AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2024, Vol. 153, No. 9, 2216–2229
https://doi.org/10.1037/xge0001617

Cross-Cultural Conceptions of Third-Party Intervention Across Childhood

Julia Marshall[1], Kellen Mermin-Bunnell[2], Anton Gollwitzer[3, 4], Jan Retelsdorf[5], and Paul Bloom[6]

[1] Department of Psychology and Neuroscience, Boston College
[2] Emory University School of Medicine
[3] Department of Leadership and Organizational Behaviour, BI Norwegian Business School
[4] Center for Adaptive Rationality, Max Planck Institute for Human Development
[5] Department of Educational Psychology, University of Hamburg
[6] Department of Psychology, University of Toronto

Third-party intervention is a cornerstone of cooperative societies, yet we know little about how children develop an understanding of this social behavior. The present work generates a cross-cultural and developmental picture of how 6-, 9-, and 12-year-olds ($N = 447$) across four societies (India, Germany, Uganda, and the United States) reason about third-party intervention. To do so, we measured children's obligation judgments and unstructured descriptions of third-party intervention. Although some cultural differences emerged, 6-year-olds largely considered bystanders as obligated to respond to wrongdoing, regardless of the bystander's social position. In contrast, 9- and 12-year-olds were more likely to exclusively ascribe this social responsibility to people in positions of authority. Despite these age differences, children of all ages generated role-specific descriptions of third-party intervention, with authority figures intervening in distinct ways from peers. For authority figures, children in India and Uganda described third-party intervention as involving corporal punishment or unspecified punishment, whereas children in the United States described such intervention as involving only verbal intervention (i.e., telling someone to stop). For peers, children in all societies described third-party intervention as involving reporting misdeeds to an authority. Collectively, these data show that early conceptualizations of third-party intervention are rooted in shared notions of obligation yet are also subject to cultural and contextual influences.

Public Significance Statement

The present research examines how 6-, 9-, and 12-year-olds in different countries (India, Germany, Uganda, and United States) reason about obligations to respond to wrongdoing and how they describe third-party intervention. Six-year-olds generally considered all bystanders, regardless of social context, as obligated to respond to wrongdoing, while, for the most part, 9- and 12-year-old children were more selective when ascribing obligations. Children in Uganda were the exception: they considered all agents obligated to respond to wrongdoing regardless of age. Despite age differences in obligation judgments, all children described third-party intervention in role-dependent ways. These findings highlight the importance of considering the culture- and age-specific aspects of third-party intervention, a key cooperative behavior.

Keywords: child development, punishment, obligations, norm enforcement, cross-cultural

Supplemental materials: https://doi.org/10.1037/xge0001617.supp

Megan M. Saylor served as action editor.

Julia Marshall https://orcid.org/0000-0002-7981-0481

The work was funded by the Klaus J. Jacobs Research Prize awarded by the Jacobs Foundation. The authors thank the Mind and Development Lab for helpful feedback on the present study. The authors thank the children, parents, schools, and teachers who participated in this research. The authors would also like to thank the people who helped in the data collection efforts for this project, including Ilona Diewold, Jonathan Kirner, Noerine Kaleeba, Nicholas Sselu, Christine Kaleeba, Thadeus Kasumba, Julius Mpagi, Caroline Atuk, Doreen Twenatween, Alexa Sacchi, Maya Rodriguez, Arianna Hairston, Emily Gerdin, Karli Cecil, Saba Baldiwala, Sakina Dahodwala, and Mahesh Srinivasan. This work was presented as part of a symposium at the Society for Research in Child Development in March of 2023. The authors have no conflicts of interest to declare.

These studies were approved by the Yale Human Subjects Committee (Protocol No. 1302011578), and we obtained informed parental consent and verbal assent from all participants. All data have been made publicly available at the Open Science Framework and can be accessed at https://osf.io/6mgpy/.

Julia Marshall played a lead role in conceptualization, formal analysis, funding acquisition, and writing–original draft and an equal role in data curation. Kellen Mermin-Bunnell played a supporting role in conceptualization, project administration, and writing–review and editing and an equal role in data curation and resources. Anton Gollwitzer played a supporting role in writing–review and editing and an equal role in data curation. Jan Retelsdorf played a supporting role in writing–review and editing and an equal role in data curation and supervision. Paul Bloom played a supporting role in writing–review and editing and an equal role in conceptualization, data curation, and supervision.

Correspondence concerning this article should be addressed to Julia Marshall, Department of Psychology and Neuroscience, Boston College, 275 Beacon Street, Chestnut Hill, MA 02467, United States. Email: marshaau@bc.edu

Responding to wrongdoing is a central feature of cooperative human societies. Consider a child who bullies a classmate, a college student who cheats on a test, or a colleague who evades taxes—in all these cases, we often feel that such wrongdoing should not go unnoticed. The desire to respond to wrongdoing as a third-party observer is referred to as *third-party intervention*. Such intervention can take on many different forms. For example, one could directly chastise the child who bullies a classmate, report the child to a higher authority, comfort the victim, or some combination of these responses.

The most prominently studied third-party response to transgressions is punishment, where punishment refers to the act of imposing a cost on a transgressor (such as inducing shame or taking away privileges; Vidmar & Miller, 1980). Over the past several decades, researchers have largely focused on third-party punishment for three reasons. First, a willingness to punish third parties plays a role in maintaining cooperative societies by promoting prosocial behavior and deterring antisocial behavior (Balliet et al., 2011; van Dijk et al., 2015; Wu et al., 2022). Second, research from a variety of fields—including social psychology, behavioral economics, and anthropology—has found that adults will engage in third-party punishment even when doing so requires making personal sacrifices (Fehr & Fischbacher, 2004; Henrich et al., 2006). Finally, third-party punishment is not present in nonhuman primates (Riedl et al., 2012), suggesting that third-party punishment may be a unique feature of human psychology (Kanakogi et al., 2022). Taken together, third-party punishment is an important mechanism for upholding societal cooperation and appears to be a unique characteristic of human psychology.

That adults engage in third-party punishment and nonhuman primates do not raises a developmental question: What do children understand about third-party punishment and at what point in development does such an understanding emerge? Current research suggests that infants, toddlers, and children hold specific expectations about who is likely to and who has an obligation to pursue punishment (see Marshall & McAuliffe, 2022, for a review). For example, 9-month-olds expect a bystander to direct punishment toward an unfair distributor rather than a fair one (Geraci & Surian, 2023), and 21-month-olds also expect a bystander to direct punishment toward someone who did not defend another agent against physical harm compared to someone who did (Geraci & Surian, 2021). Beyond judgment, children are also willing to punish transgressors themselves. At ages as young as 4, children will punish a variety of transgressions, including selfishness, property destruction, and theft among many others (i.e., McAuliffe et al., 2015; Riedl et al., 2015; Yudkin et al., 2020), and this propensity similarly strengthens with increasing age across diverse societies (House et al., 2020).

Limitations of Prior Work

The work to date is limited in three ways. First, most developmental research has focused on direct third-party punishment as the primary response to wrongdoing. In reality, though, third-party intervention can take on many other forms, such as compensating victims (e.g., FeldmanHall et al., 2014), encouraging forgiveness (e.g., McCullough & Witvliet, 2002), or dissolving a relationship (e.g., Martin & Cushman, 2015). Although we know that children between ages 5 and 9 view punishing a transgressor as less positive than helping a victim (Lee et al., 2022; Lee & Warneken, 2020), we do not know whether children believe that other people will respond to wrongdoing via punishing or helping. Addressing this question can help determine whether children at these ages believe that punishment is the primary way that people respond to wrongdoing, or whether children express an interest in punishment only because it is emphasized by experimenters at the expense of alternative ways of intervention (Marshall & McAuliffe, 2022).

Second, cooperation is not maintained in a vacuum. Instead, it is usually sustained within a broader social structure characterized by specific roles that denote particular responsibilities. For example, at least in some societies, people in a position of authority (such as parents, teachers, or law enforcement) are often the ones endowed with a duty to respond to certain sorts of wrongdoing, whereas ordinary citizens are not (Binder, 2002; Cushman, 2015). From a theoretical point of view, many have argued that rules delineating who has to respond to wrongdoing help facilitate cooperative interactions, especially in large-scale societies where individuals' social responsibilities are more distributed (Baldassarri & Grossman, 2011; O'Gorman et al., 2009; Traulsen et al., 2012). It is unclear, though, whether children across different societies recognize how social structures might govern third-party intervention.

Some work does touch on this issue. For example, 17-month-olds in the United States expect leaders to rectify a wrongdoing but do not hold such expectations for nonleaders, suggesting that even infants may recognize that social context matters for third-party intervention (Stavans & Baillargeon, 2019). Whether this line of thinking extends to obligations is unclear, however. For example, it is possible that older children may not expect peers to respond to wrongdoing but nonetheless judge them as obligated to do so. In support of this possibility, 4- to 5-year-olds in the United States consider peers as similarly obligated to punish transgressors as authority figures and only begin to distinguish between authority figures' and peers' obligations to punish when they are a couple of years older (Marshall, Mermin-Bunnell, & Bloom, 2020). In summary, although social roles are very important to understanding obligations to respond to wrongdoing, we do not know whether children grasp these roles.

Third, most research investigating how children respond to wrongdoing relies on samples from Western societies (for discussions of this sampling bias, see Amir & McAuliffe, 2020; Nielsen et al., 2017). This emphasis distorts our broader understanding of third-party intervention in part because the types of intervention that bystanders pursue in response to wrongdoing are shaped by cultural norms and vary across societies (Eriksson et al., 2021; Garfield et al., 2023; Kanngiesser et al., 2022; Marlowe et al., 2008; for a review, see Molho et al., 2024). For example, those who live in countries with higher (vs. lower) median incomes are less willing to endorse physical confrontation and more willing to endorse gossip as appropriate third-party interventions (Eriksson et al., 2021). By examining how children across different societies reason about interventions, we can gain insight into how this understanding is influenced by culture.

Present Research

Here, we integrate all three of these considerations—the need to look at interventions other than punishment, to attend to social roles, and to go beyond Western samples—so as to generate a more comprehensive picture of children's conceptualization of third-party intervention. In particular, we explore children's conceptualization of third-party intervention in samples collected in Germany, India,

Uganda, and the United States. In each country, we told children about situations where a wrongdoing occurred and asked them their judgments about (a) whether an individual—a teacher or a peer, for instance—has an obligation to respond and (b) what sort of intervention children were expecting. Critically, we intentionally did not place emphasis on punishment when we asked these questions.

With respect to our selected age range, we tested children starting from early childhood (~6 years), when most kids are generally first socialized into formal schooling, to later childhood (~12 years), when kids have had greater exposure to social norms by virtue of having been in school for several years. We also examined an intermediate age group (9-year-olds) to track developmental change more continuously. In addition to differences in socialization, 6-year-olds also differ from 9- and 12-year-olds in their cognitive development. For instance, older children (~11-year-olds) are better able to understand complex relationships among objects (such as $w < x > y < z$) than 6-year-olds, who struggle to comprehend these relationships (Davidson et al., 2006; for a review, see Zelazo et al., 2003).

With respect to our participant populations, our goal was to ensure that our selected cultures were divergent from one another. This is because we were primarily concerned with questions about *generalizability*. Accordingly, the sampled locations varied on many theoretically relevant variables, such as community size (Henrich et al., 2010; Marlowe et al., 2008; Singh & Garfield, 2022) and collectivism (Eriksson et al., 2017); we further expand on these differences in the Method section. Additionally, it is important to acknowledge that our selection of testing sites was driven in part by practical limitations (McAuliffe, 2024)—we did not decide ahead of time which four countries would best test for generalization, but rather we were limited to countries where we had, or could get, access to sites for working with children.

Hypotheses

The present research aimed to uncover the ways in which social and cultural contexts, both independently and in combination, shape children's emerging social sense of obligation. While our approach was largely exploratory, we did formulate some general hypotheses.

Our first prediction pertained to developmental changes in children's obligation judgments. In particular, we expected that social role (e.g., whether a bystander is a teacher or a peer) and cultural context (e.g., whether a participant lives in the United States or Uganda) would play a comparatively less meaningful role in children's obligation judgments at younger ages (6-year-olds) than at older ages (12-year-olds); we were not sure whether 9-year-olds would more closely resemble 6-year-olds or 12-year-olds.

The prediction about increasing attention toward social role with age was based on past research in the United States that found that younger children (~4-year-olds) are less inclined to consider social role when ascribing obligations to engage in third-party punishment compared to older children (~7-year-olds; Marshall, Mermin-Bunnell, & Bloom, 2020).

The prediction about greater cultural variability at older ages was based on two lines of research. First, children's punitive behaviors come into alignment with normative cultural practices during middle childhood, meaning that younger children's punitive behaviors across cultures are more similar to one another than older children's (House et al., 2020). Second, research on children's obligation judgments in a different domain than intervention—in the domain of helping—finds greater differences across cultures in older children's (~9 years) judgments of whether a stranger is obligated to help someone in need compared to younger children's judgments (~6 years; Marshall et al., 2022; Miller et al., 1990). Helping, however, may be different from responding to wrongdoing—helping is more straightforwardly prosocial (sharing, comforting) compared to responding to wrongdoing, which may involve less prosocial actions (such as punishing). Indeed, research has found that helping versus punishing others are differentially influenced by social considerations (e.g., friendship, future interaction; O'Gorman et al., 2005). On the whole, then, it is possible that the way in which children reason about helping compared to intervening in response to wrongdoing over the course of development may be quite different—a possibility that the present work seeks to answer.

Our second prediction pertained to children's unstructured descriptions of third-party intervention. We predicted that—despite age-related variation in obligation judgments—children across ages would provide role-specific descriptions for bystanders, similar to what has been found in previous research (Marshall, Mermin-Bunnell, & Bloom, 2020). More specifically, we predicted that children would describe authority figure intervention as involving different actions than peer intervention. We predicted that children, regardless of age, would indicate that teachers intervene via directly punitive measures, such as sending the transgressor to time-out or removing resources from the transgressor, whereas peers intervene by reporting the transgressor to an authority figure or by helping the victim. We also predicted some degree of cross-cultural variation, a finding that would align with past work with adult samples (e.g., Eriksson et al., 2021). In particular, based on conversations with our local collaborators, descriptions of physical punishment (corporal punishment) may emerge in some sites, namely India and Uganda, but not in others, such as Germany and the United States.

Method

Transparency and Openness

All verbatim materials are provided on the Open Science Framework at https://osf.io/6mgpy/ (Marshall, 2024). That repository also includes all data collected from this study. The present work was not preregistered.

Participants

We tested 447 total participants—106 in Germany (52 males, 54 females), 108 in India (56 males, 52 females), 111 in Uganda (50 males, 61 females), and 122 in the United States (48 males, 74 females). Within each site, we sought to test approximately 36 6-year-olds, 36 9-year-olds, and 36 12-year-olds, aiming for ~108 participants per site.[1] These sample sizes per age group were initially determined by practical constraints; we tested the maximum number of children possible within the data collection timeframe. Gender was ascertained by asking participants their gender; the same was the case for age. No other demographic information was collected about participants.

[1] In the Germany sample, 15 (out of 106) participants had previously participated in another conceptually similar study (Marshall et al., 2022) because of resource constraints at the place of recruitment. This overlap did not occur at any other testing locations.

Table 1
Descriptions of Testing Sites

Societal characteristic	Germany	India	Uganda	United States
(a) Variation in demographic characteristics				
% Urban	46%	25%	14%	58%
Urban vs. rural	Urban	Urban	Rural	Urban
Population count (1 km resolution)	1,348.79	13,204.32	254.03	2,576.74
Religion	Christian	Hinduism	Christian	Christian
% Indicating religion is important	36%	82%	95%	54%
% elementary school or less	15%	68%	47%	5%
Mean income	$24,046	$1,314	$1,166	$31,974
(b) Variation in psychological/sociological characteristics				
Tight looseness	83	44	35	58
Individualism	79	24	30	60
(c) Variation in attitudes about law and order				
% Confidence in judicial system	70%	82%	40%	45%
% Confidence in police	88%	81%	61%	75%
% Reported money/property stolen	8%	11%	46%	14%

A sensitivity analysis revealed that we had 95% power ($\alpha = .05$) to detect a large within–between interaction (Age Group × Social Role) effect ($\eta_p^2 = .14$) within each site, which had approximately 108 participants. For precise age breakdowns and information about exclusions, see Supplemental Table S1. Twenty-one additional children were tested but excluded because they did not fall within the predetermined age bins.

Age Range Selection

We settled on these age groups for three reasons. First, the youngest and oldest age groups were determined based on practicality considerations (McAuliffe, 2024). The youngest age group we could reliably test across all of our sites was 6 years old, and the oldest age group was 12. We tested an intermediate age group, 9-year-olds, to capture developmental change.

Second, the tested age groups (6-, 9-, and 12-year-olds) generally align with the age range among samples in related research (e.g., House et al., 2020). Importantly, that work found that children's social behaviors appear to align with adults' social behaviors around age 12 (our oldest group), meaning that testing adolescents may add little additional value to understanding at what point children express an adult-like understanding of third-party intervention.

Third, we aimed to test children across an age range that captured variability in cognitive capacities and skills, such as executive function. Although core features of executive functioning appear by age 3 (Hughes, 1998), children between 6 and 12 years old still appear to undergo substantial cognitive development in domains such as cognitive flexibility, which may relate to an ability to recognize that certain agents have social responsibilities while others do not (Davidson et al., 2006). Although the present work does not measure these facets of cognitive development specifically, research such as ours that tests children across a wide age range can set the stage for such investigations in the future by providing initial insight into potential age-related changes.

Country Selection

With respect to the particular countries tested, we aimed to collect data across societies that vary on many dimensions, including Westernization, socioeconomic status, and community size, within the confines of the lab's monetary and practical limitations. This is an example of what Amir and McAuliffe (2020) have called "the breadth" approach in developmental science.

We outline many of the ways in which our samples differ in Table 1. In particular, we characterized our sample on different dimensions under three broad categories, including (a) variation in demographics, (b) variation in psychological/sociological constructions, and (c) variation in beliefs pertaining to law and order. Some of these variables have been linked to variability in punishment behavior and beliefs, while some of them have not. For example, higher population count has been linked to greater punishment behavior (e.g., Marlowe et al., 2008; although see Kanngiesser et al., 2022). Religious beliefs have been linked to variation in punishment behavior (Laurin et al., 2012). Median income has been linked to variations in adult norms about appropriate third-party intervention (Eriksson et al., 2021).

Some of the measures included in Table 1 were determined via the Gallup (2022) World Poll: % urban (question: urban/rural), religion (question: religion), % indicating religion is important (question: religion important), % elementary school or less (question: education level), % confident in judicial system (question: confidence in judicial system), % confident in police (question: confidence in local police), % reported money/property stolen (question: money/property stolen). "Tight looseness" was determined via Uz (2015). Mean income was ascertained via *World Population Review* (https://worldpopulationreview.com/country-rankings/median-income-by-country). Individualism scores were determined from The Culture Factor Group (2024); Uganda was not included in this website, so values were generated from Rarick et al. (2013). Individualism was scored from 1 to 100, with higher values indicating greater individualism. Each of these values are reported on the country level. Urban versus rural was determined via our collaborators who worked in the particular population and refers to the specific location where we worked. Population count, calculated using https://WorldPop.com, refers to the number of people living around the latitude and longitude of the specific testing sites (at a 1-km resolution).

Because the chosen samples vary in many ways from one another, we did not have specific predictions about how older children's

(9- and 12-year-olds) judgments and descriptions of third-party intervention may specifically vary across countries. Rather, we see the present investigation as a first step to uncovering the extent to which younger children's understanding of third-party punishment varies across different cultures.

Materials and Procedure

The procedure involved three components: (a) obligation and expectation judgments, (b) attention checks, and (c) unstructured descriptions of third-party intervention.

Obligation and Expectation Judgments

Participants were presented with six different stories—three featuring females ("female story set") and three featuring males ("male story set"). We included these two story sets to ensure the present effects generalize beyond a single gender. For example, it could be the case that children expect women bystanders to intervene less than male bystanders. Each story featured three individuals: (a) a bystander, (b) a transgressor, and (c) a victim (Figure 1). For all sites, we adjusted the images of the scenes to portray children who resembled children in each community and ensured that all children in the stories had names similar to those in their community. All materials were translated by a native speaker and back-translated by a separate native speaker, then reviewed for accuracy by yet another native speaker. In each story, the transgressor always pushed the victim; we chose this transgression as children almost universally view physical violence negatively (Buon et al., 2014). Within each story set, we manipulated the social role of the bystander to be either: (a) a teacher (authority adult), (b) another student's parent (nonauthority adult), or (c) a peer (nonauthority child); within each story set, the order in which these three conditions appeared was randomized, and whether the female story set came first was counterbalanced. The inclusion of the nonauthority adult condition, contrary to similar previous work that only included an authority and a peer (Marshall, Mermin-Bunnell, & Bloom, 2020), ensures that any effects of authority are not a result of an age confound.

After the presentation of a story, participants either responded to questions about obligations ("do you think s/he *has* to do something about this?") or questions about expectations ("do you think s/he *will* do something about this?"). We did not want children to respond to the obligation and expectation questions for the same story back-to-back, as we thought doing so could anchor participants to a specific response option. To resolve this problem, we presented each of the six stories followed by the obligation question in a single block followed by another block of all six stories followed by the expectation question, meaning that children heard all six stories twice; we counterbalanced which questions participants responded to first.

We note four additional components. First, even though we were primarily interested in obligation judgments because such notions are thought of as central to social judgment (Tomasello, 2020), we also measured expectation judgments because of previous research showing that children's expectations about social behavior diverge from their obligation judgments (Marshall, Wynn, & Bloom, 2020). Second, the experimenter followed up with participants if they indicated "yes" to either the obligation or expectation question by asking, "How much do you think s/he has to (will)? A tiny bit, a little bit, or a lot?" We did so because we wanted to allow children to exhibit nuance in their judgments while also measuring their initial "yes" versus "no" judgments. Third, we used the specific "have to" language for the obligation questions based on previous developmental work that has found that children distinguish between what people have to do from what they usually do or like to do (Kalish & Lawson, 2008; Marshall et al., 2022) and also use such language when referring to others' obligations (Mammen et al., 2021).

Attention Checks

After the obligation and expectation blocks, the experimenter presented participants with two attention check items in randomized order loosely based off of previous research (Miller et al., 1990). The primary role of these items was to ensure that children were providing sensible responses to straightforward questions that used "have to" language.

In one of the questions, the experimenter showed an image of a boy and said, "This is Matt. He wants to buy this shirt. Does Matt have to steal this shirt?" If participants responded "yes," the experimenter followed up by asking, "How much do you think he has to? A tiny bit, a little bit, or a lot?" In the second question, the experimenter showed an image of a girl and said, "This is Marie. Her mom asked her a question. Does Marie have to lie to her mom?" If participants responded "yes" the experimenter followed up by asking, "How much do you think she has to? A tiny bit, a little bit, or a lot?" These questions ensured that children were willing, under certain circumstances, to respond "no" to a question regarding obligations and that they understood the materials presented to them.

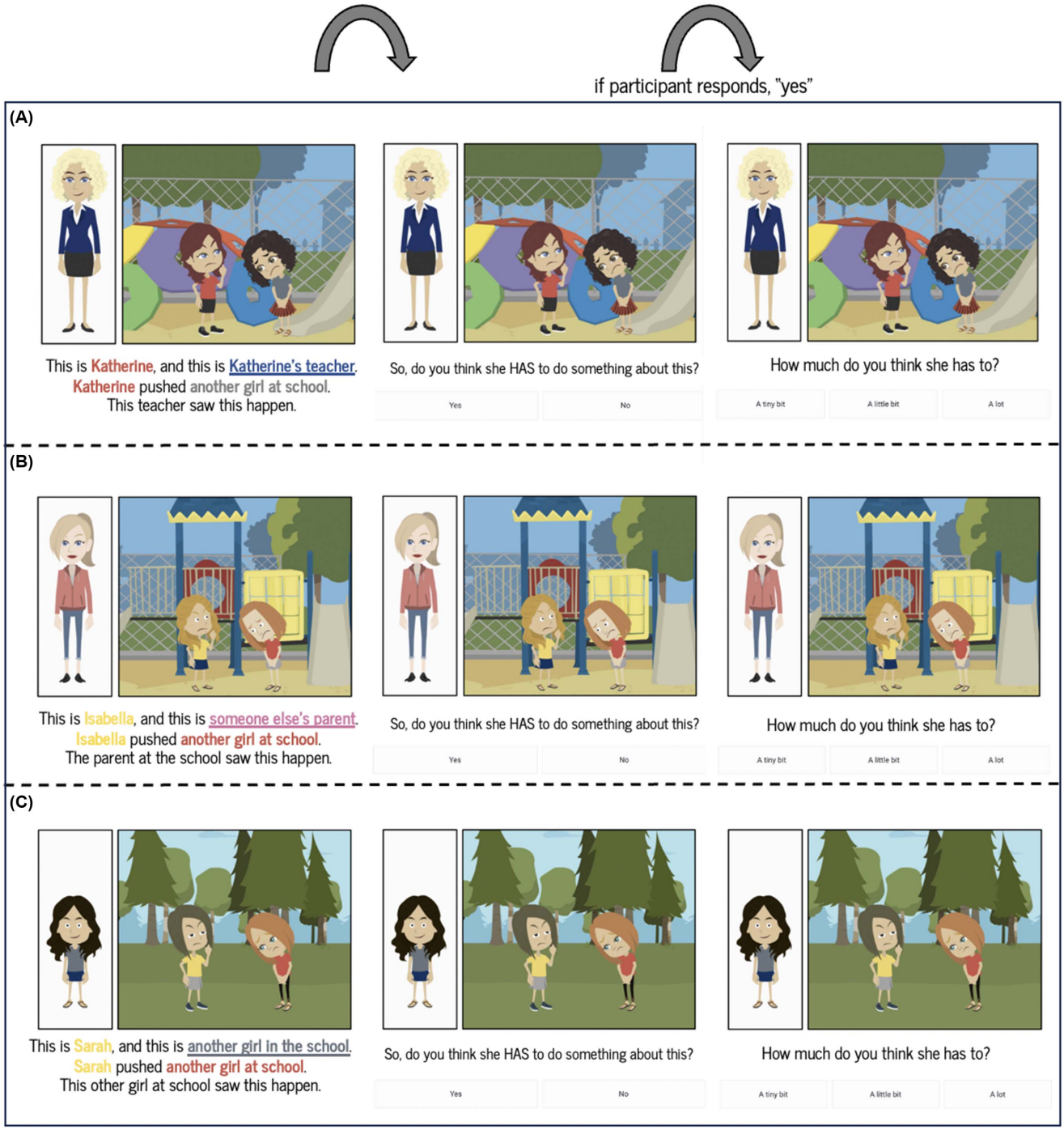
Unstructured Descriptions of Third-Party Intervention

Finally, the experimenter presented three unstructured description items about each bystander (teacher, another student's parent, a peer) in randomized order. For these questions, the experimenter returned only to the male story set (and not the female story set) as we needed to minimize the length of the study. Specifically, the experimenter briefly redescribed each story (i.e., "Now, remember this story about Robert and his teacher? Let's say the teacher does do something about this. Can you tell me what you think he'd do?"). Because we were not able to acquire videotape permission in Germany or Uganda, experimenters were instructed to write down exactly what participants said and record it in a datasheet. These instructions were not followed for all experimenters in Germany and, as a result, the majority of the responses reflected experimenters' interpretations of the child's statements. For this reason, we do not include the German responses in the unstructured description analysis. For sites where we were able to acquire videotape permission (India, United States), we transcribed children's responses verbatim and then translated them into English (for India).

Results

We first examined children's obligation judgments followed by their unstructured descriptions. For brevity's sake, we have relegated children's expectation judgments to the online Supplemental Material because the primary research question pertained to children's obligation judgments; while children's expectation judgments overlapped with their obligation judgments, their expectation

Figure 1
Study Material Example



Note. (A) Example of female story for "Teacher" condition with the obligation question, (B) Example of female story for "Adult" condition with obligation questions, (C) Example of female story for "Peer" condition with obligation question. © 2019 GoAnimate, Inc. Images are copyrighted by and used by permission of VYOND™. Characters licensed from Vyond, is a trademark of GoAnimate, Inc., registered in Argentina, Australia, Brazil, Chile, the European Union, Hong Kong, India, Indonesia, Israel, Japan, Mexico, New Zealand, Norway, African Intellectual Property Organization, the Philippines, Russia, Singapore, Switzerland, the United Kingdom, and Vietnam. See the online article for the color version of this figure.

judgments did not account for the present results. We also relegated participants' responses to the attention items to the online Supplemental Material; for those items, as expected, children overwhelmingly judged individuals as not having to engage in immoral behavior.

Obligation Judgments

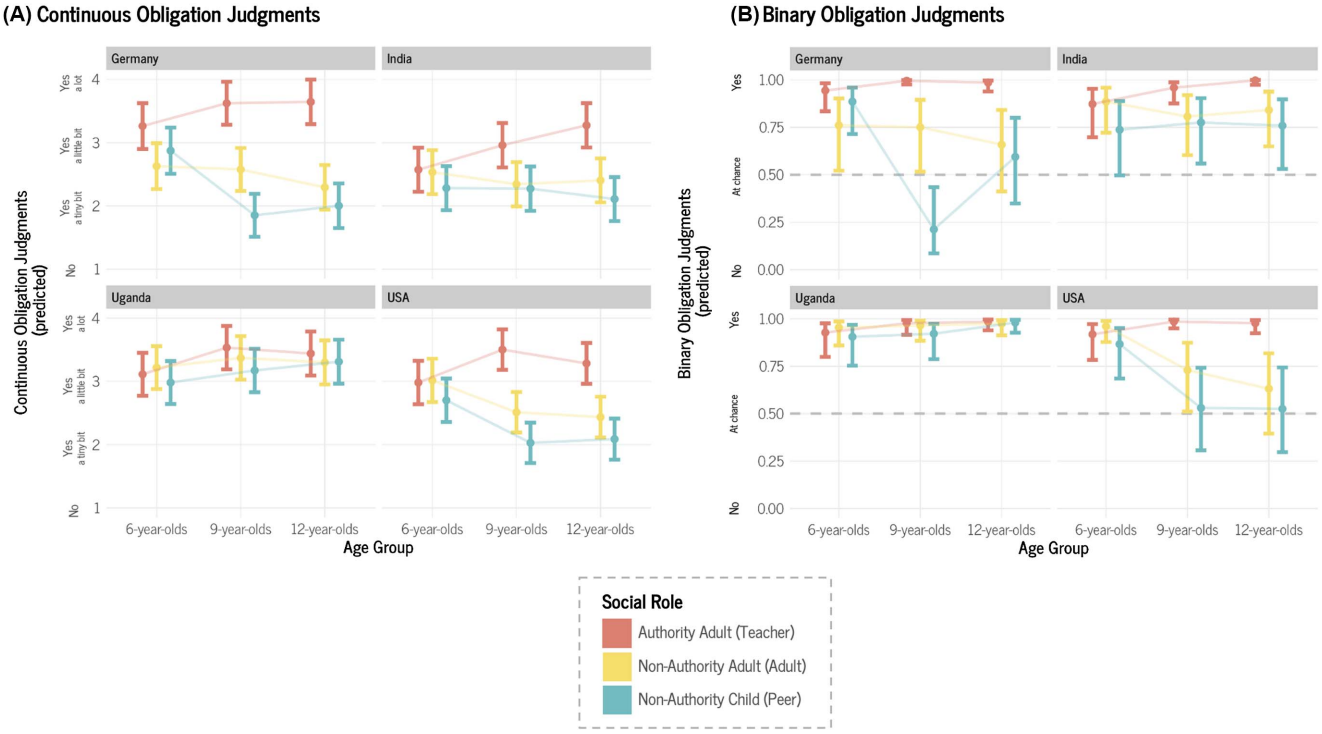
Using the lme4 package (Bates, 2010), we fit linear mixed-effects models with participant ID included as a random intercept to account for individual variability across participants and model the correlation within a participant's responses. Obligation judgments were the outcome variable ("Do you think s/he has to do something about this …"; 1 = *no*; 2 = *yes*, *a tiny*; 3 = *yes*, *a little bit*; 4 = *yes*, *a lot*). Four separate models were fit and compared. Model 1 included the main effects of age group (6-year-olds, 9-year-olds, 12-year-olds; between subjects), social role (authority adult [hereinafter, teacher], nonauthority adult [hereinafter, adult], nonauthority child [hereinafter, peer]; within subjects), country (Germany, India, Uganda, United States; between subjects), and scenario type (male stories, female stories; within subjects). Model 2 added all the two-way interactions among these variables. Model 3 added all the three-way interactions among these variables. Model 4 add the four-way interaction among these variables. Notably, we did not predict that scenario type would matter and only included this predictor in our models to account for differences between scenarios when looking at the main effects and interactions of interest among the theoretically relevant variables (age group, social role, country). We also fit the same models predicting participants' obligation judgments in a binary format (0 = *no*, 1 = *yes*); results are depicted here, but modeling details are in the Rmarkdown in the online Supplemental Material.

First, we conducted model comparisons. Model 2, which included the two-way interactions between the predictor variables, exhibited improved model fit beyond Model 1, $\chi^2(23) = 119.19$, $p < .001$. Model 3, which included the three-way interactions between the predictor variables, exhibited improved model fit beyond Model 2, $\chi^2(28) = 52.36$, $p = .003$. Model 4, which included the four-way interaction between the predictor variables, did not exhibit improve model fit beyond Model 3, $\chi^2(12) = 4.20$, $p = .980$, which indicates the lack of meaningful scenario type interactions. These results indicate that Model 3 bests fits the data.

Second, we examined Model 3 further. It revealed an Age Group × Social Role × Country interaction, $F(12, 2178) = 3.87$, $p < .001$; all other three-way interactions were not significant (all $ps > .751$). See Rmarkdown in the online Supplemental Material for full model outputs. Figure 2 depicts this interaction both for participants' continuous obligation judgments and their binary ones. Table 2 contains the means and 95% confidence intervals. Generally speaking, Figure 2 shows that children's obligation judgments differed depending on age, such that children generally became more sensitive

Figure 2

Participants' (A) Continuous and (B) Binary Obligation Judgments



Note. Panel A portrays the predicted values from Model 3 for the Age Group × Country × Social Role interaction predicting obligation judgments (continuous). Panel B portrays the predicted values from a model including the Age Group × Country × Social Role interaction predicting obligation judgments (binary). The dotted line indicates "at chance" responding. See Rmarkdown in the online Supplemental Material for full analysis of binary data. Error bands represent ± confidence intervals. See the online article for the color version of this figure.

to social roles with age, but importantly, those developmental trajectories also varied depending on country.

Third, we unpacked the Age Group × Social Role × Country interaction by fitting a model with Age Group × Social Role interaction predicting obligation judgments for each country. In this model, we also included the two-way interactions between Scenario × Social Role and Scenario × Age Group to control for any scenario-related effects, and we also included participant ID as a random intercept to account for individual variability across participants and model the correlation within a participant's responses. This approach revealed an Age Group × Social Role interaction for Germany, India, and the United States: Germany: $F(4, 517) = 11.81$, $p < .001$; India: $F(4, 524) = 4.95$, $p < .001$; the United States: $F(4, 597) = 11.96$, $p < .001$; however this approach did not reveal an Age Group × Social Role interaction for Uganda, $F(4, 544) = 1.69$, $p = .150$. These results indicate that children's obligation judgments varied depending on age and social role in Germany, India, and the United States but not in Uganda. For all estimates and confidence intervals associated with the Age Group × Social Role interaction in each country, see Supplemental Table S2.

Fourth, we examined the simple effect of social role for each age group in Germany, India, and the United States (Table 2; social role effect) to fully understand the nature of the observed Age Group × Social Role interactions. Note that we do not discuss Uganda here because we did not find an Age Group × Social Role interaction, although the simple effects are in Table 2. In Germany and the United States, we found similar patterns. Although 6-year-olds in Germany and the United States were sensitive to social role when ascribing obligations, they were considerably less sensitive compared to both 9- and 12-year-olds. In India, we found a similar pattern, but in this case, 6-year-olds were not sensitive to social role; only 9- and 12-year-olds were. These findings confirm our first prediction that younger children's obligation judgments are relatively immune to social and cultural context and our second prediction that older children's judgments are comparatively more sensitive to social and cultural context.

Fifth, we examined the simple effect of age group for each social role in Germany, India, and the United States (again, we omitted Uganda here). In doing so, we can better understand why age-related changes are emerging. In Germany, we found an effect of age group for peers, $F(2, 223) = 11.84$, $p < .001$, but not teachers or parents ($ps < .197$). In India, we found an effect of age group for teachers, $F(2, 290) = 5.74$, $p = .004$, but not peers or adults ($ps < .298$). In the United States, we found an effect of age group for both the adult, $F(2, 199) = 5.55$, $p = .005$, and peer, $F(2, 200) = 6.01$, $p = .003$, but not for teachers, $F(2, 199) = 1.65$, $p = .194$. These findings demonstrate how developmental changes in obligation judgments emerge for different reasons in varying cultural contexts. That is, developmental change is driven by variation in judgments about peers in Germany, teachers in India, and both adults and peers in the United States.

Finally, we also fit a model specifically for Uganda in which we examined the main effects of age group, social role, and scenario type in predicting obligation judgments (given the lack of an observed interaction among these variables). We found that children's obligation judgments did not vary depending on age group, $F(2, 108) = .71$, $p = .493$, suggesting that children regardless of age described bystanders as obligated to respond to wrongdoing. Their obligation judgments did vary depending on social role, $F(2, 552) = 5.36$, $p = .005$: children were more inclined to describe a teacher as obligated to respond to wrongdoing relative to a peer, $p = .004$; the other pairwise comparisons were not significant ($ps < .116$). There was not an effect of scenario type, $p = .914$. These results indicate that, although children did distinguish between the bystanders, the developmental pattern of differentiation varied from the three other countries. Indeed, children in Uganda rated each of the bystanders as obligated to respond to wrongdoing, which was not the case in either Germany or the United States.

Overall, these results align with our general predictions that younger children are generally less sensitive to social context when ascribing social obligations compared to older children. These findings suggest that culture tends to play a more potent role in

Table 2

Descriptive Statistics of Participants' Obligation Judgments (Continuous and Binary)

	Teacher		Adult		Peer		
Country	Continuous	Binary	Continuous	Binary	Continuous	Binary	Social role effect
Germany							
6-year-olds	3.26 [2.90, 3.63]	.94 [.83, .98]	2.63 [2.27, 3.00]	.76 [.52, .90]	2.87 [2.51, 3.24]	.89 [.71, .96]	$F(2, 517) = 6.54$, $p = .002$
9-year-olds	3.62 [3.28, 3.96]	1.00 [.98, 1.00]	2.58 [2.24, 2.92]	.75 [.52, .89]	1.85 [1.51, 2.19]	.21 [.09, .43]	$F(2, 517) = 68.35$, $p < .001$
12-year-olds	3.65 [3.29, 4.00]	.99 [.94, 1.00]	2.29 [1.94, 2.65]	.66 [.41, .84]	2.00 [1.65, 2.36]	.59 [.35, .80]	$F(2, 517) = 58.93$, $p < .001$
India							
6-year-olds	2.57 [2.22, 2.92]	.87 [.70, .95]	2.53 [2.19, 2.88]	.89 [.72, .96]	2.28 [1.93, 2.63]	.74 [.50, .89]	$F(2, 523) = .95$, $p = .387$
9-year-olds	2.96 [2.61, 3.31]	.96 [.88, .99]	2.34 [1.99, 2.69]	.81 [.60, 92]	2.27 [1.92, 2.62]	.78 [.56, .90]	$F(2, 525) = 9.72$, $p < .001$
12-year-olds	3.28 [2.92, 3.63]	1.00 [.97, 1.00]	2.40 [2.06, 2.75]	.84 [.65, .94]	2.11 [1.76, 2.46]	.76 [.53, .90]	$F(2, 524) = 24.59$, $p < .001$
Uganda							
6-year-olds	3.11 [2.77, 3.45]	.93 [.80, .98]	3.22 [2.88, 3.56]	.95 [.86, .99]	2.98 [2.64, 3.32]	.91 [.75, .97]	$F(2, 544) = 2.93$, $p = .054$
9-year-olds	3.53 [3.19, 3.88]	.98 [.92, .99]	3.37 [3.03, 3.72]	.96 [.88, .99]	3.17 [2.83, 3.52]	.92 [.79, .97]	$F(2, 544) = 4.80$, $p = .009$
12-year-olds	3.44 [3.09, 3.79]	.98 [.94, 1.00]	3.30 [2.95, 3.65]	.97 [.91, .99]	3.31 [2.96, 3.66]	.98 [.93, .99]	$F(2, 544) = 1.01$, $p = .365$
United States							
6-year-olds	2.98 [2.64, 3.33]	.92 [.78, .97]	3.02 [2.67, 3.36]	.96 [.88, .99]	2.70 [2.36, 3.05]	.87 [.69, .95]	$F(2, 597) = 3.11$, $p = .045$
9-year-olds	3.50 [3.18, 3.82]	.99 [.95, 1.00]	2.51 [2.19, 2.83]	.73 [.51, .87]	2.03 [1.71, 2.35]	.53 [.31, .74]	$F(2, 597) = 64.77$, $p < .001$
12-year-olds	3.29 [2.96, 3.61]	.98 [.92, .99]	2.44 [2.11, 2.76]	.63 [.39, .82]	2.09 [1.76, 2.41]	.53 [.30, .74]	$F(2, 597) = 41.04$, $p < .001$

Note. The mean value is displayed first, with confidence intervals in brackets. The "social role effect" refers to the effect of social role within a particular country and age group and pertains to the continuous analyses.

shaping obligation judgments at older ages. In particular, we find that 6-year-olds either do not consider social role when ascribing obligations to respond to wrongdoing (India, Uganda) or are considerably less inclined to do so compared to older children (Germany, United States). Furthermore, 6-year-olds' binary responses in all societies revealed that they consider all agents obligated to respond to wrongdoing. In cases where a developmental shift emerged (Germany, India, United States), it was because older children considered social role considerably more than younger children when ascribing obligations to respond to wrongdoing. The underlying reasons for these shifts differed across societies, however. In Germany, this developmental shift was driven by older children judging peers as less obligated to intervene; in India, this shift was driven by older children judging teachers as more obligated to intervene; in the United States, this shift was driven by older children's judgment on both teachers and adults as less obligated to intervene.

Unstructured Descriptions of Third-Party Intervention

We examined children's unstructured descriptions of third-party intervention. We generated a coding scheme based off of previous work (Marshall, Mermin-Bunnell, & Bloom, 2020; Marshall & McAuliffe, 2022). This scheme established four higher order categories; please see Table 3: (a) stop-gap measures, (b) punishment-specific responses, (c) punishment-related responses, and (d) victim-related responses. What distinguished the punishment-specific responses category from the punishment-related responses category was that the former clearly referenced a cost imposed on the transgressor (i.e., directly harming), whereas the latter did not (Marshall & McAuliffe, 2022). For instance, although telling a teacher about another person's transgression often leads to punishment, tattling in and of itself is not punishment per se and thus would be categorized as punishment-related responses. From a theoretical standpoint, punishment-related responses can also be closely aligned with notions of "blame" as distinct from "punishment" (Malle, in press).

Table 3
Different Description Categories and Corresponding Examples

Description type	Example
Stop-gap measures	"Split them up" "Stop them from pushing each other"
Punishment-specific responses	
Corporal punishment	"Caning them with three strokes"
Noncorporal punishment	"Give him a timeout"
Unspecified punishment	"They will punish him"
Punishment-related responses	
Reporting to a higher authority	"Tell the teacher"
Verbal scolding	"The teacher will scold him"
Verbal intervention	"The teacher will explain that people should not push at school"
Victim-related responses	
Reconciliation	"Try to get them to get along"
Forgiveness	"Tell him to ask for forgiveness"
Helping the victim	"Help the other boy"
Irrelevant responding	
I don't know	"I don't know"
Nothing	"Nothing can be done"
Other	"The teacher will be sad"
Not recorded	NA

Note. NA = not applicable.

Having established this coding scheme, two independent coders assessed all participants' responses. We had 330 descriptions from India (110 per role), 333 descriptions from Uganda (111 per role), and 366 descriptions from the United States (122 per role; for a total of 1,029 descriptions). When a description drew on multiple categories, we advised coders to assign the description a primary code depending on which theme appeared first in the description. Coders also coded for supplementary codes, although primary codes were used for analyses. By conventional standers, we observed high interrater reliability, $k = .94$. Because of this, we opted to use just one coder's assignments, which was determined randomly. Before conducting our main analyses, we eliminated responses that were characterized as irrelevant ($n = 223$; 22% of participants' responses); see online Supplemental Material for full details.

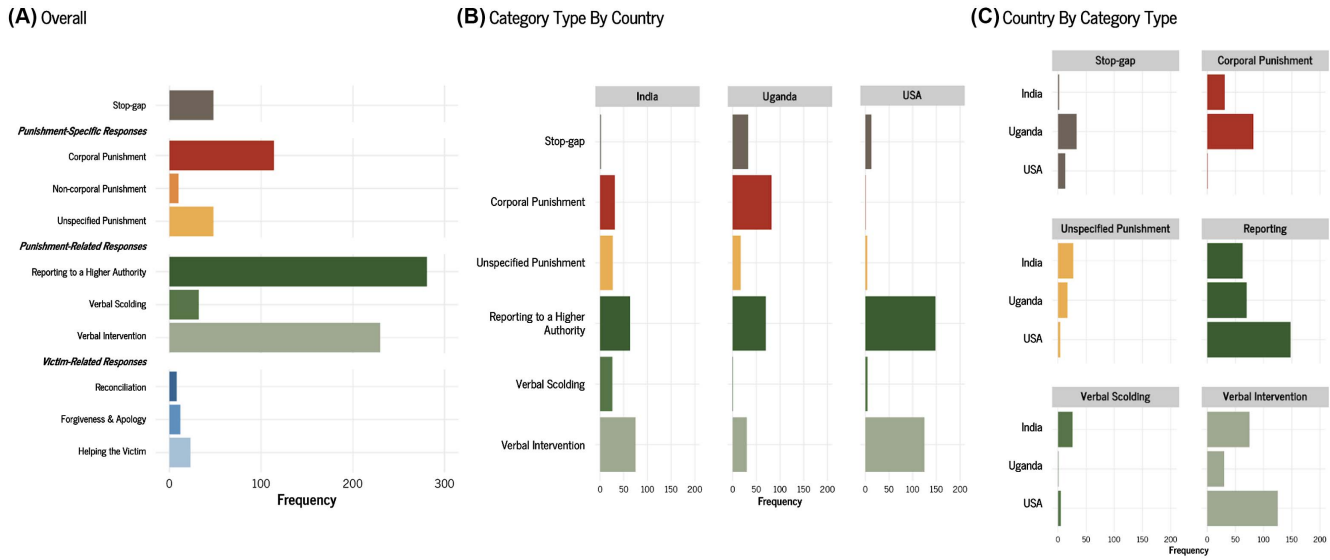
We found that participants conceptualized intervention as involving *reporting to a higher authority* the most ($n = 281$; 34.86%), followed by *verbal intervention* ($n = 230$; 28.54%), *corporal punishment* ($n = 114$; 14.14%), *unspecified punishment* ($n = 48$; 5.96%), *stop-gap measures* ($n = 48$; 5.96%), *verbal scolding* ($n = 32$; 3.97%), *helping the victim* ($n = 23$; 2.85%), *forgiveness and apology* ($n = 12$; 1.49%), *noncorporal punishment* ($n = 10$; 1.24%), and *reconciliation* ($n = 8$; .99%; Figure 3A). Generally speaking, these findings suggest that, when setting aside other factors (such as social role), participants most often conceptualized third-party intervention as involving Punishment-Related Responses, such as reporting to a higher authority or intervening verbally.

We continued by examining children's descriptions (category type: stop-gap, corporal punishment, noncorporal punishment, etc.) as a function of the included theoretically relevant factors: country (India, Uganda, United States), social role (teacher, adult, peer), and age group (6-year-olds, 9-year-olds, 12-year-olds). We eliminated categories that were generated at very low frequencies to simplify the analyses; we focused solely on categories that comprised more than 15% of descriptions within the smallest level of analyses (i.e., Ugandan 6-year-olds for teachers). Doing so retained stop-gap measures, corporal punishment, noncorporal punishment, reporting to a higher authority, verbal scolding, and verbal intervention (see online Supplemental Material for full descriptive statistics; Supplemental Table S5).

With respect to country, a Fisher exact test revealed an overall association between category type and country ($p < .001$), suggesting that the sorts of descriptions children generated systematically differed by country. We followed up on these results in two ways. First, we examined the effect of category type for each country (Figure 3B). These analyses revealed that the sorts of descriptions children in India provided differed from both Uganda and the United States, and the sorts of descriptions children in Uganda provided differed from the United States ($ps < .001$). Second, we examined the effect of country for each category type. We found an effect for every category type ($ps < .001$; Figure 3C). Pairwise comparisons revealed that children in Uganda were particularly likely to reference stop-gap and corporal punishment descriptions compared to children in India and the United States (all $ps < .001$). Children in India were particularly likely to reference unspecified punishment and verbal scolding descriptions compared to children in Uganda and the United States ($ps < .001$). Children in the United States were particularly likely to reference reporting to a higher authority and

Figure 3

Participants' Descriptions of Third-Party Intervention



Note. (A) Frequencies of participants' descriptions by category type. (B) Frequencies of participants' descriptions by category type separated by country. (C) Frequencies of participants' descriptions by country separated by category type. See the online article for the color version of this figure.

verbal intervention compared to children in India and the Uganda (*p*s = .046).

Next, we examined whether children's conceptualization of third-party intervention differs by social role. Fisher's exact tests revealed that participants within each country generated different descriptions depending on social role (all *p*s < .001). These findings suggest that participants in each country recognized that bystanders who occupy different social roles engage in distinctive forms of third-party intervention. In all countries, we found an effect of social role for reporting to a higher authority (all *p*s < .001): Children were more inclined to describe intervention on behalf of a peer as involving reporting to an authority compared to an adult or a teacher. In India and Uganda, we found an effect of social role for corporal punishment (India: $p = .029$; Uganda: $p < .001$) and unspecified punishment (India: $p < .001$; Uganda: $p = .001$): Children were most inclined to describe intervention on behalf of a teacher as involving these sorts of actions followed by an adult and then a peer (Supplemental Table S6). In the United States, we found an effect of social role for verbal intervention ($p < .001$): Participants were most inclined to describe intervention on behalf of a teacher as involving verbal intervention followed by an adult and then a peer. Collectively, these findings suggest that children recognize that people in a position of authority (i.e., teachers) are likely to respond to wrongdoing via corporal punishment or unspecified punishment (for participants in India or Uganda) or verbal intervention (for participants in the United States). For peers, children in all societies indicated that peers intervene by reporting a transgressor to someone else.

Finally, we also examined the extent to which age influences the present results. We did not find age-related effects for any category type within each social role in Uganda and the United States. We only found age-related effects in India specifically for corporal punishment and verbal intervention in the teacher and parent conditions. For a full explanation of age-related effects, see online Supplemental Material and Supplemental Figure S3.

General Discussion

We investigated how children across four societies conceptualize third-party intervention. In general, we found that children at younger ages (i.e., 6-year-olds) were more inclined than children at older ages (i.e., 9- and 12-year-olds) to describe third-party intervention as obligatory regardless of the social role of a bystander. We also found that older children in Germany, India, and the United States were particularly likely to relegate the responsibility to intervene exclusively to someone in a position of authority (e.g., a teacher). This pattern of results did not emerge in Uganda, however. Relative to children in the other countries studied, children in Uganda exhibited a more consistent propensity across ages to consider all bystanders as obligated to intervene. On the whole, these findings confirm our prediction that younger children's (6-year-olds) obligation judgments are less sensitive to social role and cultural contexts compared to older ages (12-year-olds).

Beyond obligation judgments, we found that children describe third-party intervention differently depending on social role and culture. For the most part, children in India and Uganda indicated that third-party intervention on behalf of authority figures (such as teachers) involves corporal punishment (hitting the perpetrator), whereas children in the United States indicated that such third-party intervention involves verbal intervention (telling the perpetrator that it is wrong to hit others). When it comes to peers, children across ages and societies largely referenced actions that involved telling an authority figure (i.e., tattling). These findings suggest that children from a young age (6 years old) recognize that certain bystanders have different tools at their disposal to engage in third-party intervention depending on their social role.

When it comes to descriptions of third-party intervention, we did not find many age-related effects with the exception of India. In India, younger children referenced corporal punishment (like children in Uganda) and older children referenced verbal intervention (like children in the United States) when asked how teachers intervene in response to physical violence. Perhaps children in India learn as they get older that it is not appropriate to engage in corporal punishment. Alternatively, it is possible that children at older ages simply experience corporal punishment less and therefore mention it less. The present data cannot determine which of these two explanations best explains the age effect.

What factors explain children's varied references to corporal punishment across cultures? One explanation may be variation in income—indeed, adults in lower income countries are more inclined to endorse physical punishment, while adults in higher income countries are more inclined to endorse gossiping (Eriksson et al., 2021). As shown in Table 1, both India and Uganda have far lower mean incomes relative to Germany and the United States, and accordingly, we find that children in India and Uganda are more inclined to describe third-party intervention as involving corporal punishment compared to children in Germany and the United States. That being said, Table 1 also demonstrates how the four countries in which we tested vary on many other dimensions as well. We hope future work will be better suited to address the observed cross-cultural differences by testing sites other than the four that we worked with here.

Theoretical Implications

Our findings have theoretical implications for understanding children's developing sense of obligation (Tomasello, 2020). In particular, recent research suggests that children at younger ages exhibit less discriminate reasoning in terms of obligations compared to older children and adults (Dahl et al., 2020; Geraci et al., 2021; Marshall & Wilks, 2024; Marshall et al., 2022; Marshall, Mermin-Bunnell, & Bloom, 2020; Marshall, Wynn, & Bloom, 2020; Miller et al., 1990). For example, children around the age of 5 tend to think that bystanders have an obligation to help others regardless of whether the person in need is a family member, a friend, or a stranger. Only older children make robust distinctions. For instance, 9-year-olds, like adults, tend to consider bystanders as strongly obligated to help family members and to a lesser extent friends—but not obligated to help strangers (Marshall et al., 2022). To our knowledge, though, all previous work showing this pattern has explored obligations to help. Our current results suggest that children's reasoning about obligations appears to be consistent across varying forms of intervention (third-party helping and third-party intervention in response to wrongdoing).

Most research on how children and adults respond to wrongdoing has exclusively focused on punishment rather than considering a broader range of options, such as compensating victims or encouraging forgiveness (e.g., FeldmanHall et al., 2014). Many have argued that this focus has led to misguided conclusions regarding the prominence of punishment relative to these other third-party interventions (Van Doorn & Brouwers, 2020a, 2020b). Given the open-ended nature of our description measure, our findings suggest that the conclusion that punishment is a central tool used in responding to wrongdoing might not have been entirely misguided after all—punishment is indeed a key aspect of how children conceptualize third-party intervention on behalf of authority figures. This finding aligns with work arguing that direct punishment is especially relevant in the context of societies that are governed by social hierarchies (Binder, 2002; Malle, in press). In terms of peers, though, children were not particularly inclined to supply explanations that aligned with direct punishment. Instead, they were more likely to produce explanations that involved calling out an individual for their bad behavior or bringing in another person of higher authority, such as a teacher. Notably, though, these explanations were neither explicitly prosocial nor directly punitive given that such explanations infrequently referenced helping victims or punishing perpetrators.

Relatedly, cultural variation in children's descriptions of third-party intervention points toward the power of social learning in shaping children's understanding of how people respond to wrongdoing. Corporal punishment was rarely mentioned as a form of third-party intervention in the United States. This finding aligns with previous work that suggests that children in the United States do not tend to enact corporal punishment themselves when given the opportunity (Marshall et al., 2019). However, corporal punishment was frequently referenced in India and Uganda. This result indicates that children generally consider bystanders as responding to wrongdoing in culturally specific ways and, in our view, likely reflects the types of third-party intervention that children witness and experience in their everyday lives.

Outstanding Questions

The present findings raise many questions. First, why are younger children more inclined than older children and adults to describe all agents as obligated to respond to wrongdoing? We suspect that the answer to this question requires the synthesis of a variety of factors related to social and cognitive development. In terms of cognitive factors, it is possible that children at a younger age are more likely to make moral and social judgments on the basis of a singular principle or heuristic due to cognitive simplicity (Berndt & Berndt, 1975; Walden, 1982). In other words, younger children may have more difficulty making complex judgments that require coordinating among different factors, such as social role, because they have less refined executive functioning skills (Zelazo et al., 2003). This possibility is consistent with the finding, mentioned above, that children are similarly indiscriminate when it comes to their judgments about who is obligated to help.

In terms of social factors, children may receive ample messaging from trusted others (i.e., parents, teachers) reinforcing the negativity of wrongdoing and the importance of intervening. For example, children may repeatedly hear that transgressions are inappropriate and wrong, and this message may leave little room for nuance about precisely who is supposed to intervene—especially when heard through the ears of young children. As children get older, they may hear more nuanced conversations about responding to wrongdoing, such as not being a "tattle tale."

A second question relates to cultural differences. We discussed differences in children's intuitions about corporal punishment above. Here, we turn to the question of why younger children in Uganda are no more differentiating on the basis of social role than older children when ascribing obligations to intervene, while children in each of the other countries exhibited increasing sensitivity to social context with age. We did not predict this effect, but we can discuss a few speculative proposals.

One consideration is population size. As illustrated in Table 1, our Ugandan sample stands out among the other samples in its much smaller population density. Past work has argued that larger, more urban societies require greater institutionalized punishment systems in part because larger groups are more anonymous and therefore require designating certain individuals as "punishers" (e.g., police) to enforce laws rather than relying on more egalitarian and perhaps informal policing systems (Baldassarri & Grossman, 2011; O'Gorman et al., 2009; Traulsen et al., 2012). The present data seem to generally align with this argument—at older ages, children in larger, more urban communities (Germany, United States, India) were increasingly more likely to recognize that individuals within certain social positions, such as teachers, bear the responsibility to "right wrongs." It is the children in the smaller, more rural community (Uganda) that have comparatively different intuitions. That children in Uganda are not increasingly sensitive to social role may then reflect the reality of Ugandan children's environment whereby obligations to respond to wrongdoing are indeed less delineated on the basis of social role.

Another consideration is religion. As illustrated by Table 1, our Ugandan sample is unique in that the overwhelming majority of individuals in Uganda (95%) indicate that religion is important to them and also adhere to the Christian religion. It is possible that religious beliefs in Uganda play a role in maintaining children's heightened sense that peers are obligated to respond to wrongdoing relative to other countries. For example, children may learn that they should be their "brother's keeper." That is, children may learn that there is social and moral value in looking over others, even when not in a position of authority. Individuals in less religious communities may not hear these sorts of messages as frequently and, as a result, may be less inclined to consider peers as obligated to police one another. Consistent with this, research in the United States finds that more religious individuals tend to exhibit a greater willingness to punish transgressors (Laurin et al., 2012). Future work could test this possibility and also the possibility related to community size by collecting data in more samples across a larger number of countries with varying population sizes and religious affiliations.

Limitations

The present work is not without limitations. First, we did not include a control condition in which children were presented with a neutral or positive situation and asked to determine whether a bystander is obligated to do something about it. For that reason, the current data does not necessarily measure children's baseline willingness to respond "yes" to questions about obligation. Importantly, though, we did ask children to describe third-party intervention and found that, across the board, children referenced theoretically relevant interventions, such as punishment, so we do know that they were not just reflexively assenting to the experimenter's questions without any understanding. Also, the primary goal of the present work was to look at variation in children's judgments as a function of age, social role (i.e., a teacher vs. peer), and culture (i.e., Uganda vs. United States). Because the obligation question was the same across these conditions, the effects of age group, social role, and culture in addition to their interactions still stand and represent meaningful psychological variation.

Second, there are concerns about statistical power. Although we did find that older children are more sensitive to contextual factors (such as Social Role) compared to younger children, we did not find this age-related pattern in Uganda. As indicated by the sensitivity analyses described above, it is possible that such an interaction exists but that it is small. If this were the case, we would not have had adequate power to detect such an interaction. Additionally, even though we were able to acquire unstructured descriptions from most children, the sample sizes within the smallest level of analysis were quite small. As a result, the findings regarding developmental changes within a particular society for a specific age group in response to a certain social actor should be interpreted with caution.

Concluding Remarks

The present data demonstrates that young children across many divergent societies construe bystanders, regardless of their social role (e.g., teacher, peer), as obligated to respond to wrongdoing. However, these younger children do recognize that people who occupy different social roles (i.e., teachers compared to peers) engage in distinctive forms of third-party intervention. Older children in some societies (India, Germany, United States) eventually relegate such a responsibility to authority figures and are less committed to considering peers as obligated to engage in third-party intervention. Nonetheless, children in other societies (Uganda), despite recognizing that agents pursue different forms of third-party intervention depending on their role, maintain that all bystanders are obligated to respond to wrongdoing.

Constraints on Generality

One constraint on the generalizability of our findings is that we focused on children's responses to a singular moral violation, hitting. We suspect that different results may emerge for alternate transgressions, especially if these transgressions differ in severity. For example, how children think about third-party intervention may be quite different in situations where the wrongdoing has to do with unfairness, such as when an employer gives more money to one person than another, even if both have done the same work. If we had used such violations (e.g., Lee & Warneken, 2020), for instance, children may have more readily generated third-party intervention descriptions that align with victim compensation, and there would be less emphasis on punishment.

A second constraint pertains to our sample population. A strength of the present work is that it allows us to gain traction on questions of generality by testing children from many different countries, including non-Western ones. By doing so, we show that younger children's judgments (6-year-olds) are less sensitive to cultural surroundings compared to older children (9- to 12-year-olds), at least in the present experimental context. Nonetheless, we only worked with children in four countries and only did so at one site within each country. Because of this, we are limited in the sorts of inferences we can draw. In particular, we cannot say whether the between-country differences documented here are a result of true "country-level" differences or are instead a result of other differences between the sites that happen to covary with country. Future research could address this by examining more sites within a country that vary in theoretically relevant ways.

References

Amir, D., & McAuliffe, K. (2020). Cross-cultural, developmental psychology: Integrating approaches and key insights. *Evolution and*

Human Behavior, *41*(5), 430–444. https://doi.org/10.1016/j.evolhumbehav.2020.06.006

Baldassarri, D., & Grossman, G. (2011). Centralized sanctioning and legitimate authority promote cooperation in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(27), 11023–11027. https://doi.org/10.1073/pnas.1105456108

Balliet, D., Mulder, L. B., & Van Lange, P. A. (2011). Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin*, *137*(4), 594–615. https://doi.org/10.1037/a0023489

Bates, D. M. (2010). *lme4: Mixed-effects modeling with R*. https://people.math.ethz.ch/~maechler/MEMo-pages/lMMwR.pdf

Berndt, T. J., & Berndt, E. G. (1975). Children's use of motives and intentionality in person perception and moral judgement. *Child Development*, *46*(4), 904–912. https://doi.org/10.2307/1128396

Binder, G. (2002). Punishment theory: Moral or political? *Buffalo Criminal Law Review*, *5*(2), 321–372. https://doi.org/10.1525/nclr.2002.5.2.321

Buon, M., Jacob, P., Margules, S., Brunet, I., Dutat, M., Cabrol, D., & Dupoux, E. (2014). Friend or foe? Early social evaluation of human interactions. *PLOS ONE*, *9*(2), Article e88612. https://doi.org/10.1371/journal.pone.0088612

Cushman, F. (2015). Punishment in humans: From intuitions to institutions. *Philosophy Compass*, *10*(2), 117–133. https://doi.org/10.1111/phc3.12192

Dahl, A., Gross, R. L., & Siefert, C. (2020). Young children's judgments and reasoning about prosocial acts: Impermissible, suberogatory, obligatory, or supererogatory? *Cognitive Development*, *55*, Article 100908. https://doi.org/10.1016/j.cogdev.2020.100908

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

Eriksson, K., Strimling, P., Andersson, P. A., Aveyard, M., Brauer, M., Gritskov, V., Kiyonari, T., Kuhlman, D. M., Maitner, A. T., Manesi, Z., Molho, C., Peperkoorn, L. S., Rizwan, M., Stivers, A. W., Tian, Q., Van Lange, P. A. M., Vartanova, I., Wu, J., & Yamagishi, T. (2017). Cultural universals and cultural differences in meta-norms about peer punishment. *Management and Organization Review*, *13*(4), 851–870. https://doi.org/10.1017/mor.2017.42

Eriksson, K., Strimling, P., Gelfand, M., Wu, J., Abernathy, J., Akotia, C. S., Aldashev, A., Andersson, P. A., Andrighetto, G., Anum, A., Arikan, G., Aycan, Z., Bagherian, F., Barrera, D., Basnight-Brown, D., Batkeyev, B., Belaus, A., Berezina, E., Björnstjerna, M., … Van Lange, P. A. M. (2021). Perceptions of the appropriate response to norm violation in 57 societies. *Nature Communications*, *12*(1), Article 1481. https://doi.org/10.1038/s41467-021-21602-9

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, *25*(2), 63–87. https://doi.org/10.1016/S1090-5138(04)00005-4

FeldmanHall, O., Sokol-Hessner, P., Van Bavel, J. J., & Phelps, E. A. (2014). Fairness violations elicit greater punishment on behalf of another than for oneself. *Nature Communications*, *5*(1), Article 5306. https://doi.org/10.1038/ncomms6306

Gallup. (2022). *"Education level"*. Gallup World Poll.

Garfield, Z. H., Ringen, E. J., Buckner, W., Medupe, D., Wrangham, R. W., & Glowacki, L. (2023). Norm violations and punishments across human societies. *Evolutionary Human Sciences*, *5*, Article e11. https://doi.org/10.1017/ehs.2023.7

Geraci, A., Rigo, P., Simonelli, A., Di Nuovo, S., & Simion, F. (2021). Preschoolers' evaluations of comforting actions towards third parties in different relationship contexts. *Journal of Applied Developmental Psychology*, *76*, Article 101315. https://doi.org/10.1016/j.appdev.2021.101315

Geraci, A., & Surian, L. (2021). Toddlers' expectations of third-party punishments and rewards following an act of aggression. *Aggressive Behavior*, *47*(5), 521–529. https://doi.org/10.1002/ab.21979

Geraci, A., & Surian, L. (2023). Preverbal infants' reactions to third-party punishments and rewards delivered toward fair and unfair agents. *Journal of Experimental Child Psychology*, *226*, Article 105574. https://doi.org/10.1016/j.jecp.2022.105574

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., & Ziker, J. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, *327*(5972), 1480–1484. https://doi.org/10.1126/science.1182238

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., & Ziker, J. (2006). Costly punishment across human societies. *Science*, *312*(5781), 1767–1770. https://doi.org/10.1126/science.1127333

House, B. R., Kanngiesser, P., Barrett, H. C., Yilmaz, S., Smith, A. M., Sebastian-Enesco, C., Erut, A., & Silk, J. B. (2020). Social norms and cultural diversity in the development of third-party punishment. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1925), Article 20192794. https://doi.org/10.1098/rspb.2019.2794

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, *16*(2), 233–253. https://doi.org/10.1111/j.2044-835X.1998.tb00921.x

Kalish, C. W., & Lawson, C. A. (2008). Development of social category representations: Early appreciation of roles and deontic relations. *Child Development*, *79*(3), 577–593. https://doi.org/10.1111/j.1467-8624.2008.01144.x

Kanakogi, Y., Miyazaki, M., Takahashi, H., Yamamoto, H., Kobayashi, T., & Hiraki, K. (2022). Third-party punishment by preverbal infants. *Nature Human Behaviour*, *6*(9), 1234–1242. https://doi.org/10.1038/s41562-022-01354-2

Kanngiesser, P., Schäfer, M., Herrmann, E., Zeidler, H., Haun, D., & Tomasello, M. (2022). Children across societies enforce conventional norms but in culturally variable ways. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(1), Article e2112521118. https://doi.org/10.1073/pnas.2112521118

Laurin, K., Shariff, A. F., Henrich, J., & Kay, A. C. (2012). Outsourcing punishment to God: Beliefs in divine control reduce earthly punishment. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1741), 3272–3281. https://doi.org/10.1098/rspb.2012.0615

Lee, Y. E., Marshall, J., Deutchman, P., McAuliffe, K., & Warneken, F. (2022). Children's judgments of interventions against norm violations: COVID-19 as a naturalistic case study. *Journal of Experimental Child Psychology*, *221*, Article 105452. https://doi.org/10.1016/j.jecp.2022.105452

Lee, Y. E., & Warneken, F. (2020). Children's evaluations of third-party responses to unfairness: Children prefer helping over punishment. *Cognition*, *205*, Article 104374. https://doi.org/10.1016/j.cognition.2020.104374

Malle, B. F. (in press). Blame and punishment: Two distinct mechanisms for regulating moral behavior. In B. F. Malle & P. Robbins (Eds.), *Cambridge handbook of moral psychology*. Cambridge University Press.

Mammen, M., Köymen, B., & Tomasello, M. (2021). Young children's moral judgments depend on the social relationship between agents. *Cognitive Development*, *57*, Article 100973. https://doi.org/10.1016/j.cogdev.2020.100973

Marlowe, F. W., Berbesque, J. C., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., Ensminger, J., Gurven, M., Gwako, E., Henrich, J., Henrich, N., Lesorogol, C., McElreath, R., & Tracer, D. (2008). More 'altruistic' punishment in larger societies. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1634), 587–590. https://doi.org/10.1098/rspb.2007.1517

Marshall, J. (2024). *Cross-cultural conceptions of third-party intervention across childhood*. https://osf.io/6mgpy/

Marshall, J., Gollwitzer, A., Mermin-Bunnell, K., Shinomiya, M., Retelsdorf, J., & Bloom, P. (2022). How development and culture shape intuitions about prosocial obligations. *Journal of Experimental Psychology: General*, *151*(8), 1866–1882. https://doi.org/10.1037/xge0001136

Marshall, J., Gollwitzer, A., Wynn, K., & Bloom, P. (2019). The development of corporal third-party punishment. *Cognition*, *190*, 221–229. https://doi.org/10.1016/j.cognition.2019.04.029

Marshall, J., & McAuliffe, K. (2022). Children as assessors and agents of third-party punishment. *Nature Reviews Psychology*, *1*(6), 334–344. https://doi.org/10.1038/s44159-022-00046-y

Marshall, J., Mermin-Bunnell, K., & Bloom, P. (2020). Developing judgments about peers' obligation to intervene. *Cognition*, *201*, Article 104215. https://doi.org/10.1016/j.cognition.2020.104215

Marshall, J., & Wilks, M. (2024). Does distance matter? How physical and social distance shape our perceived obligations to others. *Open Mind: Discoveries in Cognitive Science*, *8*, 511–534. https://doi.org/10.1162/opmi_a_00138

Marshall, J., Wynn, K., & Bloom, P. (2020). Do children and adults take social relationship into account when evaluating people's actions? *Child Development*, *91*(5), e1082–e1100. https://doi.org/10.1111/cdev.13390

Martin, J. W., & Cushman, F. (2015). To punish or to leave: Distinct cognitive processes underlie partner control and partner choice behaviors. *PLOS ONE*, *10*(4), Article e0125193. https://doi.org/10.1371/journal.pone.0125193

McAuliffe, K. (2024). Sampling decisions in developmental psychology. *Nature Reviews Psychology*, *3*, 297–298. https://doi.org/10.1038/s44159-024-00306-z

McAuliffe, K., Jordan, J. J., & Warneken, F. (2015). Costly third-party punishment in young children. *Cognition*, *134*, 1–10. https://doi.org/10.1016/j.cognition.2014.08.013

McCullough, M. E., & Witvliet, C. V. (2002). The psychology of forgiveness. In C. R. Snyder Shane & J. Lopez (Eds.), *Handbook of positive psychology* (Vol. 2, pp. 446–455). Oxford University Press.

Miller, J. G., Bersoff, D. M., & Harwood, R. L. (1990). Perceptions of social responsibilities in India and in the United States: Moral imperatives or personal decisions? *Journal of Personality and Social Psychology*, *58*(1), 33–47. https://doi.org/10.1037/0022-3514.58.1.33

Molho, C., De Petrillo, F., Garfield, Z. H., & Slewe, S. (2024). Cross-societal variation in norm enforcement systems. *Philosophical Transactions of the Royal Society B*, *379*(1897), Article 20230034. https://doi.org/10.1098/rstb.2023.0034

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31–38. https://doi.org/10.1016/j.jecp.2017.04.017

O'Gorman, R., Henrich, J., & Van Vugt, M. (2009). Constraining free riding in public goods games: Designated solitary punishers can sustain human cooperation. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1655), 323–329. https://doi.org/10.1098/rspb.2008.1082

O'Gorman, R., Wilson, D. S., & Miller, R. R. (2005). Altruistic punishing and helping differ in sensitivity to relatedness, friendship, and future interactions. *Evolution and Human Behavior*, *26*(5), 375–387. https://doi.org/10.1016/j.evolhumbehav.2004.12.006

Rarick, C., Winter, G., Nickerson, I., Falk, G., Barczyk, C., & Asea, P. K. (2013). An investigation of Ugandan cultural values and implications for managerial behavior. *Global Journal of Management and Business Research Administration and Management*, *13*(9), 1–9.

Riedl, K., Jensen, K., Call, J., & Tomasello, M. (2012). No third-party punishment in chimpanzees. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(37), 14824–14829. https://doi.org/10.1073/pnas.1203179109

Riedl, K., Jensen, K., Call, J., & Tomasello, M. (2015). Restorative justice in children. *Current Biology*, *25*(13), 1731–1735. https://doi.org/10.1016/j.cub.2015.05.014

Singh, M., & Garfield, Z. H. (2022). Evidence for third-party mediation but not punishment in Mentawai justice. *Nature Human Behaviour*, *6*(7), 930–940. https://doi.org/10.1038/s41562-022-01341-7

Stavans, M., & Baillargeon, R. (2019). Infants expect leaders to right wrongs. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(33), 16292–16301. https://doi.org/10.1073/pnas.1820091116

The Culture Factor Group. (2024, February). *The country comparison tool*. https://www.hofstede-insights.com/country-comparison-tool

Tomasello, M. (2020). The moral psychology of obligation. *Behavioral and Brain Sciences*, *43*, Article e56. https://doi.org/10.1017/S0140525X19001742

Traulsen, A., Röhl, T., & Milinski, M. (2012). An economic experiment reveals that humans prefer pool punishment to maintain the commons. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1743), 3716–3721. https://doi.org/10.1098/rspb.2012.0937

Uz, I. (2015). The index of cultural tightness and looseness among 68 countries. *Journal of Cross-Cultural Psychology*, *46*(3), 319–335. https://doi.org/10.1177/0022022114563611

van Dijk, E., Molenmaker, W. E., & de Kwaadsteniet, E. W. (2015). Promoting cooperation in social dilemmas: The use of sanctions. *Current Opinion in Psychology*, *6*, 118–122. https://doi.org/10.1016/j.copsyc.2015.07.006

Van Doorn, J., & Brouwers, L. (2020a). Third-Party responses to injustice: A review on the preference for compensation. In D. Canter & D. Youngs (Eds.), *Reviewing crime psychology* (pp. 315–333). Routledge.

Van Doorn, J., & Brouwers, L. (2020b). Third-Party responses to injustice: A review on the preference for compensation. *Crime Psychology Review*, *3*(1), 59–77. https://doi.org/10.1080/23744006.2018.1470765

Vidmar, N., & Miller, D. T. (1980). Social psychological processes underlying attitudes toward legal punishment. *Law & Society Review*, *14*(3), 565–602. https://doi.org/10.2307/3053193

Walden, T. A. (1982). Mediation and production deficiencies in children's judgments of morality. *Journal of Experimental Child Psychology*, *33*(1), 165–181. https://doi.org/10.1016/0022-0965(82)90012-1

Wu, J., Luan, S., & Raihani, N. (2022). Reward, punishment, and prosocial behavior: Recent developments and implications. *Current Opinion in Psychology*, *44*, 117–123. https://doi.org/10.1016/j.copsyc.2021.09.003

Yudkin, D. A., Van Bavel, J. J., & Rhodes, M. (2020). Young children police group members at personal cost. *Journal of Experimental Psychology: General*, *149*(1), 182–191. https://doi.org/10.1037/xge0000613

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., Chiang, J. K., Hongwanishkul, D., Schuster, B. V., & Sutherland, A. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, *68*(3), 7–137. https://doi.org/10.1111/j.0037-976x.2003.00260.x

Received October 13, 2023
Revision received April 15, 2024
Accepted May 17, 2024 ■