of these effects, these results should be taken with caution. Future studies could make the authority of the speaker more salient, which can be achieved either by changing the informants' identity visually (e.g., presenting children with a more senior person instead of an informant in her early twenties), or giving more explicit verbal cues when introducing the informant (e.g., having the experimenter introduce the informant as someone who is putatively granted a

superior status such as teachers), and see if Chinese children might be more receptive to these authority-related cues.

Although there is observational evidence suggesting that parent-child conversations may support moral development (Dahl, & Brownell, 2019), the role of adult testimonial influence has rarely been examined experimentally. Because children depend greatly on others to learn about the world around them, it is important that we understand the mechanisms by which they actively optimize their social learning in moral contexts. The results in Study 1 demonstrate that 3- to 5-year-old Chinese and U.S. children judged novel, distress-inducing actions to be morally wrong independently, but they shifted their judgments upon hearing counterintuitive explanations that either reasoned about consequences for the transgressor or emphasized the informant's authority. With age, children's moral judgments were less influenced by Authoritarian explanations, but Utilitarian explanations continued to exert a powerful effect. This study is one of the first to explore the types of moral explanations that can guide children to believe what they are told, and the developmental and cultural differences in children's acceptance of different types of explanations commonly utilized in moral education (Leman, 2005).

# CHAPTER 3

## Study 2: Children's Evaluations on Sources of Moral and Empirical Knowledge

### Introduction

As reviewed, the breadth and depth of children's knowledge substantially depends upon the testimony provided by others (Harris et al., 2018; Mills, 2013; Tong, Wang, & Danovitch, 2020). From testimony supplied by adults, children acquire a wealth of

factual knowledge, including the meaning of new words (Koenig, Clément, & Harris, 2004; Sobel, & Corriveau, 2010; Stephens, & Koenig, 2015), function of artifacts (Lane, & Harris, 2015), object location (Jaswal, Croft, Setia, & Cole, 2010; Liu, Vanderbilt, & Heyman, 2013), scientific facts such as the habitats of animals and the existence of germs (Landrum, & Mills, 2015; Clegg, Cui, Harris, & Corriveau, 2019), as well as culturally-specific phenomenon such as religious norms and entities (Harris, & Koenig, 2006). Although testimony has been treated as an undisputed source of empirical knowledge, an important question concerns how children evaluate testimony as a way to justify knowledge. It also remains unclear whether children can appreciate the limits of testimony when it comes to one less descriptive type of knowledge: moral knowledge. Here, I aim to assess whether reliance on testimony, a source of second-hand knowledge, is treated as epistemically inferior to acquiring knowledge for oneself. I also explore whether children and adults treat deference to testimony in the moral domain differently than deference in other domains of empirical knowledge.

**Reliance on testimony versus independent thinking.** Children acquire knowledge about the world in a variety of ways. From a young age, children can identify and to selectively learn from competent or knowledgeable sources over less knowledgeable ones (e.g., Birch, Vauthier, & Bloom, 2008; Koenig et al., 2004). At the same time, they can gather evidence through exploration (Bonawitz et al., 2011; Schulz & Bonawitz, 2007; Yu, Landrum, Bonawitz, & Shafto, 2018), experimentation (Cook, Goodman, & Schulz, 2011; Lapidow, & Walker, 2020), and question-asking (Callanan & Oakes, 1992; Kurkul & Corriveau, 2017; for a review, see Ronfard, Zambrana, Hermansen, & Kelemen, 2018). However, little is known about *children's own*

*perspectives* on the best ways to acquire knowledge. Specifically, how do children evaluate testimony as a source of knowledge, especially when compared to acquiring knowledge on one's own?

When children are on the receiving end of information, much recent work on testimonial learning has shown that children are able to calibrate self-reliance and dependence on testimony based on their level of confidence. When adult claims violate children's causal or perceptual knowledge, children often show strong resistance to testimony and favor their own first-hand perception and intuitions (Koenig & Echols, 2003; Jaswal, 2004; Ma & Ganea, 2010, for a review, see Lane, & Harris, 2014). For example, toddlers actively correct adults who mislabel common objects (Koenig & Echols, 2003). Three- to 8-year-old children often distrust adult testimony that runs counter to their own beliefs about the causal structure of the word (e.g., a ball going through a solid wall, Lane & Harris, 2014; Woolley & Ghossainy, 2013; Shtulman, 2009). When faced with counter-perceptual or counter-intuitive claims with a certain level of ambiguity (e.g., the identity of a hybrid object that looked like both a spoon and a key, or the contents of an opaque box), however, preschool children can also go against their own intuition and readily accept adult testimony (e.g., Chan & Tardif, 2013; Hermansen, Ronfard, Harris, & Zambrana, 2021; Jaswal, 2004; Jaswal & Markman, 2007; Ma & Ganea, 2010; Ronfard, Chen, & Harris, 2021). In fact, when Chan and Tardif (2013) presented children with learning tasks that involve both familiar prototypical objects (e.g., a button) and ambiguous objects (e.g., a button that also looks like a wheel), Chinese and U.S. American children were more likely to accept adult

counterintuitive testimony when they had weaker prior knowledge about the objects, compared to cases when they had stronger conflicting intuitions.

Recent research has also begun to pit independent inference and testimony directly against each other, finding that children are flexible when it comes to balancing the use of their own reasoning and testimonial knowledge. When there are conflicts between children's own observation of statistical evidence and adult testimony, children take both the strength of observational evidence and the reliability of the informant into account when deciding what to trust (Bridgers et al., 2016; Gualtieri et al., 2019; McLoughlin et al., 2021). For example, when asked to predict the color of a dog's collar in a park, children trusted relevant statistical evidence (e.g., observing eight dogs wearing blue collars and two dogs wearing yellow) more than testimony when the speaker was presented as inaccurate (e.g., only accurate half of the time at predicting the colors of collars), but they overrode the observable inferential information and accepted what the speaker had said when she was presented as accurate and reliable (Gualtieri et al., 2019). Similarly, when deciding which block made a machine go, children trusted observational evidence more than testimony when the observational data is deterministic (e.g., seeing one block activating the machine 100% of the time), and relied more on testimony when the observed data was less conclusive (e.g., seeing one block activating the machine 66.67% of the time). In addition to acquiring factual knowledge, children also demonstrated similar levels of flexibility when it comes to learning social information and conventional norms. For example, when speaker reliability is unspecified, 7-year-old children relied more on first-hand observation more than second-hand gossip when

deciding whether an agent is helpful or not (Haux, Engelmann, Herrman, & Tomasello, 2016).

When children are asked to compare the two sources of knowledge as third party evaluators, they treat both independent reasoning (e.g., "Backpacks hold books. I think there is a backpack in there.") and learning from testimony by a reliable informant (e.g., "My teacher told me there's a ball in the box. I think there's a ball in there.") as reasonable ways of thinking (Koenig, 2012). However, there is some evidence showing that children sometimes prefer to learn from an informant who has independent knowledge compared to one who acquires second-hand knowledge from others. For example, 4 - and 5-year-old children in the UK preferred to seek new information from a speaker who accurately labeled familiar animals independently over a speaker who had always relied on the help of others (Einav & Robinson, 2011). It is possible that children were able to infer that while both informants were accurate, the speaker who was able to come to his own conclusions was more knowledgeable. Similarly, Einav (2017) found that when it comes to accepting claims about facts in a faraway country, 8- and 9-year-old children preferred to trust the information provided by an independent group of three adults (who responded privately) over that of a non-independent consensus (who had heard each other's answers). Moreover, children also preferred to learn from an agent who was able to figure out how to solve a problem (e.g., activating a music box) by herself compared to one who did so after hearing direct instructions (Bridgers, Gweon, Bretzke, Ruggeri, 2018). These results might indicate that while children treat both testimony and independent thinking as good sources of empirical knowledge, they might

have an emerging understanding that individuals who have independent knowledge are more epistemically competent than individuals with second-hand knowledge.

**The limits of testimony in the moral domain.** As reviewed above, much recent work has found that children use both independent reasoning and testimony as important sources of empirical knowledge. When these two sources of knowledge are pitted against each other, children's own inclinations to trust testimony often depend on the relative strengths of children's own prior knowledge and speaker reliability, and they sometimes treat an individual with first-hand empirical knowledge to be more competent than an individual with second-hand empirical knowledge. However, it remains unclear whether this preference to learn from independent thinkers extends to the moral domain. Is it reasonable to depend on testimony to justify one's moral beliefs? Should moral testimony be treated differently from testimony in the empirical domain?

As reviewed in Chapter 1, although no studies to date have investigated children's views on reliance on testimony in the moral realm, much developmental research has taken a constructivist or nativist approach, arguing that children either develop moral competencies independent of authority by actively assessing their social environment (Turiel,1983; Smetana,1981;1985; Smetana et al., 2012), or are equipped with nascent moral intuitions that are independent of social influences (Hamlin, Wynn, & Bloom, 2007; Kuhlmeier, Wynn, & Bloom, 2003; Geraci & Surian, 2011; Sloane, Baillargeon, & Premack, 2012). Since these developmental accounts prioritize the innate origins and authority-independent nature of moral development, these approaches leave open questions about the amount of weight children put on moral testimony.

The few studies that have explored children's acceptance of moral testimony have found similar patterns of results as in the empirical domain: children assess their level of uncertainty of the judgment when deciding whether to trust testimony. While children are less receptive to counterintuitive information about familiar harm-related moral actions (Kim, Chen, Smetana, & Greenberger, 2016), they readily depend on testimony when the given situation is more ambiguous (Rottman et al., 2017; Li et al., 2019). For example, children accepted adult testimony and concluded that an unfamiliar, self-directed action is harmful (e.g., painting one's own face white, Rottman et al., 2017), and decided that a novel action that induces distress in a new context is more morally permissible (Li et al., 2019). When making both empirical judgments (i.e., function of an object) and moral judgments (i.e, the exclusion of a peer), Chinese and Spanish preschoolers were more inclined to accept a counterintuitive opinion of a unanimous group of peers in the object function context than the social evolution context (Sebastián-Enesco, Guerrero, & Enesco, 2020). Thus, it is possible that there are domain-specific differences in children's acceptance of testimony (Harris, Koenig, Corriveau, & Jaswal, 2018). Do children judge testimony as a more legitimate source of empirical knowledge than as a source of moral knowledge?

To my knowledge, the only study that has directly assessed the legitimacy of relying on moral testimony was done by Andow (2019), who asked adults to evaluate the legitimacy of forming descriptive beliefs and moral beliefs on the basis of first-hand experience and testimony. Results revealed that beliefs formed in the testimony cases were generally evaluated as less legitimate than independently acquired beliefs. Moreover, moral beliefs formed on the basis of testimony were judged to be less

legitimate than testimonially transmitted beliefs about descriptive matters. Although this work has attempted to fill a gap in our understanding of the asymmetry between moral and descriptive testimony, it is not clear whether children, who rely heavily on testimony to acquire knowledge, would also demonstrate the same level of preference for independent thinking and experience.

**Cultural differences in independence and interdependence.** Children's social learning is deeply embedded in socially shaped learning experiences, as well as culturally shaped beliefs about what it means to be a good learner. Based on the great cultural variability in the extent to which children are invited, encouraged, and expected to participate in conversations in their own learning (e.g., Chavajay & Rogoff, 1999; Correa-Chavez & Rogoff, 2009; Rogoff, Correa-Chávez, & Silva, 2011), another interesting question concerns whether there are any cultural variations in the openness to testimony as a source of empirical and moral knowledge.

Comparisons of learning-related beliefs between Western and Eastern countries found that a Socratic approach to learning is generally endorsed in the U.S., which places an emphasis on active engagement, inquiry, and independent insights, whereas Chinese cultures tend to embrace a Confucianism tradition which stresses intellectual humility, listening attentively to sources of authorized knowledge, and speaking only when fully sure (Chan & Elliott, 2004; Pratt et al., 1999, .J. Li, 2005; J. Li & Fischer, 2004; 2007). Additionally, since China is often considered as a more collectivist culture than U.S., it is possible that there is a stronger emphasis on social relationships and connections, respectful deference, and accommodation to other people, whereas Western cultures might place a higher value on personal opinions and judgments (Chen & French, 2008;

Markus & Kitayama, 1991). As a result, these cultural differences may manifest in differences between Western and East Asian parental socialization with their children, with East Asian parents playing more values on conformity, power and intellectual modesty than Western parents (Suizzo & Cheng, 2007). In fact, recent developmental evidence has shown that children with East Asian heritages were often more likely to privilege information provided by an adult authority in social learning contexts (e.g., Corriveau & Harris, 2010; Corriveau et al., 2013; DiYanni et al., 2015; Li, Harris, & Koenig, 2019). On these bases, it is possible that parents might transmit their epistemic stance to their children, leading Chinese children to be more open to attributing knowledge to others, and to favor reliance on testimony over independent thinking across domains. However, there is also evidence showing that Chinese parents tend to engage in more didactic talk about moral standards and social norms than U.S. parents (Doan & Wang, 2010; Wang, Leichtman, & Davies, 2000; Wu, & Honig, 2010). Thus, this stronger emphasis on the learning of moral rules can also guide Chinese children to be more vigilant towards agents who need to rely on others for moral knowledge, leading them to prefer the agent with first-hand moral knowledge.

The current study was designed with three goals in mind. The primary goal was to determine whether children would treat justifying one's beliefs on the basis of testimony as a less desirable way of thinking than coming to one's own conclusions. Here, I asked Chinese and U.S. American children and adults to evaluate two speakers: one who showed independence in her thinking and the other who relied on testimony to make judgments. I then assessed whether participants would consistently credit coming to one's own conclusions as a better way of thinking, and whether participants would show

learning and social preferences for the independent speaker over the speaker who deferred to testimony. The second goal of the study is to explore whether children judge dependence on testimony in the moral domain as less legitimate than dependence on testimony in the empirical domain. In Study 2, participants were randomly assigned to one of two between-subjects conditions, a *Moral Knowledge condition* in which agents make judgments about familiar moral actions or an *Empirical Knowledge condition* in which agents make judgments about the contents of containers. My third goal was to explore whether there would be cross-cultural differences in children's judgments. To pinpoint specific environmental factors that can contribute to cross-cultural differences in children's deference to testimony, parents and adult participants from both countries completed questionnaires assessing parent authoritarianism (Feldman & Stenner, 1997) and independent versus interdependent cultural values (Singelis, 1994).

## Method

**Ethics Statement.** This study was approved by the Institutional Review Board at the University of Minnesota ("Sources of Knowledge", IRB ID: STUDY00010855). A parent or guardian provided consent. The data in the U.S. was collected between September 2019 and August 2020; the data in China was collected between June 2021 to August 2021. All children received written consent from their parents prior to participation.

**Participants.** Participants included 261 4- to 6-year-old children in the United States and China. The 128 U.S. participants ($M_{age}$ = 68.22 months; range = 54.24 – 83.78 months, 63 boys and 65 girls) were recruited through a university-managed database of families from the greater Twin Cities areas in Minnesota. The 133 Chinese participants

($M_{age}$ = 65.64 months; range = 48.36 – 84.30 months, 73 boys and 60 girls) were recruited through online social media platforms and word of mouth across 16 provinces in mainland China. I recruited 4- to 6-year-old children because these appear to be the ages at which children have an emerging understanding of various ways of thinking and reasoning, and the ages at which children start to show preferences for strong arguments concerning contingent facts (e.g., Corriveau & Kurkul, 2014; Koenig, 2012; Mercier, Bernard, & Clément, 2014).

Information on ethnicity collected from parental questionnaires showed that 87.5% of the U.S. American children were White (1.56% Black, 1.56% Asian, and 9.38% identified their child with multiple races or ethnicities), and 100% of the Chinese children were of Chinese descent (86.47% identified as being of Han ethnicity, 5.26% of Man ethnicity and 8.27% of other minority ethnicity groups including Daur, Hui, Korean, Miao, Mongolian, Tujia, Qiang and Yao). Self-reported social-economic status was high for U.S. parents: 10.94% of the mothers had a Doctorate, 47.66% a Master's degree, 33.59% a Bachelor's degree, 3.91% an Associate degree, and 3.91% completed some college. For the participants' fathers or other parents, 11.72% had a Doctorate, 22.66% a Master's degree, 45.31% a Bachelor's degree, 10.94% an Associate degree, 7.81% completed some college, and 1.51% had a high school diploma. U.S. parent reported income was primarily middle to upper class, with 72.13% of the parents reporting an annual income greater than $10,000 per year (122 out of 128, or 95.31% of parents answered this question). The Chinese sample was comparable in family background. Parents reported on the level of education they and their partner had completed (129 out of 133, or 97% of parents answered this question) and on their income level (168 out of

133, or 93.23% of parents answered this question). 7.52% of the mothers had a Doctorate, 19.55% a Master's degree, 52.63% a Bachelor's degree, 10.53% a Professional degree, and 4.51% finished elementary or secondary school. For the participant's fathers, 12.03% had a Doctorate, 16.54% a Master's degree, 51.13% a Bachelor's degree, 9.77% a Professional degree, and 7.81% completed some college, and 7.52% had a high school diploma or below. Family monthly income was measured on a 9-point scale ranging from less than 3000 RMB to more than 24,000 RMB per month (scale adapted from McBrid-Chang et al., 2012; Liu et al., 2018). The median monthly income is between 18,000 to 20,999 RMB, and approximately 69.17% of the families earned higher monthly incomes than the average for urban households across the country in 2021 (10,758.5 RMB; National Statistics Bureau, 2021, based on a family of three in a typical Chinese family). All Chinese and U.S. child participants were tested in online sessions via video-conferencing. Informed consent was obtained from the child's parents in advance of testing.

Adult comparison groups were also recruited to access whether adults had the same intuitions as children. Eighty-four U.S. residents ($M_{age}$ = 31.13; range = 20-57; 45 female, 33 male and 6 non-binary; 73.8% White, 10.7% Asian, 4.8% Black, 8.3% more than one race, 2.4% another race, 10.7% Hispanic of any race) were recruited online through *Prolific*. Seventy-nine Chinese residents ($M_{age}$ = 35.42; range = 20-57; 46 female and 33 male; all ethnically Chinese, 93.7% of the Han ethnicity) were recruited online through social media.

**Procedure.** Children were randomly assigned to two between-subjects conditions, a *Moral Knowledge* condition (i.e., in which agents make judgments about moral actions

such as pushing or helping others) or an *Empirical Knowledge* condition (i.e., in which agents make judgments about the hidden contents of containers, adapted from Koenig, 2012). The experimental session for each condition consisted of four tasks: A *reason evaluation* task, a *selective learning* task, a *selective preference* task and an *explicit judgment* task. The order of the last two tasks were counterbalanced. To pinpoint possible environmental factors that may contribute to cross-cultural differences in children's evaluations of reliance on testimony, parents also filled out questionnaires assessing their own authoritarianism tendency (Feldman, & Stenner, 1997) and independent versus interdependent cultural values (Singelis, 1994).

At the beginning of the session, children were introduced to a pair of unfamiliar adult females on the computer screen ("Daisy" and "Iris" for the U.S. participants, "Ms. Wang" and "Ms. Zhang" for the Chinese participants). The characters were generated in an animation website (Animaker, see Figure 4). They wore the same T-shirt in different colors (blue and yellow) and were presented side-by-side on the screen. Two pairs of female native speakers of Chinese and English completed prerecorded voice-overs for the two agents. After introducing the agents, in the *Moral Knowledge* condition, the experimenter said: "You will see that Daisy and Iris are talking about *whether it is good or bad to do something*." In the *Empirical Knowledge* condition, the experimenter said: "You will see that Daisy and Iris are talking about *what's inside these containers*." In both conditions, this statement was followed by: "…and they have different reasons, or

different ways of thinking. I want you to figure out who has better reasons, or who is thinking in a better way, okay?"

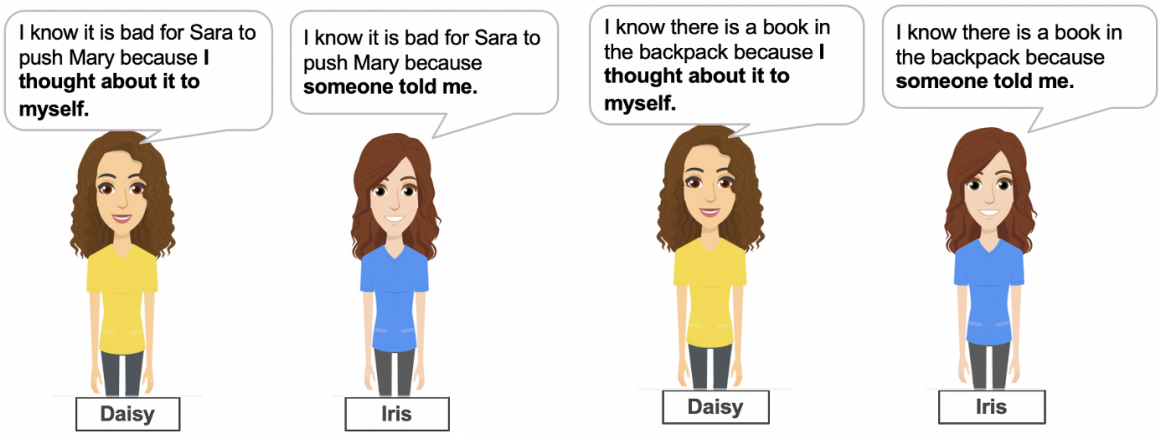


*Figure 4.* Example of Speaker Statements Used in Study 2 in the United States. In Both Conditions, One Speaker Shows Independence in her Thinking, and the Other Relies on Testimony to Make Judgments.

***Reason evaluation task.*** In each of the four trials in the *reason evaluation* task, children first went through a familiarization phase in which short vignettes were presented, and the two speakers then demonstrated different ways of thinking. In the *Moral Knowledge* condition, children heard a story in each trial in which one child engaged in either a familiar prosocial or antisocial action with the aid of a picture depicting the action (e.g., "One day, Johnny shoved Brian, and Brian got upset and started crying.", see Table 5 in Appendix for a full list of stories). After hearing each story, children were invited to confirm their opinion about the wrongness or niceness of the action (e.g., "What do you think? Is it bad for Johnny to shove Brian?"). Afterwards, the experimenter said: "Okay thanks! Now let's listen to Daisy and Iris and see what they say." In each of the four trials, a video was then shown to the child in which the two

speakers always made the same correct moral judgment, but one of the speakers consistently did so after articulating her own independent deliberation, whereas the other always justified her judgment by saying that someone had told her about the action. For example, the speaker who showed independence in her thinking would say: "I know [it is bad for Johnny to shove Brian] because *I thought about it to myself*.", and the speaker who relied on testimony to make moral judgments would say: "I know [it is bad for Johnny to shove Brain] because *someone told me*."

In the *Empirical Knowledge* condition, children were presented with a story in each trial about two protagonists discussing the content of a container, and the container itself usually gave some clues about what might be inside (e.g., "One day, Johnny showed Brian a paper bag, and Brian said: "Is there a sandwich in here?"). I decided to include specific containers that had some indications of what might be inside (e.g., a paper bag, a pencil box) so that both speakers, as well as the child, would have some ground to make assumptions and evaluate the claims. Since familiar moral actions can be relatively more transparent in nature, this would also make the two conditions more directly comparable to each other. Similar to the *Moral Knowledge* condition, after telling each story, the experimenter asked children what they thought (e.g., "What do you think? Is there a sandwich in the paper bag?"), and showed a video of one speaker who came to her own conclusions (e.g., "I know [there is a sandwich in the paper bag] because *I thought about it to myself*.") and another speaker who relied on another's testimony to figure out the contents of different containers ("I know [there is a sandwich in the paper bag] because *someone told me*." ). The specific speaker who engaged in independent

thinking, the side that each speaker was on and the order in which they spoke were counterbalanced across participants.

After each pair of claims, children were asked a *reason evaluation* question to indicate the agent with the best way of thinking, for example: "So Daisy says [she knows shoving is bad/there is a sandwich in the paper bag] because she thought about it, and Iris says [she knows shoving is bad/there is a sandwich in the paper bag] because someone told her. Who do you think *has the best reason* to know that [shoving is bad/there is a sandwich in the paper bag]? Daisy or Iris?"

After children had chosen one way of thinking, the experimenter also asked a *justification* question: "Why do you think [Daisy] has a better reason than [Iris]?" If children replied with "I don't know" or remained silent, they were invited to take some time to think about it and the question was repeated one more time. Children's answers to the *justification* questions were transcribed from videotaped sessions and coded by two research assistants blind to the hypotheses of the study. Both coders were also blind to children's age and condition. To develop the coding categories, one research assistant drew a random sample of 50% of the justifications. Reliability coding was then calculated on the basis of 25% of the justifications. Interrater reliability was high for both samples (Chinese sample $\boldsymbol{\kappa} = 91.95$; U.S. sample $\boldsymbol{\kappa} = 91.48$.) Discrepancies were resolved through discussion. I coded the remaining explanations.

***Selective learning task***. To further probe whether children's evaluations of the agents extended to a testimonial learning context, I also included a *selective learning* task (adapted from Doebel & Koenig, 2013; Koenig, 2012). Specifically, to align with the type of judgments the speakers made before across the two conditions, children in both

the *Empirical Knowledge* and *Moral Knowledge* conditions were asked to choose from one of the two speakers to learn about the contents of new containers (i.e., *novel content* trials) and about novel moral actions (i.e., *novel moral actions* trials, see Table 6 in Appendix). The order in which the two types of trials were presented was counterbalanced across participants.

Before each of three *novel content* trials, participants saw an image of a generic container (i.e., a box, a paper cup, a bag) which made it difficult to tell the content from the outside. In each trial, children were first presented with an *ask* question: "Hmmm, I wonder what is in this [box]? I bet one of these people can tell us. Who would you like to ask?" After children had made a selection, the experimenter showed a video with the two speakers making conflicting claims about what was inside the container (e.g., "There is a *biff* in the bag ." "There is a *zazz* in the bag."). Children were then be asked to endorse one of the two claims made by the speakers (*endorse* question), for example: "[Daisy] said there is a *biff* in the bag. [Iris] said there is a *zazz* in the bag. What do you think, is there a *biff* or a *zazz* in the bag?"

The procedure for the three *novel moral actions* trials was similar. In each trial, children were first presented with pictures of two children standing side by side with neutral expressions. Then, the experimenter would introduce the novel action by saying, for example: "See these children here? This is Devon, and this is Casey. One day, Devon *lepped* Casey." This was followed by an *ask* question: "Hmmm, I wonder if it is good or bad to *lep* someone. I bet one of these people can tell us. Who would you like to ask?" After children had made a choice, a video would show one speaker saying the novel action was good, and the other claiming that the action was bad. In the *endorse* question

that followed, the experimenter repeated the two speakers' moral judgments, and asked for the child's judgment, for example: "Daisy said it is good to *lep*. Iris said it is bad to *lep*. What do you think, is it good or bad to *lep*? " For both types of *selective learning* trials, the order of which informant spoke first and their specific judgments were also counterbalanced across participants.

***Social preference task.*** To assess whether children's preference for one way of thinking would affect their social and affiliative judgments toward the two agents, I also included a *social preference* task. In this task, children were asked to make choices about which of the two agents they liked more and would like to play with more. They were also asked who they would prefer to share a cookie with, and who they would like to help move some boxes.

***Control questions.*** To assess whether participants had an overwhelming preference for one speaker over the other no matter what the issue was, I included four control questions which were irrelevant to the task (e.g., asking children which of the two speakers lived in the house with a black and not white door, and who rode bus to school and not a bike). The pictures used in this task were adjusted for U.S. and Chinese participants respectively to so that the type of objects that they were most familiar with were presented.

***Explicit judgment task.*** At the end of the study, I also asked children to make a more general judgment on the person who consistently had a better way of thinking: "These people told us a lot about [whether it was good or bad to do something/what was inside these different things]. Who had the best reasons to know [whether it is good or bad to do something/what was inside these things], Daisy or Iris?" Children were also

asked to explicitly evaluate the two agents' epistemic competence by answering questions on who was smarter and who knew more.

***Parent measures.*** To explore how individual differences in parents' self-reported values might affect children's tendency to privilege independent thinking, one of the parents completed the same parent authoritarianism questionnaire as in Study 1 (Feldman & Stenner, 2005), as well as the Self-Construal Scale (SCS, Singelis, 1994) assessing cultural values related to parental independent and interdependent self-construal. The SCS consists of 30 statements that assess cultural values related to independent (e.g., "I act the same way no matter who I am with.") and interdependent (e.g., "Even when I strongly disagree with group members, I avoid an argument.") self-values on a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree). Following the SCS scoring system (Singelis, 1994), mean scores were computed for each dimension by dividing each dimension's total score by the number of items.

**Results**

Since all outcome variables were coded as binary (i.e., choosing the speaker who relied on testimony = 1, choosing the speaker who engaged in independent thinking = 0) and that each task involved multiple trials, I analyzed participants' responses using Generalized Linear Mixed- Effect Models (GLMMs) with binomial error distribution and logistic link function. All models were fit using the *glmer* function in the *lme4* package (Bates et al., 2015) in R (version 1.2.5019; R Core Team, 2019) with Age Group (children ages 4-6, adults), Culture (U.S. vs. China), Condition (Moral knowledge vs. Empirical knowledge), and their interactions as fixed effects, and participant as a random effect. Parental independent and interdependent values, parental authoritarianism was also added sequentially to investigate the contributions of parenting values on children's

preferences. I also ran separate analyses for children to examine the effects of children's Age (continuous, centered). For all measures, I report Odd Ratios (ORs) and their confidence intervals as the most easily interpreted effect size measures for ordinal logistic regressions. ORs indicate the relative change in the odds of different outcomes occurring per unit change in a predictor.

**Reason evaluation task ("Who has the better reason to know that…?")** The primary question of interest for Study 2 is whether participants judged one of the agents to have a better way of thinking. The best model combining responses from children and adults found a three-way interaction between Age Group, Culture and Condition (OR = 32.1, 95% CI = [4.01, 257]; $p = 0.001$). Following the interaction, I ran separate mixed-effect binary regression models for children and adults, and by country.

For adults, I found a significant interaction between Culture and Condition, OR = 33, 95% CI = [5.14, 212]; $p = .0002$, indicating that compared to Chinese participants, the difference between adults' responses in the Moral and Empirical conditions was larger for the U.S. participants. Comparisons against chance with Benjamini-Hochberg adjustment revealed that adults from both countries chose reliance on testimony at significantly lower than chance levels in the Morality condition (U.S. adults 7.3%. $p < .001$; Chinese adults 24.4%. $p < .001$). By contrast, in the Empirical condition, U.S. adults chose reliance on testimony at significantly higher than chance levels (74.4%, $p < .00\ 1$); but Chinese adults were at chance (57.5%, *ns*).

When examining children's responses, I found that children in China performed below chance in the proportion of trials they chose the speaker who relied on testimony as having a better way of thinking: Empirical condition, 38.3% choices of testimony as

the better reason, $p = .003$[1]; Moral condition, 33.7% choices of testimony as the better reason, $p < .001$. By contrast, children in the U.S. were at chance in choosing testimony as the better reason (Empirical condition 49.6%, *ns*; Moral condition 45.3%, *ns,* see Figure 5). A mixed-effects logistic regression including children's Age, Culture, and Condition as fixed effects, and participant as a random effect, confirmed that there was a significant main effect of Culture on children's likelihood of choosing testimony, OR = 2.39, 95% CI = [1.34, 4.29]; $p = .003$. Thus, American children were more likely than Chinese children to choose reliance on testimony as a better reason to know something over independent thinking, regardless of the type of knowledge. There was also a moderate main effect of Age, OR = 0.742, 95% CI = [0.554, 0.993]; $p = .04$, indicating that in general, older children were less likely to choose testimony as a better reason compared to younger children. I did not find a main effect of Condition, $p = .25$.

When running separate mixed-effect binary regression models for Chinese and U.S. participants, I found a significant Age Group × Condition interaction for the Chinese sample (OR = 0.12, 95% CI = [0.03, 0.53]; $p = .005$). Interestingly, I did not find any significant differences between Chinese children and adults' responses in the Moral condition (OR = 2.58, 95% CI = [0.65, 10.2]; $p = .18$), indicating that Chinese children in this condition performed at adult levels. But in the Empirical condition, Chinese adults had significantly higher likelihood of choosing testimony over independent thinking than Chinese children (OR = 0.3, 95% CI = [0.13, 0.7]; $p = .005$). By contrast, U.S. American children's responses significantly differed from adults in both Moral (OR = 33, 95% CI =

[1] All the comparisons to chance were exact binomial tests, the p-values were Benjamini-Hochberg adjusted.

[9.37, 116]; $p < .001$) and Empirical conditions (OR = 0.17, 95% CI = [0.07, 0.46]; $p = .0004$).

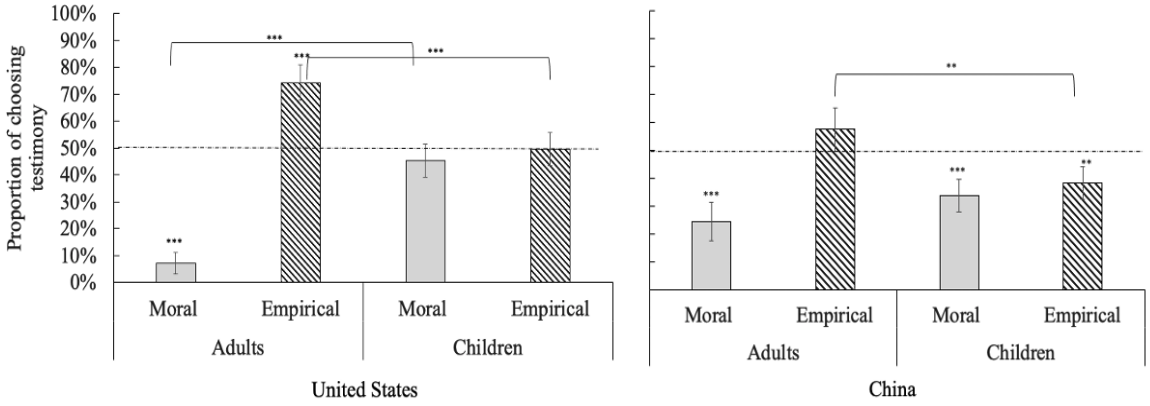


*Figure 5.* Proportion of Trials in which Participants Chose Testimony as a Better Reason for Each Country, as a Function of Age Group and Condition in Study 2.

**Reason Justifications.** One exploratory measure is to assess the types of justifications participants gave for preferring one type of reason over the other. Here, I divided participants' responses into three levels. Level 1 responses were ones in which the participants restated what the agent said without further elaboration (e.g., "Because someone told her."). Level 2 responses were ones in which participants explicitly stated that they had preference or non-preference for one way of thinking over the other, but did not further elaborate on their choice (e.g., "It's better that Daisy knows this is bad already."). Level 3 responses were ones in which participants provided specific reasons or explanations for their choices. Specifically, these justifications were coded into four categories for preferring independent thinking: (1) Speaker knowledge, intelligence or virtue; (2) common ground knowledge or inference; (3) independent agency; and (4) potential errors of testimony; and three categories for preferring reliance on testimony:

(1) Credible informant; (2) importance of learning and (3) potential errors of independent thinking (see definitions and examples of the categories in Table 5).

Table 5. *Participants' Justification Categories for Choices of Reasons in Study 2.*

| **Levels** | **Level/Category Name** | **Description** | **Example** |
|---|---|---|---|
| **Level 1** | Restatement | Repeating what the agents said, restating their choices or making irrelevant comments | "Because someone told her"; "Because she thought about it" |
| **Level 2** | Explicit preference or non-preference | Explicit preference or non-preference for one way of thinking over the other | *Child:* "Because it is usually better to just think of things."; *Adult:* "It's better that Daisy knows this is bad already." |
| **Level 3** Providing reasons for preferring the independent speaker | Speaker knowledge, intelligence or virtue | Stating that the speaker already knew the information, how one agent was smarter than the other, or commenting on the agents' moral characters and virtues. | *Child:* "Because she has a better brain"; *Adult*: "Iris has come to her decision on her own, which says something about Iris's own moral character." |
| | Common ground knowledge or inference | Stating that a concept should be known or understood by everyone, or that a fact should be easily inferred. | *Child*: "You're supposed to know kindness"; *Adult*:"This should be an intuitive reaction, there must be books in backpacks, people should assume so." |
| | Independent agency | Stating that independent epistemic or moral agency was needed, or that merely relying on others was not enough. | *Child*: "Although Ms. Wang might not be right, she is very right to have used her own brain. Ms. Zhang always has to be told by others, and she doesn't have any right intuitions about things, and needs others to tell her everything, that isn't very right." *Adult*: "I don't think it's always best to base judgements just on what you've been told, and personal reflection is important in questions about morality and ethics." |

| | | | |
|---|---|---|---|
| | Potential errors of testimony | Arguing that relying on testimony was prone to errors, for example when the addressee had encountered unreliable informants. | *Child*: “Because I don’t really know if the person who told Daisy is telling the truth or not”; *Adult*: “Because she ‘thought about it’ might mean that she reasoned her way to that conclusion using premises she didn't tell us whereas ‘someone told me about it’ is uninvestigated secondhand information and also untrue in this case.” |
| **Level 3** Providing reasons for preferring the deferring speaker | Credible informant | Stating that the informant who told the deferring agent had knowledge, was more credible or authoritative. | *Child*: “Because Iris just thought about it, Daisy actually got told from the person who owned it” *Adult*: “Daisy’s beliefs are most likely attributed to an authority figure telling her that helping is good whereas Iris may not really know why it is good despite thinking about it.” |
| | Importance of learning | Commenting on how certain facts, norms or values needed to be learned or socialized. | *Child*: “If you don’t know something, you need to learn.” *Adult*: “I chose Daisy because I think of how with younger kids, they are usually told by their parents or teachers that sharing is good. A lot of the times we don't like to share, so having someone tell us it is good, helps us understand it more.” |
| | Potential errors of independent thinking | Arguing that independent thinking was prone to errors, for example, arguing that the independent agent was merely guessing. | *Child*: “if you think about it yourself, you could be wrong. If someone told you, you could probably be like 'I agree'”; *Adult*: “Iris can think about it all she wants, unless she has the power to investigate jars with her mind then she's really just guessing." |

For child participants, around 60% of the responses across both countries fell under Level 1, and 8-12% of the responses fell under Level 2 (see Table 6). For the children who provided Level 3 responses, it is noticeable that *speaker knowledge, intelligence or virtue* was the most common category (8%-12%). It is also interesting that 12% of the U.S. children used *common ground knowledge or inference* justifications in the Moral Condition. For adults, the most frequently used category was *independent agency* in the Moral condition (25% for Chinese adults and 42% for U.S. adults), and *credible informant* in the Empirical condition (28% for Chinese adults and 32% for U.S. adults), indicating that adults may have shared the intuition that independent agency is particularly important when it comes to justifying moral beliefs.

Table 6. *Participants' Justifications for Choice of Reasons Across Countries and Conditions in Study 2.*

| | | Chinese children | | U.S. children | | Chinese adults | | U.S. adults | |
|---|---|---|---|---|---|---|---|---|---|
| | | Moral Condition | Empirical Condition | Moral Condition | Empirical Condition | Moral Condition | Empirical Condition | Moral Condition | Empirical Condition |
| Level 1 responses | | 62.08% | 58.37% | 57.01% | 61.76% | 19.38% | 18.97% | 10.30% | 19.79% |
| Level 2 responses | | 14.17% | 11.43% | 10.41% | 7.56% | 11.25% | 8.05% | 11.52% | 11.46% |
| Level 3 responses (preferring independent thinking) | Speaker knowledge, intelligence, or virtue | 7.92% | 8.16% | 12.22% | 10.50% | 11.25% | 0.57% | 29.09% | 9.38% |
| | Common ground knowledge or inference | 4.58% | 4.90% | 12.22% | 1.68% | 10.00% | 25.86% | 0.00% | 8.85% |
| | Independent agency | 3.75% | 1.63% | 1.36% | 5.04% | 25.00% | 4.02% | 41.82% | 0.52% |
| | Potential errors of testimony | 2.08% | 2.45% | 1.36% | 3.78% | 2.50% | 1.15% | 0.61% | 7.81% |
| Level 3 responses (preferring testimony) | Credible informant | 2.50% | 4.90% | 3.62% | 5.88% | 5.63% | 27.59% | 1.21% | 32.29% |
| | Importance of learning | 0.00% | 1.63% | 0.45% | 0.00% | 13.75% | 0.57% | 5.45% | 0.00% |
| | Potential errors of independent thinking | 2.92% | 6.53% | 1.36% | 3.78% | 1.25% | 13.22% | 0.00% | 9.90% |

**Selective learning task**. To explore whether participants demonstrated selectivity in their learning from the two agents, I analyzed their responses in the *selective learning task*. Preliminary analyses revealed no significant effects of question type (*ask* vs. *endorse* questions) or the orders of the informants in the selective trust task, so I collapsed *ask* and *endorse* responses for all subsequent analyses (as in, e.g., Doebel & Koenig, 2013; Elashi, & Mills, 2014).

For the analyses with the full sample, I included the following fixed effects: Age Group (adults, children), Culture (China, U.S.), Condition (Empirical, Moral) and Trial Type (novel content trials, novel moral actions trials). This revealed significant main effects of Condition (OR = 3.9, 95% CI = [2.43, 6.27], $p < .001$), Age Group (OR = 2.24, 95% CI = [1.46, 3.43], $p = .0002$) and Country (OR = 1.49, 95% CI = [1.11, 1.99], $p = .007$), and a significant interaction between Age Group (Adults, Children) and Condition (Empirical, Moral), OR = 3.57, 95% CI = [1.96, 6.5], $p < .007$. To follow up on the interaction, I conducted separate analyses for adults and children.

For the adult participants, the model revealed a significant main effect of Condition, OR = 4.49, 95% CI = [2.5, 8.04], $p < .001$ and a main effect of Trial Type, OR = 1.52, 95% CI = [1.22, 1.89], $p = .0002$, indicating that adults were more likely to selectively learn from the agent who relied on testimony in the Empirical condition than the Moral condition, and that across both conditions, they were also more likely to rely on this agent in the novel content trials ($M = 49.2\%$, $SE = 1.6\%$) compared to the novel moral actions trials ($M = 42.5\%$, $SE = 1.6\%$). I also found a main effect of Culture, OR = 2.07, 95% CI = [1.16, 3.69], $p = .014$. Interestingly, U.S. adults ($M = 51.2\%$, $SE = 1.6\%$) were more likely to learn from the agent who relied on testimony than Chinese adults ($M = 40.2\%$, $SE = 1.6\%$). None of the interactions were significant. Comparisons against chance revealed that adults from both countries selectively learned from the agent who relied on testimony at significantly lower than chance levels in the Morality condition (U.S. adults 40%. $p < .001$; Chinese adults 28%. $p < .001$). In the Empirical condition, U.S. adults chose reliance on testimony at significantly higher than chance levels (61.8%, $p < 0.001$); but Chinese adults were at chance (52.1%, *ns*).

For children, a generalized linear mixed-effect model with Age (continuous), Culture, Condition and Trial Type revealed a significant main effect of Age on children's selective learning choices, OR = 1.25, 95% CI = [1.06, 1.47], $p$ = .007. Interestingly, with age, children were more likely to learn from the speaker who relied on testimony, regardless of Condition. There were no other main effects, and adding interaction terms did not significantly impact model fit. However, when comparing children's selective learning responses to chance, I found that children in China were below chance in the proportion of trials they chose to learn from the speaker who relied on testimony in the Moral condition ($M$ = 42.9%, $SE$ = 1.7%, $p$ = .0003), but they were at chance in the Empirical condition ($M$ = 49.2%, $SE$ = 1.8%, *ns*). In line with their responses in the reason evaluation task, children in the U.S. were at chance both conditions (Empirical condition 48.2%, *ns*; Moral condition 52%, *ns*), indicating that they were equally likely to ask for and endorse information from the two speakers.

**Social preference task.** To examine participants' social preferences between the two agents, I pooled children's and adults' responses to the liking, play, helping and sharing questions. The main GLMM with Age Group, Culture, and Condition as fixed effects, and participant ID as the random effect revealed a significant Age Group · Condition interaction, OR = 23.3, 95% CI = [5.56, 97.5], $p < 0.01$. For adults, I found a significant main effect of Condition, OR = 36, 95% CI = [6.83, 190], $p < 0.01$, indicating that adults from both countries had higher social preference for the agent relying on testimony in the Empirical condition ($M$ = 49.1%, $SE$ = 2.7%) than in the Moral condition ($M$ = 20.6%, $SE$ = 2.3%). U.S. and Chinese adults were below chance in preferring the adult who relied on testimony in the Moral condition (U.S.: 16.5%, $p <$

0.001; China: 25%, $p < 0.001$). As for the Empirical condition, U.S. adults were at chance, whereas Chinese adults were moderately below chance in choosing the agent who deferred to testimony (U.S.: 56.4%, *ns* ; China: 41.2 %, $p = 0.04$ ).

By contrast, for the child participants, there were no significant main effects of Age ($p = 0.92$), Country ($p = 0.08$) or Condition ($p = 0.62$). None of the interaction terms improved model fit. U.S. children's responses were at chance regardless of Condition (Empirical: 51.4%, *ns*, Moral: 51.2%, *ns*), indicating that they did not have a clear social preference for one agent over the other based on their different ways of thinking. However, Chinese children were less likely to choose the deferring agent in the Empirical condition (Empirical: 41.7/%, $p = 0.03$, Moral: 44.6%, $p = 0.17$).

**Control questions.** These questions tested for the possibility of an unexplained general preference in favor of one character over the other. There was no effect of Age Group, Condition, but an unexpected, moderate main effect of Culture was found, OR = 1.28, 95% CI = [1.01, 1.61], $p = 0.04$. Importantly, adults and children in both conditions responded at chance level (Chinese participants: Empirical: 47.9/%, Moral: 44.9%; U.S. participants: Empirical: 51.1%, Moral: 53%).

**Explicit judgment task.** Here, I pooled participants' responses to the explicit judgment questions (i.e., who has a better reason, who is smarter, who knows more). The best model revealed a significant Age Group × Condition interaction, OR = 19.1, 95% CI = [5.64, 64.8], $p < .001$, and a moderate Country × Condition interaction, OR = 0.3, 95% CI = [0.09, 0.93]; $p = .04$. For adults, I found a significant main effect of Condition, OR = 36.9, 95% CI = [9.2, 147], $p < .001$, indicating that adults from both countries judged the agent relying on testimony to be more knowledgeable and smarter in the Empirical

condition (*M* = 55.8%, *SE* = 3.2%) than the Moral condition (*M* =18.3%, *SE* = 2.5%). Specifically, comparisons against chance revealed that adults from both countries chose reliance on testimony at significantly lower than chance levels in the Moral condition (U.S. adults 7.3%. $p < .0001$; Chinese adults 29.9%. $p < .0001$). By contrast, in the Empirical condition, U.S. adults chose reliance on testimony at significantly higher than chance levels (65.9%, $p = 0.003$); but Chinese adults were at chance (45%, *ns*).

For child participants, the mixed-effects logistic regression model revealed a significant main effect of Country on children's likelihood of choosing the speaker who relied on testimony, OR = 2.39, 95% CI = [1.19, 4.76]; $p = .014$, indicating that Chinese, but not U.S. American children judged the agent who engaged in independent thinking as having more epistemic competence. Specifically, similar to children's performance in the *reason evaluation* task, Chinese children chose the speaker who relied on testimony at below chance levels in Moral condition (36.4%, $p = .0006$), and their choice of this speaker was also below chance in the Empirical condition (40.9%, $p = .025$). By contrast, children in the U.S. were at chance when choosing between the two agents (Empirical condition 48.4%, *ns*; Moral condition 49.7%, *ns*).

**Authoritarianism, independent and interdependent self-construal.** Surveys for parents and adult participants were used to explore individual variation in authoritarian values and independent and interdependent self-construal. On average, U.S. adults (*M*= 1.05 out of 4, *SD*= 1.21) and U.S. parents (*M*= 0.78 out of 4, *SD*= 1.02) were lower in authoritarianism compared to Chinese adults (*M*= 1.52, *SD*= 0.92) and Chinese parents (*M*= 1.48, *SD*= 0.88). A linear regression using Group (adult participants versus parents) and Culture (China versus U.S.) as predictors and authoritarian score as the

response found a significant main effect of Culture, $B$ = -0.61, $SE(B)$ = 0.1, $p < .001$, indicating that Chinese parents and adult participants had higher authoritarian scores than U.S. participants.

Parents and adult participants also completed the Self-Construal Scale (Singelis, 1994) which consisted of two higher-order dimensions, Independence and Interdependence. As predicted, a linear regression revealed a significant main effect of Culture for the Interdependence score, $B$ = -0.49, $SE(B)$ = 0.07, $p < .001$, Chinese participants (Chinese adults: $M$= 5.22, $SD$= 0.70; Chinese parents: $M$=5.06, $SD$= 0.70) had higher average Interdependence scores than U.S. participants (U.S. adults: $M$= 4.60 , $SD$= 0.77; U.S. parents: $M$= 4.65, $SD$= 0.59). Surprisingly, Chinese participants (Chinese adults: $M$ = 5.03, $SD$ = 0.60; Chinese parents: $M$ = 5.11, $SD$ = 0.72) also had higher Independence scores than U.S. participants (U.S. adults: $M$= 4.82, $SD$= 0.85; U.S. parents: $M$= 4.72, $SD$= 0.69), $B$ = -0.32, $SE(B)$ = 0.07, $p < .001$. I also found that authoritarianism was not significantly correlated with Independence, but it was significantly and positively correlated with Interdependence, $r$ = 0.25, $p < .001$.

When Likelihood Ratio Tests were performed, I found that individually adding parental authoritarianism, parental independent and interdependent self-values did not significantly improve the fit of the models for the *reason evaluation* task, *selective learning* task, *social preference* task and *explicit judgment* task.

**Discussion**

In this study, I was interested in the ways in which U.S. and Chinese adults and children evaluate different bases of moral and empirical knowledge, namely autonomous judgments and testimony. I presented participants with two agents who made the same moral or empirical judgments, but who justified their beliefs either on the basis of

independent thinking or testimony. As predicted, I found that U.S. and Chinese adults demonstrated a domain-specific tendency to judge reliance on testimony as inferior to self-reliance in the Moral condition, but not in the Empirical condition. By contrast, children were less likely to distinguish between the two domains, and Chinese children were more likely to show a marked preference for independent thinking across the two domains compared to the U.S. American children.

The finding that adults judged reliance on testimony to be less legitimate in the moral domain than in the empirical domain is consistent with the philosophical account that moral testimony has limits that other forms of testimony do not share: While testimony is an acceptable source of empirical beliefs, it is problematic to base one's moral beliefs on testimony (e.g., *asymmetry thesis*, Hopkins, 2007). Specifically, adults credited someone with knowing a moral principle under their own power, not simply because they were told about it and deferred to an informant's judgment. (Jones 1999; Williams, 1972). It is possible that U.S. and Chinese adults judged justifying beliefs about familiar moral actions with testimony to be wrong either because it runs against the central aspects of moral agency such as moral worth and moral virtue (e.g., McGrath, 2011; Hills, 2009, 2013; Nickel, 2001), or because it threatens one's epistemic agency when acquiring moral understanding (Hills, 2009; McGrath, 2011). These possibilities were supported by adults' explanations of their choices in the *reason evaluation* task. Specifically, some adults commented on the agents' knowledge or moral character (e.g., "Iris has come to her decision on her own, which says something about Iris's own moral character."; "I think it's important for people to make up their own mind about what is right and wrong. It also ideally involves some empathy and being able to put yourself in

the position of the person who was harmed.") or the importance of independent agency and appreciating the reasoning behind a moral belief (e.g., "I don't think it's always best to base judgements just on what you've been told, and personal reflection is important in questions about morality and ethics."; "It is better to think about an action in order to understand why it is bad. If you understand why an action is bad then you can recognize a similar bad action in the future. If you only think it is bad because someone told you might not understand why it is bad."). Interestingly, compared to Chinese adults, the difference between adults' evaluations of the deferring agent in the Moral condition and Empirical condition was larger for U.S. adults. In particular, U.S. adults were more likely than Chinese adults to choose the agent who deferred to testimony in the Empirical condition ($p = 0.02$), but less likely than Chinese adults to choose this agent in the Moral condition ($p = 0.01$). One possible explanation is that the U.S. American adults might have had a firmer sense about what was considered right when choosing between conflicting stances, whereas Chinese adults may be more likely to have sought a middle way between two opposing views (e.g., Peng, & Nisbett, 1999). Specifically, it is possible that Chinese adults were relatively more open to testimony as a source of knowledge in the moral domain, likely because of the cultural emphasis on intellectual humility and the importance of learning, whereas U.S. adults may have felt stronger about independent moral and epistemic agency when it comes to making moral judgments because of cultural values of individualism (Markus & Kitayama, 1991; Wu et al., 2002).

This study was also the first to directly assess how children judge the legitimacy of independent thinking and reliance on testimonial input as sources of moral and

empirical knowledge, and to examine developmental and cultural differences in children's preferences. In a way, the fact that U.S. American children, but not adults, treated both ways of thinking as legitimate makes a case for the importance of testimony for young learners, indicating that skepticism towards moral testimony and appreciation for independent thinking for adults is likely a learned phenomenon acquired from cultural socialization. Interestingly, I found that Chinese children treated reliance on testimony as less ideal than coming to one's own conclusions across both the Moral and Empirical domain. This pattern of results also extended to children's responses in the *explicit judgment* task, with Chinese children judging the agent who engaged in independent thinking as more epistemically competent, and U.S. American children choosing between the two agents at chance levels. Why did Chinese children show a more marked preference for independent thinking than U.S. children? Since my study represents the first attempt to assess children's own evaluations of independently acquired knowledge and testimonial-based knowledge across cultures, what was driving this interesting cross-cultural difference still remains an open question. Here, I offer some – albeit speculative – explanations. It is common for cross-cultural studies to interpret cultural differences in terms of the collectivist versus individualist dimensions (Triandis, 1993), and previous work has found that parental socialization in the U.S. is more likely to encourage independent decision making, whereas Chinese parents are more likely to value conformity as a goal (e.g., Suizzo & Cheng, 2007). While it seemed counterintuitive for Chinese children to value social information less than U.S. children, one possibility is that the specific informant who provided testimony for the deferring agent is left unclear, and children across the two cultures might have had different interpretations on the

identity and reliability of the unspecified informant. Perhaps Chinese children were only socialized to prioritize learning from adult or high-status authorities. In fact, there is cross-cultural evidence showing that Chinese children were more likely than Spanish children to endorse an adult consensus's counterintuitive moral judgment about the exclusion of a peer (Enesco et al., 2016), but they were less likely to follow the same testimony given by a majority of peers compared to Spanish children (Sebastián-Enesco, Guerrero, & Enesco, 2020). By implication, it is possible that while Western children are more open to accepting a variety of sources acquiring second-hand knowledge, Chinese children are more selective among informants and more vigilant about their credibility. In the Study 2 design, I intentionally used a generic statement (i.e., "because someone told me") as a first step to explore children's own perspectives when directly comparing the two sources of knowledge. If adult testimony is interpreted as more authoritative than that of peers by Chinese children, future work is needed to explore whether Chinese children's evaluations of the two types of reasoning would be shifted if the deferring agent specified that the testimony was given by an adult authority (e.g., "Because my teacher told me.") or endorsed by a collective authority (e.g., "Because that's what everyone else thinks.").

Another possible explanation for the cross-cultural difference is related to the emphasis placed on knowledge in parental socialization in the Chinese culture. Influenced by the Confucianism tradition where knowledge is considered to be the first step toward perfection (Legge, trans.1971), Chinese parents have been found to emphasize diligence and persistence in the acquisition of knowledge, and treat the pursuit of knowledge as a moral virtue (Li, 2005; Luo et al, 2013). Thus, while both ways of

justifying moral and empirical knowledge might be more or less legitimate to Chinese children, they might be more sensitive to differences in the agents' epistemic competence than U.S. children, leading them to have a stronger preference for the agent who demonstrated first-personal authority in her knowledge. This possibility is supported by cross-cultural work on theory of mind, revealing that Chinese preschoolers may understand that people can be knowledgeable versus ignorant before understanding how they might differ in their beliefs (i.e., passing the knowledge access task before the diverse belief task), whereas U.S. American children understand that individuals can differ in what they think or believe earlier than they understand knowledge (i.e., passing the diverse belief task before passing the knowledge access task, Wellman et al, 2006). The finding that children who preferred independent thinking often referred to the agents' knowledge and intelligence (e.g., "She has a better brain." "She already knows.") also lends some support to this claim.

Although Chinese children demonstrated a strong preference for coming to one's own conclusions when it comes to justifying moral and empirical beliefs, their responses to the *selective learning* questions did not reveal the same level of selectivity for the agent who engaged in independent thinking. By contrast, adults from both countries were less likely to learn from the agent who relied on testimony in the Moral condition than the Empirical condition. These findings warrant further exploration. Consistent with my prior discussion, although Chinese, but not U.S. children treated the person who engaged in independent thinking as having more epistemic competence, children from both countries may have treated both independent deliberation and reliance on testimony as legitimate ways of acquiring knowledge. Thus, children might have thought that both adult agents

were acceptable to learn novel information from. This possibility is consistent with recent work suggesting that ignorance by itself may not lead children to doubt speakers when they do claim to know something (Kushnir, & Koenig, 2017). Specifically, when it comes to learning new information, preschool-aged American children did not distinguish between a speaker who had been previously accurate at labeling common objects and one who had professed ignorance (Kushnir, & Koenig, 2017), indicating that children may not discredit ignorant agents who need to rely on others for basic moral knowledge and empirical facts when it comes to learning new information. Therefore, children in the current study may have treated the deferring speaker's ignorance as a situational constraint, a constraint that did not extend to new learning contexts. An additional and not mutually exclusive possibility is that children were more "ends-driven" than "means-driven", meaning that their overall impression of the agents were less focused on their specific ways of thinking, and were instead more focused on the same correct conclusions reached by them (e.g., that shoving others is wrong and that there is an eraser in the pencil box). As a result, children might have treated both agents as adequate informants to learn new moral or epistemic information from.

Similar to the patterns of results in the *selective learning* task, while adults demonstrated a similar domain-specific aversion to the agent who relied on testimony in the Moral condition in the *social preference* task, Chinese and U.S. children did not have a clear social preference for one agent over the other. By implication, recognizing that one agent had better reasons to justify her moral and empirical beliefs did not affect Chinese children's decisions to help, share and affiliate with both speakers. Based on work revealing that children are less likely to affiliate with individuals with a history of

antisocial or immoral behavior (e.g., Hetherington et al., 2014), it is possible that children's equivalent treatment of the two agents in Study 2 signals that they did not judge the deferring agent as seriously lacking the moral virtue of knowing a moral truth under one's own power (Hopkins, 2007). To further explore whether children may have an emerging understanding that acting virtuously requires responsible agency, future studies should more explicitly ask children to evaluate the two agents' moral characters (e.g., who is nicer).

Across these tasks, it is also interesting that children's responses are domain-general, meaning that unlike adults, there were not any marked differences between the Moral condition and Empirical condition. One possible explanation is that unlike adults, children have treated the familiar moral beliefs similarly as factual beliefs (Heiphetz, & Young, 2017; Wainryb, Shaw, Langley, Cottam, & Lewis, 2004), and it is not until later in development that children develop a more nuanced understanding of the unique characteristics of the moral domain, and possible limitations of blind deference to moral testimony. Since older children in my study were more likely to endorse independent thinking than younger children, it is important for future studies to replicate the study in older age groups, and see if children would respond in more adult-like ways starting from middle childhood, and treat reliance on testimony to be particularly problematic in the moral domain.

Children's moral development is contingent upon both independent construction and getting testimonial input. How do children balance these two sources of knowledge? Study 2 constitutes one of the earliest attempts to assess the way in which children evaluate different bases of moral and empirical knowledge. I found that Chinese and

U.S.adults judged that it was less legitimate to rely on testimony to justify moral beliefs compared to empirical beliefs, and they demonstrated higher learning and social preferences for the independent agent over the deferring agent in the Moral condition. By comparison, U.S. American children were less selective in both conditions, implying that they may have considered independent deliberation and reliance on testimony as equally good ways of thinking. Chinese children, however, showed a stronger preference for independent thinking, and considered the independent agent to be more competent. These findings add an important twist to prior cross-cultural research on children's reliance on testimony. The fact that Chinese children treated reliance on testimony as epistemically inferior to independently acquiring good judgment may indicate an earlier appreciation for independent epistemic and moral agency. This work also has important implications for children's developing ability to think critically about the different ways to acquire moral knowledge, and opens interesting new avenues of research on children's development of moral agency and moral learning.

# CHAPTER 4

## General Discussion

Across two studies, I examined the role of testimony in the early ontogeny of morality. Specifically, I assessed (1) the types of moral explanations that most influence children to believe what they are told; (2) the ways in which children evaluate two bases of moral knowledge (e.g., autonomous judgments versus testimony); and (3) the cultural influences on children's judgments about reliance on testimony. With regards to the first aim, in Study 1, 3- to 5-year-old Chinese and U.S. American children were presented

with countervailing Authoritarian and Utilitarian explanations from an adult. Results revealed that both types of testimony moved children's moral judgments of the permissibility of novel action, and children who accepted these explanations were also likely to retain these judgments in the *true belief* task. Moreover, I found that while both types of explanations were similarly effective for younger children, 5-year-old children across the two countries were more influenced by Utilitarian explanations which reasoned about positive consequences for the transgressor, and were less likely to be convinced by Authoritarian explanations in which they were asked to believe on an adult's authority. Adults' judgments were not moved by either type of explanation. These results suggest that adult testimony is powerful in changing children's moral judgments, and it is possible that children treat all types of moral testimony indiscriminately at first, and become more discerning with age.

To explore the second aim, participants in Study 2 were asked to judge deference to testimony and coming to one's own conclusions in the moral domain and the empirical domain. I found that while adults judged reliance on testimony to be less legitimate when justifying moral beliefs compared to empirical beliefs, children did not distinguish between the two domains. Specifically, Chinese children showed skepticism towards reliance on testimony when evaluating the two ways of thinking and making explicit judgments about the agents; but U.S. children did not privilege independent thinking, and showed the same level of preference for independent thinking and testimony across both domains.

Finally, with regards to the third goal, I found an interesting pattern of cross-cultural differences: As addresses of information in Study 1, 4- and 5-year-old Chinese

children were more receptive to Authoritarian explanations then U.S. children; but when it comes to evaluating the use of testimony from a third-party perspective in Study 2, 4- and 6-year-old Chinese children were more likely to judge reliance on testimony to be inferior than independent deliberation compared to their U.S. counterparts. As discussed, Chinese children may value authority more in their learning, leading them to be more receptive to Authoritarian explanations in Study 1 on the one hand, and be more wary of the informant's credentials in Study 2 on the other hand. It is also possible that Study 1 presented children with new and ambiguous harm-related moral situations, whereas Study 2 presented children with familiar moral beliefs (e.g., pushing, sharing). When it comes to learning new information, perhaps Chinese children were open to deferring to an adult's judgments and take their words. But when justifying familiar moral beliefs, Chinese children were more vigilant when it comes to evaluating agents who did not know how to judge these basic moral actions.

Taken together, this dissertation project suggests that while adults appreciate the distinct features of moral testimony, children accept moral testimony at high rates, even when others' claims run counter to their own moral intuitions. It is possible that children may gradually appreciate the first-personal authority of moral testimony over the course of development, and this appreciation is influenced by cultural socialization. Moral testimony presents a particularly fascinating case because of the theoretical and practical implications of depending on others' claims in the realm of morality. While testimony can be essential in the cultivation of children's capacities for moral understanding and promote children's prosocial development, it can also lead children to think or act in a self-interested, biased, ignorant or antisocial manner. Therefore, research focused on the

role of testimony in children's moral development can have important implications for educational and parenting practices meant to support children's moral learning.

## References

Aguiar, N. R., Stoess, C. J., & Taylor, M. (2012). The development of children's ability to fill the gaps in their knowledge by consulting experts. *Child Development*, *83*, 1368-1381.

Andow, J. (2019). Why don't we trust moral testimony? *Mind & Language, 35,* 456-474.

Anonymous. (1971a). The great learning. In J. , Trans. Legge (Ed.), *The Confucian analects, the great learning & the doctrine of the mean*. New York, NY: Dover Publications (Original work published in 1893).

Bandura, A. (1986). Social foundations of thought and action. *Englewood Cliffs, NJ*, *1986*, 23-28.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... & Grothendieck, G. (2011). Package 'lme4'. *Linear mixed-effects models using S4 classes. R package version*, *1*(6).

Bauer, D. J., & Sterba, S. K. (2011). Fitting multilevel models with ordinal outcomes: performance of alternative specifications and methods of estimation. *Psychological methods*, *16*, 373.

Beijing Statistics Bureau (2021). *Beijing statistics yearbook (in Chinese).* Beijing, China: Chinese Statistics Press (Retrieved from http://tjj.beijing.gov.cn/tjsj_31433/yjdsj_31440/jmsz_32036/2021/202107/t20210719_2439211.html)

Benozio, A., & Diesendruck, G. (2017). Parochial compliance: Young children's biased consideration of authorities' preferences regarding intergroup interactions. *Child development*, *88*, 1527-1535.

Bernard, S., Castelain, T., Mercier, H., Kaufmann, L., Van der Henst, J. B., & Clément, F. (2016). The boss is always right: Preschoolers endorse the testimony of a dominant over that of a subordinate. *Journal of experimental child psychology*, *152*, 307-317.

Birch, S. A., Vauthier, S. A., & Bloom, P. (2008). Three-and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition*, *107*, 1018-1034.

Blatt, M. M., & Kohlberg, L. (1975). The effects of classroom moral discussion upon children's level of moral judgment. *Journal of Moral Education*, *4*, 129-161.

Bloom, P. (2010). How do morals change? *Nature, 464*(7288), 490.

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, *120*, 322-330.

Bond, R. & Smith, P.B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin, 119,* 111-137.

Boseovski, J. J., & Thurman, S. L. (2014). Evaluating and approaching a strange animal: Children's trust in informant testimony. *Child Development*, *85*, 824-834.

Bridgers, S., Buchsbaum, D., Seiver, E., Griffiths, T. L., & Gopnik, A. (2016). Children's causal inferences from conflicting testimony and observations. *Developmental psychology*, *52*, 9.

Bridgers, S., Gweon, H., Bretzke, M., & Ruggeri, A.(2018). How you learned matters: The process by whichothers learn informs young children's decisions aboutwhom to ask for help. In C. Kalish, M. Rau, J. Zhu, &T. Rogers (Eds.),*Proceedings of the 40th annual conferenceof the Cognitive Science Society*(pp. 1402–1407). Austin,TX: Cognitive Science Society.

Callaghan, T., Moll, H., Rakoczy, H., Warneken, F., Liszkowski, U., Behne, T., ... & Collins, W. A. (2011). Early social cognition in three cultural contexts. *Monographs of the society for research in child development*, i-142.

Chalik, L., & Rhodes, M. (2015). The communication of naïve theories of the social world in parent–child conversation. *Journal of Cognition and Development, 16,* 719-741.

Chan, K. W., & Elliott, R. G. (2004). Epistemological beliefs across cultures: Critique and analysis of beliefs structure studies. *Educational Psychology*, *24*, 123-142.

Chan, C. C., & Tardif, T. (2013). Knowing better: the role of prior knowledge and culture in trust in testimony. *Developmental Psychology*, *49*, 591.

Chang, L., Mak, M. C. K., Li, T., Wu, B. P., Chen, B. B., & Lu, H. J. (2011). Cultural adaptations to environmental variability: An evolutionary account of East-West differences. *Educational Psychology Review, 23,* 99–129.

Callanan, M. A., & Oakes, L. M. (1992). Preschoolers' questions and parents' explanations: Causal thinking in everyday activity. *Cognitive Development*, *7*, 213-233.

Castelain, T., Bernard, S., Van der Henst, J. B., & Mercier, H. (2016). The influence of power and reason on young Maya children's endorsement of testimony. *Developmental science*, *19*, 957-966.

Charafeddine, R., Mercier, H., Clément, F., Kaufmann, L., Berchtold, A., Reboul, A., & Van der Henst, J. B. (2015). How preschoolers use cues of dominance to make sense of their social environment. *Journal of Cognition and Development*, *16*, 587-607.

Chavajay, P., & Rogoff, B. (1999). Cultural variation in management of attention by children and their caregivers. *Developmental Psychology*, *35*, 1079.

Chen, X., & French, D. C. (2008). Children's social competence in cultural context. *Annual Review of Psychology*, 59, 591–616.

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

Cholbi, M. (2007). Moral expertise and the credentials problem. *Ethical Theory and Moral Practice, 10,* 323-334.

Christensen, R. H. B. (2015). ordinal—regression models for ordinal data. *R package version*, *28*, 2015.

Clegg, J. M., Cui, Y. K., Harris, P. L., & Corriveau, K. H. (2019). God, germs, and evolution: Belief in unobservable religious and scientific entities in the US and China. *Integrative Psychological and Behavioral Science*, *53*, 93-106.

Clegg, J. M., Kurkul, K. E., & Corriveau, K. H. (2019). Trust me, I'ma competent expert: developmental differences in children's use of an expert's explanation quality to infer trustworthiness. *Journal of Experimental Child Psychology*, *188*, 104670.

Clegg, J. M., Wen, N., & Legare, C. H. (2017). Is non-conformity WEIRD? Cultural variation in adults' beliefs about children's competency and conformity. *Journal of Experimental Psychology: General*, *146*, 428.

Cook, C., Goodman, N. D., & Schulz, L. E. (2011). Where science starts: Spontaneous experiments in preschoolers' exploratory play. *Cognition*, *120*, 341-349.

Correa-Chávez, M., & Rogoff, B. (2009). Children's attention to interactions directed to others: Guatemalan mayan and european american patterns. *Developmental psychology*, *45*, 630.

Corriveau, K. H., Fusaro, M., & Harris, P. L. (2009). Going with the flow: Preschoolers prefer nondissenters as informants. *Psychological science, 20,* 372-377.

Corriveau, K. H., & Harris, P. L. (2010). Preschoolers (sometimes) defer to the majority in making simple perceptual judgments. *Developmental Psychology, 46*, 437.

Corriveau, K. H., Kim, E., Song, G., & Harris, P. L. (2013). Young children's deference to a consensus varies by culture and judgment setting. *Journal of Cognition and Culture, 13,* 367-381.

Corriveau, K.H., & Kurkul, K.E. (2014). 'Why does rain fall?' Children prefer to learn from an informant who uses noncircular explanations. *Child Development, 85,* 1827–1835

Corriveau, K. H., Meints, K., & Harris, P. L. (2009). Early tracking of informant accuracy and inaccuracy. *British Journal of Developmental Psychology, 27*, 331-342.

Costin, S. E., & Jones, D. C. (1992). Friendship as a facilitator of emotional responsiveness and prosocial interventions among young children. *Developmental Psychology, 28,* 941.

Cowell, J. M., Lee, K., Malcolm-Smith, S., Selcuk, B., Zhou, X., & Decety, J. (2017). The development of generosity and moral cognition across five cultures. *Developmental science, 20,* e12403.

Crisp, R. (2014). Moral Testimony Pessimism: A Defense. In *Aristotelian Society Supplementary Volume* (Vol. 88, No. 1, pp. 129-143). Oxford, UK: Oxford University Press.

Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology*, *60*, 47-69.

Dahl, A. (2015). The developing social context of infant helping in two U.S. samples. *Child Development*, *86*, 1080– 1093.

Dahl, A. (2018). How, not whether: Contributions of others in the development of infant helping. *Current Opinion in Psychology*, *20*, 72–76.

Dahl, A., & Brownell, C. A. (2019). The social origins of human prosociality. *Current Directions in Psychological Science, 28,* 274–279.

Damon, W., and Killen, M. (1982). Peer interaction and the process of change in children's moral reasoning. *Merrill Palmer Q.* 28, 347–367.

Danovitch, J. H., & Keil, F. C. (2004). Should you ask a fisherman or a biologist?: Developmental shifts in ways of clustering knowledge. *Child Development*, *75*, 918-931.

DiYanni, C. J., Corriveau, K. H., Kurkul, K., Nasrini, J., & Nini, D. (2015). The role of consensus and culture in children's imitation of inefficient actions. *Journal of Experimental Child Psychology*, *137*, 99-110.

Doan, S. N., & Wang, Q. (2010). Maternal discussions of mental states and behaviors: Relations to emotion situation knowledge in European American and immigrant Chinese children. *Child development*, *81*, 1490-1503.

Doebel, S., & Koenig, M. A. (2013). Children's use of moral behavior in selective trust: Discrimination versus learning. *Developmental Psychology*, *49*, 462.

Driver, J. (2006). Autonomy and the asymmetry problem for moral expertise. *Philosophical Studies*, *128*, 619-644.

Einav, S. (2017). Thinking for themselves? The effect of informant independence on children's endorsement of testimony from a consensus. *Social Development*, *27*, 73-86.

Einav, S., & Robinson, E. J. (2011). When being right is not enough: Four-year-olds distinguish knowledgeable informants from merely accurate informants. *Psychological science*, *22*, 1250-1253.

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

Enesco, I., Sebastián-Enesco, C., Guerrero, S., Quan, S., & Garijo, S. (2016). What makes children defy majorities? The role of dissenters in Chinese and Spanish preschoolers' social judgments. *Frontiers in Psychology*, *7*, 1695.

Feldman, S., & Stenner, K. (1997). Perceived threat and authoritarianism. *Political Psychology*, *18*, 741-770.

Finch, W. H., Bolin, J. H., & Kelley, K. (2014). Group membership prediction when known groups consist of unknown subgroups: a Monte Carlo comparison of methods. *Frontiers in psychology*, *5*, 337.

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

Gelfert, A. (2014). *A critical introduction to testimony.* London: Bloomsbury.

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science*, *14*, 1012-1020.

Goldman, A. (2001). Experts: Which ones should you trust? *Philosophy and Phenomenological Research, 63*, 85–111.

Grusec, J. E. (2006). The development of moral behavior and conscience from a socialization perspective. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 243–265). Mahwah, NJ: Erlbaum.

Grusec, J. E., & Goodnow, J. J. (1994). Impact of parental discipline methods on the child's internalization of values: A reconceptualization of current points of view. *Developmental psychology*, *30*, 4.

Gualtieri, S., Buchsbaum, D., & Denison, S. (2020). Exploring information use in children's decision-making: Base-rate neglect and trust in testimony. *Journal of Experimental Psychology: General*, *149*, 1527.

Haidt, J., & Bjorklund, F. (2008). Social intuitionists answer six questions about moral psychology. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 2: The cognitive science of morality: Intuition and diversity* (pp. 181–217). Cambridge, MA: MIT Press.

Haidt, J., & Joseph, C. (2007). The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. *The innate mind*, *3*, 367-391.

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*, 613.

Hamlin, J. K. (2013). Moral judgment and action in preverbal infants and toddlers: Evidence for an innate moral core. *Current Directions in Psychological Science*, *22*, 186-193.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*, 557-559.

Hamlin, K. J., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science, 13* , 923–929.

Hammond, S. I., & Carpendale, J. I. M. (2015). Helping children help: The relation between maternal scaffolding and children's early help. *Social Development*, *24*, 367–383.

Harris, P. L. (2012). *Trusting what you're told: How children learn from others*. Cambridge, MA: Belknap Press/Harvard University Press.

Harris, P. L., & Koenig, M. A. (2006). Trust in testimony: How children learn about science and religion. *Child development, 77,* 505-524.

Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. *Annual Review of Psychology*, *69*, 251–273.

Haun, D. B., Rekers, Y., & Tomasello, M. (2012). Majority-biased transmission in chimpanzees and human children, but not orangutans. *Current Biology*, *22*, 727-731.

Haux, L., Engelmann, J. M., Herrmann, E., & Tomasello, M. (2017). Do young children preferentially trust gossip or firsthand observation in choosing a collaborative partner?. *Social Development*, *26*, 466-474.

Heck, I. A., Chernyak, N., & Sobel, D. M. (2018). Preschoolers' compliance with others' violations of fairness norms: The roles of intentionality and affective perspective taking. *Journal of Cognition and Development*, *19*, 568-592.

Heiphetz, L., Lane, J. D., Waytz, A., & Young, L. L. (2018). My mind, your mind, and God's mind: How children and adults conceive of different agents' moral beliefs. *British Journal of Developmental Psychology*, *36*, 467-481.

Heiphetz, L., & Young, L. L. (2017). Can only one person be right? The development of objectivism and social preferences regarding widely shared and controversial moral beliefs. *Cognition*, *167*, 78-90.

Hermansen, T. K., Ronfard, S., Harris, P. L., & Zambrana, I. M. (2021). Preschool children rarely seek empirical data that could help them complete a task when observation and testimony conflict. *Child Development*, *92*, 2546-2562.

Hetherington, C., Hendrickson, C., & Koenig, M. (2014). Reducing an in-group bias in preschool children: The impact of moral behavior. *Developmental science*, *17*, 1042-1049.

Heyes, C. (2019). Is morality a gadget? Nature, nurture and culture in moral development. *Synthese*, 1-24.

Hills, A. (2009). Moral testimony and moral epistemology. *Ethics*, *120*, 94-127.

Hills, A. (2013). Moral testimony. *Philosophy Compass*, *8*, 552-559.

Hills, A. (2019). Moral Testimony: Transmission Versus Propagation. *Philosophy and Phenomenological Research*.

Hopkins, R. (2007). What is wrong with moral testimony?. *Philosophy and Phenomenological Research*, *74*, 611-634.

Howell, R. J. (2014). Google morals, virtue, and the asymmetry of deference. *Noûs*, *48*, 389-415.

Hussar, K., & Harris, P. L. (2010). Children who choose not to eat meat: A demonstration of early moral decision-making. *Social Development, 19,* 627–641.

Jaswal, V. K. (2004). Don't believe everything you hear: Preschoolers' sensitivity to speaker intent in category induction. *Child development*, *75*, 1871-1885.

Jaswal, V. K., Croft, A. C., Setia, A. R., & Cole, C. A. (2010). Young children have a specific, highly robust bias to trust testimony. *Psychological Science*, *21*, 1541-1547.

Jaswal, V. K., & Markman, E. M. (2007). Looks aren't everything: 24-month-olds' willingness to accept unexpected labels. *Journal of Cognition and Development*, *8*, 93-111.

Jensen, L. A. (Ed.). (2015). *The Oxford handbook of human development and culture: An interdisciplinary perspective*. New York, NY: Oxford University Press.

Johnston, A. M., Mills, C. M., & Landrum, A. R. (2015). How do children weigh competence and benevolence when deciding whom to trust? *Cognition*, *144*, 76-90.

Jones, K. (1999). Second-hand moral knowledge. *The Journal of Philosophy*, *96*, 55-78.

Kajanus, A., Afshordi, N., & Warneken, F. (2020). Children's understanding of dominance and prestige in China and the UK. *Evolution and Human Behavior*, *41*, 23-34.

Kant, I. (1959). *Foundations of the metaphysics of morals*. Indianapolis, IN: Bobbs-Merrill. (Original work published 1785).

Kim, E. B., Chen, C., Smetana, J. G., & Greenberger, E. (2016). Does children's moral compass waver under social pressure? Using the conformity paradigm to test preschoolers' moral and social-conventional judgments. *Journal of Experimental Child Psychology, 150,* 241-251.

Koenig, M. A. (2012). Beyond semantic accuracy: Preschoolers evaluate a speaker's reasons. *Child Development*, *83*, 1051-1063.

Koenig, M. A., Clément, F., & Harris, P. L. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science*, *15*, 694-698.

Koenig, M. A., & Echols, C. H. (2003). Infants' understanding of false labeling events: The referential roles of words and the speakers who use them. *Cognition*, *87*, 179-208.

Koenig, M. A., & Jaswal, V. K. (2011). Characterizing children's expectations about expertise and incompetence: Halo or pitchfork effects?. *Child development*, *82*, 1634-1647.

Koenig, M. A., Tiberius, V., & Hamlin, J. K. (2019). Children's judgments of epistemic and moral agents: From situations to intentions. *Perspectives on Psychological Science*, *14*, 344-360.

Köymen, B., & Tomasello, M. (2020). The early ontogeny of reason giving. *Child Development Perspectives*, *14*, 215-220.

Kruger, A. C., & Tomasello, M. (1986). Transactive discussions with peers and adults. *Developmental Psychology, 22,* 681–685.

Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science*, *14*, 402-408.

Kurkul, K. E., & Corriveau, K. H. (2018). Question, explanation, follow-up: A mechanism for learning from others?. *Child Development*, *89*, 280-294.

Kushnir, T., & Koenig, M. A. (2017). What I don't know won't hurt you: The relation between professed ignorance and later knowledge claims. *Developmental psychology*, *53*, 826.

Lapidow, E., & Walker, C. M. (2020). Informative experimentation in intuitive science: Children select and learn from their own causal interventions. *Cognition*, *201*, 104315.

Landrum, A. R., & Mills, C. M. (2015). Developing expectations regarding the boundaries of expertise. *Cognition*, *134*, 215-231.

Lane, J. D. (2018). Children's belief in counterintuitive and counterperceptual messages. *Child Development Perspectives*, *12*, 247-252.

Lane, J. D., Conder, E. B., & Rottman, J. (2020). The influence of direct and overheard messages on children's attitudes toward novel social groups. *Child Development*, *91*, 829-845.

Lane, J. D., & Harris, P. L. (2015). The roles of intuition and informants' expertise in children's epistemic trust. *Child Development*, *86*, 919-926.

Lane, J. D., & Harris, P. L. (2014). Confronting, representing, and believing counterintuitive concepts: Navigating the natural and the supernatural. *Perspectives on Psychological Science*, *9*, 144-160.

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

Leman, P. J. (2005). Authority and moral reasons: Parenting style and children's perceptions of adult rule justifications. *International Journal of Behavioral Development*, *29*, 265-270.

Leman, P. J., & Duveen, G. (1999). Representations of authority and children's moral reasoning. *European Journal of Social Psychology*, *29*, 557-575.

Levine, S., Mikhail, J., & Leslie, A. M. (2018). Presumed innocent? How tacit assumptions of intentional structure shape moral judgment. *Journal of Experimental Psychology: General*, *147*, 1728.

Lewis, M. (2020). A defense of the very idea of moral deference pessimism. *Philosophical Studies*, *177*, 2323-2340.

Li, J. (2005). Mind or virtue: Western and Chinese beliefs about learning. *Current Directions in Psychological Science*, *14*, 190-194.

Li, J., & Fischer, K. W. (2004). Thought and affect in American and Chinese learners' beliefs about learning. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 385–418). Mahwah, NJ: Erlbaum.

Li, P. H., Harris, P. L., & Koenig, M. A. (2019). The role of testimony in children's moral decision making: Evidence from China and United States. *Developmental Psychology*, *55*, 2603.

Li, P. H., & Koenig, M. A. (2020). Children's Evaluations of Informants and their Surprising Claims in Direct and Overheard Contexts. *Journal of Cognition and Development*, 1-22.

Liu, D., Vanderbilt, K. E., & Heyman, G. D. (2013). Selective trust: Children's use of intention and outcome of past testimony. *Developmental psychology*, *49*, 439.

Lutz, D. J., & Keil, F. C. (2002). Early understanding of the division of cognitive labor. *Child Development*, *73*, 1073-1084.

Luo, R., Tamis-LeMonda, C. S., & Song, L. (2013). Chinese parents' goals and practices in early childhood. *Early Childhood Research Quarterly*, *28*, 843-857.

Ma, L., & Ganea, P. A. (2010). Dealing with conflicting information: Young children's reliance on what they see versus what they are told. *Developmental Science*, *13*, 151-160.

Mammen, M., Köymen, B., and Tomasello, M. (2019). Children's reasoning with peers and parents about moral dilemmas. *Developmental Psychology.* 55, 2324–2335.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. Psychological Review, 98, 224–253.

Matheson, J., McElreath, S., & Nobis, N. (2018). Moral experts, deference & disagreement. In J. C. Watson & L.K. Guidry-Grimes (Eds.), *Moral expertise* (pp. 87-105). Springer, Cham.

McAuliffe, K., Raihani, N. J., & Dunham, Y. (2017). Children are sensitive to norms of giving. *Cognition*, *167*, 151-159.

McGrath, S. (2009). The puzzle of pure moral deference. *Philosophical Perspectives, 23*, 321–344.

McGrath, S. (2011). Skepticism about moral expertise as a puzzle for moral realism. *The Journal of Philosophy*, *108*, 111-137.

McLoughlin, N., Finiasz, Z., Sobel, D. M., & Corriveau, K. H. (2021). Children's developing capacity to calibrate the verbal testimony of others with observed evidence when inferring causal relations. *Journal of Experimental Child Psychology*, *210*, 105183.

Mercier, H. (2011). Reasoning serves argumentation in children. *Cognitive Development*, *26*, 177-191.

Mercier, H. (2016). The argumentative theory: predictions and empirical evidence. *Trends in Cognitive Sciences, 20*, 689-700.

Mercier, H., Bernard, S., & Clément, F. (2014). Early sensitivity to arguments: how preschoolers weight circular arguments. *Journal of Experimental Child Psychology, 125,* 102–109.

Mercier, H., Castelain, T., Hamid, N., & Marín-Picado, B. (2017) The power of moral arguments. In J.-F. Bonnefon and B. Trémolière (eds.) *Moral Inferences*. London: Routledge.

Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Sciences*, *11*, 143-152.

Mills, C. M. (2013). Knowing when to doubt: developing a critical stance when learning from others. *Developmental psychology*, *49*, 404.

Mills, C. M., & Keil, F. C. (2005). The development of cynicism. *Psychological Science*, *16*, 385-390.

Misch, A., Over, H., & Carpenter, M. (2018). The whistleblower's dilemma in young children: When loyalty trumps other moral concerns. *Frontiers in psychology, 9,* 250.

National Bureau of Statistics of China. (2021). *Statistical Communiqué of the People's Republic of China on the 2012 National Economic and Social development.* Retrieved from:. http://www.stats.gov.cn/tjsj/zxfb/202201/t20220117_1826403.html#:~:text=2021%E5%B9%B4%EF%BC%8C%E5%85%A8%E5%9B%BD%E5%B1%85%E6%B0%91%E4%BA%BA%E5%9D%87%E5%8F%AF%E6%94%AF%E9%85%8D%E6%94%B6%E5%85%A5%E4%B8%AD%E4%BD%8D,%E6%98%AF%E5%B9%B3%E5%9D%87%E6%95%B0%E7%9A%8489.3%25%E3%80%82

Nickel, P. (2001). Moral testimony and its authority. *Ethical Theory and Moral Practice*, *4*, 253-266.

Noh, J. Y., Elenbaas, L. M., Park, K. J., Chung, Y. S., & Killen, M. (2017). Opinion versus knowledge: The influence of testimony format on children's judgments in morally relevant contexts. *Early Education and Development*, *28*, 240-254.

Noh, J. Y., Jambon, M., Smetana, J. G., Lee, I. J., & Killen, M. (2020). Children's Evaluations of Necessary Harm: The Role of Maternal Disapproval and Friend Relationship Status. *Merrill-Palmer Quarterly*, *66*, 145-173.

Paddy Jane, M. (2018). Moral Testimony and Moral Understanding, *Journal of Moral Philosophy, 15,* 245-271.

Paulus, M., & Moore, C. (2014). The development of recipient-dependent sharing behavior and sharing expectations in preschool children. *Developmental psychology, 50,* 914.

Paxton, J. M., & Greene, J. D. (2010). Moral reasoning: Hints and allegations. *Topics in cognitive science*, *2*, 511-527.

Peng, K., & Nisbett, R. E. (1999). Culture, dialectics, and reasoning about contradiction. *American psychologist*, *54*, 741.

Peretz-Lange, R., Perry, J., & Muentener, P. (2021). Developmental shifts toward structural explanations and interventions for social status disparities. *Cognitive Development*, *58*, 101042.

Piaget, J. (1965). The stages of the intellectual development of the child. *Educational psychology in context: Readings for future teachers*, *63*, 98-106.

Pizarro, D. A., Uhlmann, E., & Bloom, P. (2003). Causal deviance and the attribution of moral responsibility. *Journal of Experimental Social Psychology, 39,* 653–660.

Pratt, D., Kelly, M., & Wong, W. (1999). Chinese conceptions of "effective teaching" in Hong Kong: Towards culturally sensitive evaluation of teaching. *International Journal of Lifelong Education, 18,* 241–258.

R Core Team. (2019). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rhodes, M., & Chalik, L. (2013). Social categories as markers of intrinsic interpersonal obligations. *Psychological science*, *24*, 999-1006.

Rhodes, M., & Wellman, H. (2017). Moral learning as intuitive theory revision. *Cognition*, *167*, 191-200.

Roberts, S. O., Ho, A. K., & Gelman, S. A. (2021). Should individuals think like their group? A descriptive-to-prescriptive tendency toward group-based beliefs. *Child Development*, *92*, e201-e220.

Rogoff, B., Correa-Chávez, M., & Silva, K. G. (2011). Cultural variation in children's attention and learning. *Psychology and the real world: Essays illustrating fundamental contributions to society*, 154-163.

Ronfard, S., Chen, E. E., & Harris, P. L. (2021). Testing What You're Told: Young Children's Empirical Investigation of a Surprising Claim. *Journal of Cognition and Development*, *22*, 426-447.

Ronfard, S., & Lane, J. D. (2019). Children's and adults' epistemic trust in and impressions of inaccurate informants. *Journal of experimental child psychology*, *188*, 104662.

Ronfard, S., Zambrana, I. M., Hermansen, T. K., & Kelemen, D. (2018). Question-asking in childhood: A review of the literature and a framework for understanding its development. *Developmental Review*, *49*, 101-120.

Rottman, J., & Young, L. (2015). Mechanisms of moral development. In J. Decety & T. Wheatley (Eds.), *The Moral Brain: A Multidisciplinary Perspective* (pp. 123–142). Cambridge, MA: MIT Press.

Rottman, J., Zizik, V., Minard, K., Young, L., Blake, P. R., & Kelemen, D. (2020). The moral, or the story? Changing children's distributive justice preferences through social communication. *Cognition, 205,* 1-15.

Rottman, J., Young, L., & Kelemen, D. (2017). The impact of testimony on children's moralization of novel actions. *Emotion*, *17*, 1-16.

Ruggeri, A., Luan, S., Keller, M., & Gummerum, M. (2018). The influence of adult and peer role models on children'and adolescents' sharing decisions. *Child Development*, *89*, 1589-1598.

Schäfer, M., Haun, D. B., & Tomasello, M. (2015). Fair is not fair everywhere. *Psychological science*, *26*, 1252-1260.

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: preschoolers engage in more exploratory play when evidence is confounded. *Developmental psychology*, *43*, 1045.

Sebastián-Enesco, C., Guerrero, S., & Enesco, I. (2020). What makes children defy their peers? Chinese and Spanish preschoolers' decisions to trust (or not) peer consensus. *Social Development*, *29*, 494-508.

Shtulman, A. (2009). The development of possibility judgment within and across domains. *Cognitive Development*, *24*, 293-309.

Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and social psychology bulletin*, *20*, 580-591.

Sliwa, P. (2012). In defense of moral testimony. *Philosophical Studies*, *158*, 175-195.

Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness?. *Psychological Science*, *23*, 196-204.

Slomkowski, C. L., & Killen, M. (1992). Young children's conceptions of transgressions with friends and nonfriends. *International Journal of Behavioral Development, 15,* 247-258.

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

Smetana, J. G. (1985). Preschool children's conceptions of transgressions: Effects of varying moral and conventional domain-related attributes. *Developmental Psychology, 21,* 18-29.

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

Smetana, J. G., & Ball, C. L. (2019). Heterogeneity in children's developing moral judgments about different types of harm. *Developmental psychology*, *55*, 1150.

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

Sobel, D. M., & Corriveau, K. H. (2010). Children monitor individuals' expertise for word learning. *Child development*, *81*, 669-679.

Stenner, K. (2005). *The authoritarian dynamic*. Cambridge University Press.

Stephens, E. C., & Koenig, M. A. (2015). Varieties of testimony: Children's selective learning in semantic versus episodic domains. *Cognition*, *137*, 182-188.

Stephens, E., Suarez, S., & Koenig, M. (2015). Early testimonial learning: Monitoring speech acts and speakers. *Advances in child development and behavior*, *48*, 151-183.

Sterelny, K. (2010). Moral nativism: A sceptical response. *Mind & Language*, *25*, 279-297.

Suizzo, M. A., & Cheng, C. C. (2007). Taiwanese and American mothers' goals and values for their children's futures. *International Journal of Psychology, 42,* 307-316.

Tasimi, A. (2020). Connecting the dots on the origins of social knowledge. *Perspectives on Psychological Science*, *15*, 397-410.

Thomas, A. J., Thomsen, L., Lukowski, A. F., Abramyan, M., & Sarnecka, B. W. (2018). Toddlers prefer those who win but not when they win by force. *Nature human behaviour*, *2*, 662-669.

Thomsen, L., Frankenhuis, W. E., Ingold-Smith, M., & Carey, S. (2011). Big and mighty: Preverbal infants mentally represent social dominance. *science*, *331*(6016), 477-480.

Toelch, U., Bruce, M. J., Newson, L., Richerson, P. J., & Reader, S. M. (2014). Individual consistency and flexibility in human social information use. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1776), 20132864.

Tomasello, M. (2020). The moral psychology of obligation. *Behavioral and Brain Sciences*, *43*.

Tong, Y., Wang, F., & Danovitch, J. (2020). The role of epistemic and social characteristics in children's selective trust: Three meta-analyses. *Developmental Science*, *23*, e12895.

Triandis, H. C. (1993). Collectivism and individualism as cultural syndromes. *Cross-cultural research*, *27*(3-4), 155-180.

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

Vanderbilt, K. E., Liu, D., & Heyman, G. D. (2011). The development of distrust. *Child Development*, *82*, 1372-1380.

Vasilyeva, N., Gopnik, A., & Lombrozo, T. (2018). The development of structural thinking about social categories. *Developmental Psychology*, *54*, 1735.

Wainryb, C., Shaw, L. A., Langley, M., Cottam, K., & Lewis, R. (2004). Children's thinking about diversity of belief in the early school years: Judgments of relativism, tolerance, and disagreeing persons. *Child development*, *75*, 687-703.

Walker, L. J., Hennig, K. H., & Krettenauer, T. (2000). Parent and peer contexts for children's moral reasoning development. *Child Development, 71,* 1033–1048.

Wang, J., Elicker, J., McMullen, M., & Mao, S. (2008). Chinese and American preschool teachers' beliefs about early childhood curriculum. *Early Child Development and Care*, *178*, 227-249.

Wang, Q., Leichtman, M. D., & Davies, K. I. (2000). Sharing memories and telling stories: American and Chinese mothers and their 3-year-olds. *Memory*, *8*, 159-177.

Warneken, F., & Tomasello, M. (2009). Varieties of altruism in children and chimpanzees. *Trends in cognitive sciences*, *13*, 397-402.

Wellman, H. M., Fang, F., Liu, D., Zhu, L., & Liu, G. (2006). Scaling of theory-of-mind understandings in Chinese children. *Psychological science*, *17*, 1075-1081.

Wen, N. J., Clegg, J. M., & Legare, C. H. (2019). Smart conformists: Children and adolescents associate conformity with intelligence across cultures. *Child Development*. *90.* 746 – 758.

Wiland, E. (2017). Moral Testimony. *Oxford Studies in Metaethics 12*, *12*, 51.

Williams, B. A. O. (1972). *Knowledge and Reasons. Problems in the Theory of Knowledge / Problèmes de La Théorie de La Connaissance*, ed. GH Von Wright, pp. 1–11. Berlin, Germany: Springer

Woo, B. M., Tan, E., & Hamlin, K. (2022). Early-Emerging Evaluations of Morally Relevant Behavior: Breadth, Basis, and Connections to Later Development.

Woolley, J. D., & E. Ghossainy, M. (2013). Revisiting the fantasy–reality distinction: Children as naïve skeptics. *Child development*, *84*, 1496-1510.

Wu, C. C., & Honig, A. S. (2010). Taiwanese mothers' beliefs about reading aloud with preschoolers: Findings from the parent reading belief inventory. *Early Child Development and Care*, *180*, 647-669.

Wu, P., Robinson, C. C., Yang, C., Hart, C. H., Olsen, S. F., Porter, C. L., ... & Wu, X. (2002). Similarities and differences in mothers' parenting of preschoolers in China and the United States. *International Journal of Behavioral Development*, *26*, 481-491.

Yu, Y., Landrum, A. R., Bonawitz, E., & Shafto, P. (2018). Questioning supports effective transmission of knowledge and increased exploratory learning in pre-kindergarten children. *Developmental science*, *21*, e12696.

Zahn-Waxler, C., Radke-Yarrow, M., & King, R. A. (1979). Child rearing and children's prosocial initiations toward victims of distress. *Child development*, 319-330.

Zhang, X., Anderson, R. C., Dong, T., Nguyen-Jahiel, K., Li, Y., Lin, T. J., & Miller, B. (2013). Children's moral reasoning: Influence of culture and collaborative discussion. *Journal of Cognition and Culture*, *13*, 503-522.

Zhao, L., Chen, L., Sun, W., Compton, B. J., Lee, K., & Heyman, G. D. (2019). Young children are more likely to cheat after overhearing that a classmate is smart. *Developmental Science*, e12930.

Zhao, L., Heyman, G. D., Chen, L., & Lee, K. (2018). Telling young children they have a reputation for being smart promotes cheating. *Developmental science*, *21*, e12585.

# Appendix

Table 1.
*Study 1 story vignettes.*

| | Story vignette |
|---|---|
| Trial 1 | This is Devon. He (or she) is a little boy (or girl) about your age. One day at school, he (or she) saw that one of his (or her) classmates, Casey was playing with a cool new toy. Devon wanted to play with that toy too. He (or she) felt like mibbing Casey and so he (or she) did. He (or she) mibbed Casey and Casey got very upset and started crying. |
| Trial 2 | This is Taylor. He (or she) is a little boy (or girl) about your age. One day at school, he (or she) saw that one of his (or her) classmates, Morgan was playing on the swings. Taylor wanted to go play on the swings too. He (or she) felt like tamming Morgan and so he (or she) did. He (or she) tammed Morgan and Morgan got very upset and started crying. |
| Trial 3 | This is Amari. He (or she) is a little boy (or girl) about your age. One day, Amari was playing with his/her sister Brodie. When they were taking a snack break, Amari's mommy gave Brodie the last strawberry ice cream. Amari was mad. He (or she) felt like gorping Brodie and so he (or she) did. He (or she) gorpped Brodie and Brodie got very upset and started crying. |

**Table 2.**
*Study 1 Video script (C: confederate in the video; I: informant).*

| | ***Control condition*** | ***Utilitarian testimony condition*** | ***Authoritarian testimony condition*** |
|---|---|---|---|
| *Trial 1* | **C:** What do you think? Did mibbing happen in the story? | **C:** What do you think? Was it good or bad for [Devon] to [mib]? | |
| | **I:** Yes. | **I:** It was good for Devon to mib. | |
| | **C:** Who did the mibbing, Devon or Casey? | **C:** Why was it good to mib? | |
| | **I:** Devon did the mibbing. | **I:** Because Devon could play with the new toy after mibbing. | **I:** Because I get to decide whether mibbing was good. |
| *Trial 2* | **C:** What do you think? Did tamming happen in the story? | **C:** What do you think? Was it good or bad for [Taylor] to [tam]? | |
| | **I:** Yes. | **I:** It was good for Taylor to tam. | |
| | **C:** Who did the tamming, Taylor or Morgan? | **C:** Why was it good to tam? | |

| | | | |
|---|---|---|---|
| | **I:** Taylor did the tamming. | **I:** Because Morgan could play on the swings after tamming. | **I:** Because I get to decide whether tamming was good. |
| *Trial 3* | **C:** What do you think? Did gorping happen in the story? | **C:** What do you think? Was it good or bad for [Amari] to [gorp]? | |
| | **I:** Yes. | **I:** It was good for Amari to gorp. | |
| | **C:** Who did the gorping, Amari or Brodie? | **C:** Why was it good to gorp? | |
| | **I:** Amari did the gorping. | **I:** Because Amari could eat the ice cream after gorping. | **I:** Because I got to decide whether gorping is good. |

Table 3.
*Study 1 scenarios in the context transfer questions.*

| | Alternative location | Alternative victim |
|---|---|---|
| Trial 1 | Playground | Another classmate who was playing in the sand pit |
| Trial 2 | Park | Another classmate who was riding on a bike |
| Trial 3 | Birthday party | A cousin/sibling who was having some cake. |

Table 4.
*Study 1 actions in the novel word comprehension check questions.*

| | Negative actions | Neutral actions | Positive actions |
|---|---|---|---|
| Trial 1 | taking toy away from classmate | reading next to the classmate | Giving classmate a gift |
| Trial 2 | Shoving classmate off the swing | Jumping rope next to the swing | Giving classmate a nice push on the swing |
| Trial 3 | Stealing ice cream from sibling | Eating cookie next to sibling | Giving sibling his/her own ice cream |

Table 5.
*Study 2 stories in the Moral and Empirical Conditions (U.S. version) .*

| **Moral Knowledge condition** | **Empirical Knowledge condition** |
|---|---|
| One day, Sara pushed Mary at school, and Mary said “Ouch!” and told the teacher. | One day, Sara showed Mary a backpack at school, and Mary said: “Is there a book in here?” |
| One day, Johnny shoved Brian, and Brian got upset and started crying. | One day, Johnny showed Brian a paper bag, and Brian said: “Is there a sandwich in here?” |

| | |
|---|---|
| One day, Hazel shared with Shirley in class, and Shirley smiled and gave Hazel a hug. | One day, Hazel gave Shirley a pencil box in class, and Shirley asked Hazel: “Is there an eraser in here?” |
| One day, Tom helped Raymond, and Raymond told Tom: “Thank you!” | One day, Tom gave Raymond a jar, and Raymond asked Tom: “Is there jam in here? ” |

Table 6.
*Study 2 agent statements in the selective learning trials(U.S. version).*

Novel Content Trials:

| **Agent A** | **Agent B** |
|---|---|
| There is a [biff in the bag]. | There is a [zazz in the bag]. |
| There is a [crut in the box]. | There is a [larp in the box]. |
| There is [linz] in the cup. | There is [slod] in the cup. |

Novel Moral Action Trials:

| **Agent A** | **Agent B** |
|---|---|
| It is [good] for Deven to [lep] Casey . | It is [bad] for Devon to [lep] Casey. |
| It is [good] for Taylor to [dax] Morgan. | It is [bad] for Taylor to [dax] Morgan. |
| It is [good] for Bailey to [gorp] Jamie. | It is [bad] for Bailey to [gorp] Jamie. |