Children's and Adults' Understanding of How Punishment Shapes Social Relationships

Young-eun Lee[1,2] and Larisa Heiphetz Solomon[1]

[1]Department of Psychology, Columbia University

1190 Amsterdam Ave., New York, NY 10027, United States

[2]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology

43 Vassar St., Cambridge, MA 02139, United States

Lee, Y., & Solomon, L. H. (2025). Children's and adults' understanding of how punishment shapes social relationships. *Journal of Experimental Child Psychology*, *260*, 106311. doi: 10.1016/j.jecp.2025.106331

Author Note

Correspondence concerning this article should be addressed to Young-eun Lee at yelee@mit.edu.

Acknowledgments

This project was made possible through the support of grant #61808 from the John Templeton Foundation and National Science Foundation CAREER Award #2044360 to LHS. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation or the National Science Foundation. Materials are available at https://osf.io/gf7bn/?view_only=05903f5be0a34a0a8b5f4cb1289ff317. Preregistrations are available at the following URLs: https://aspredicted.org/VT2_1HC (Study 1) and https://aspredicted.org/J43_XKL (Study 2).

Highlights

- We investigated how punishment influences children's and adults' perceptions of social relationships.
- Participants' attitudes toward punishers and targets were shaped by whether the punishment was motivated by a transgression.
- Participants generally inferred negative social relationships between punishers and their targets.

Abstract

Punishment regulates selfish behaviors and maintains cooperation. However, because punishment imposes costs on another person, it could also harm relationships. The current work asked how punishment shapes 5- to 10-year-olds' (Study 1; *n*=128) and adults' (Study 2; *n*=159) attitudes toward punishers and those who receive punishment as well as their inferences about relationships between punishers and targets. We reasoned that the motives underlying punishment might shape evaluations; punishments motivated by prosocial desires may elicit more positive responses than punishments motivated by antisocial desires. We tested both motives that were external to the punisher (the behavior that elicited the punishment) as well as internal motives (the desire to harm versus rehabilitate transgressors). The main result is that we found negative social relationships among punishers, targets, and observers. Both children and adults preferred punishers who inflicted punishment for behaviors that violated (versus did not violate) norms, preferred targets of punishment who had not (versus had) violated norms, and expected punishers and targets to dislike each other. External motives, but not internal motives, consistently influenced participants' own social preferences. In contrast, neither external nor internal motives consistently shaped participants' inferences about social relationships between punishers and their targets. Our work contributes to social cognitive development by clarifying how motives shape children's and adults' understanding of social relationships.

Keywords: morality, punishment, social cognitive development, social relationships

Children's and Adults' Understanding of How Punishment Shapes Social Relationships

People prototypically punish in response to morally negative behaviors (e.g., theft or unfairness). Adults, as well as children as young as three years of age, punish those who transgress (Fehr & Fischbacher, 2004; Henrich et al., 2006; Marshall et al., 2021; McAuliffe et al., 2015; Yudkin et al., 2020). However, the eliciting behavior need not be a transgression. For instance, people can punish others for behaving pro-socially (e.g., sharing more resources with others than the punisher deems appropriate, Herrmann et al., 2008). People can also punish neutral acts. For example, in May 2024, Georgia governor Brian Kemp signed a bill prohibiting individuals or groups from paying bail for more than three people per year, presumably a neutral act because paying bail for a small number of people remained legal (Alfonseca, 2024). Additional laws prohibit wearing masks during meetings, sitting on public sidewalks, and selling or giving away silly string (Simon, 2018). Thus, scholars define punishment as an act that harms or imposes costs on another person (Raihani & Bshary, 2019; Marshall & McAuliffe, 2022) and not in reference to the behavior that elicited the punishment.

Although children and adults actively punish even at a personal cost (e.g., Fehr & Fischbacher, 2004; Marshall et al., 2021; McAuliffe et al., 2015; Yudkin et al., 2020), they perceive punishment as aggressive and often disapprove of punishment (Eriksson et al., 2016; Gordon et al., 2014; Lee & Warneken, 2020). These prior studies suggest that both school-aged children and adults may hold mixed attitudes towards punishment. That is, they are willing to mete out punishment themselves, but show negative attitudes towards other people's enactment of punishment (e.g., Lee & Solomon, 2024). The current work tested how punishment shapes

social relationships among punishers, targets and observers in adults as well as children. In addition, we asked how motives underlying motives might shape these social relationships.

Motives Underlying Punishment

Why do adults and children show mixed attitudes towards other people's punishment? One potential reason is that attitudes towards punishment might differ depending on perceived motives underlying that punishment. For example, people might show more positive attitudes towards punishment when the punishers harm transgressors (implying prosocial motives) than when they harm people who have not transgressed (implying antisocial motives). Similarly, observers might judge others' punishment positively when they learn that the punisher punished to teach transgressors a lesson (implying relatively prosocial motives) than when they punished to harm transgressors.

Broadly, at least two major types of motives—external and internal—underlie punishment. We define external motives as those elicited by factors outside of a punisher's control and internal motives as those residing within the punisher. One example of external motives is whether or not the behavior that elicited the punishment was a transgression. Adults are more driven to enact punishment when a transgression occurred than when it did not (Fehr & Fischbacher, 2004; Henrich et al., 2006; Nelissen & Zeelenberg, 2009; Yamagishi et al., 2017). An example of internal motives is whether a punisher has retributivist goals (e.g., harming transgressors because they have harmed others) or consequentialist goals (e.g., teaching transgressors a lesson; Carlsmith et al., 2002). According to the consequentialist perspective, punishment is justifiable because it achieves a valuable social goal (typically reduction of future transgressions; Hoskins & Duff, 2024). In contrast, from a retributivist perspective, punishment is justifiable because transgressors *deserve* punishment in proportion to their transgression

(Hoskins & Duff, 2024). In this case, "punishment is an end in itself and needs no further justifications" (Carlsmith et al., 2002, p. 284). Prior work has shown that both retributivist and consequentialist motives drive punishment among adults (Carlsmith et al., 2002; Crockett et al., 2014).

Social Relationships and Punishment

While prior work focused on how external and internal motivations might drive people's own punishment behavior (e.g., Carlsmith et al., 2002; Crockett et al., 2014; Fehr & Fischbacher, 2004), it remains less clear how these motives shape (a) how people relate to punishers and those who receive their punishment (hereafter, targets) and (b) how people expect punishers and their targets to relate to each other. It is important to study these questions because punishment has social consequences. It is often an aggressive and destructive behavior directed towards another person that could potentially harm social relationships (Petersen et al., 2010). For instance, adults who receive punishment often retaliate against their punishers (Denant-Boemont et al., 2007; Fehl et al., 2012; Zheng & Nie, 2013), potentially creating a vicious circle in their relationships. Yet, social relationships among punishers, targets, and observers might differ depending on motives.

In adults, external and internal motives shape observers' preferences for punishers differently. For example, adults judge people who punish transgressions more positively than those who do not punish transgressions (Barclay, 2006). Similarly, in another study (Raihani & Bshary, 2015), adults were more likely to reward those who punished transgressions than those who did not punish transgressions. In contrast, they were more likely to reward those who did not punish people who had not transgressed than those who did punish non-transgressions. These results show that adults base their social preferences on the punisher's external motives. Internal

motives also appear to shape their preferences. Specifically, adults are more likely to trust a person who endorses consequentialist reasons for punishment compared to another person who endorses retributivist motives for punishment (Dhaliwal et al., 2022). Overall, these studies have shown that punishers' external and internal motives shape how people relate to them.

The Importance of a Developmental Approach

While the studies discussed above provide important insight into how different motives shape adults' social preferences, less is known about children's inferences. It is important to study the influence of punishment motives on children's social preferences because children often observe or receive punishment from adults such as parents or teachers. If children focus exclusively on the aggressive nature of punishment without considering punishers' motives, they may perceive punishment and punishers negatively regardless of its underlying motives. However, if children understand that punishment might have different underlying motives, then they may view punishment with pro-social motives more positively than punishment with anti-social motives. Understanding the inferences children draw about punishment can also address broader questions about children's social cognition. For instance, if children judge punishment differently depending on its underlying motives, they may form stronger social bonds with people who emphasize pro-social motives driving their punishment decisions.

Like adults, children enact punishment for both external and internal reasons. Children are more likely to punish transgressions than non-transgressions, indicating that external motives drive their punishment behavior. For example, children as young as three years of age tattle on transgressors (Vaish et al., 2011; Yucel & Vaish, 2018). School-aged children take resources away from individuals who share unfairly (Arini et al., 2021; Gummerum & Chu, 2014; House et al., 2020; Lee & Warneken, 2022; McAuliffe et al., 2015). Further, children aged three years and

older punish those who destroy another person's belonging by keeping them from enjoying a fun activity (Marshall et al., 2021; Yudkin et al., 2020). In addition to these external reasons, internal factors—including both retributivist and consequentialist motives—also drive children's punishment (Marshall et al., 2021; Twardawski & Hilbig, 2020). For instance, in one study (Marshall et al., 2021), 5- to 7-year-old children were more likely to punish transgressors when their punishment could satisfy not only retributivist motives (e.g., harming transgressors) but also consequentialist motives (e.g., teaching transgressors a lesson). These findings suggest that internal motives as well as external motives drive children's punishment behavior.

While most developmental studies have focused on how external and internal motives drive children's punishment (e.g., Arini et al., 2021; Gummerum & Chu, 2014; House et al., 2020; Lee & Warneken, 2022; Marshall et al., 2021; McAuliffe et al., 2015; Twardawski & Hilbig, 2020; Vaish et al., 2011; Yucel & Vaish, 2018), relatively fewer studies have directly examined how these motives shape social preferences for others' punishment. In one study (Lee & Warneken, 2020), with increasing age, 5- to 9-year-olds increasingly preferred third parties who punished unfair sharers over neutral ones who neither witnessed nor intervened against the unfair sharing. Yet, children did not *always* prefer punishers. Specifically, in this work, children aged 7 years and older preferred third parties who punished unfair sharers rather than those who punished fair sharers. Further, when compensation was an available intervention, children preferred third parties who compensated the victim's loss over those who punished unfair sharers (Lee & Warneken, 2020). Therefore, children do not show strong positivity toward punishers when no transgression occurs or when an alternative intervention, such as compensation, is available. These findings suggest that children's preference for punishers could vary depending on motives and contexts.

Taken together, prior studies hint that external motives (i.e., whether or not a transgression has occurred) may guide children's and adults' punishment behavior and their social preference for punishers. These findings generally support the idea of developmental continuity, in which the consideration of external motives underlying punishment is in place by middle childhood and remains stable into adulthood. However, prior studies have not focused on investigating children's perceptions of social relationships involving punishment, and this question formed a central component of the current work. By examining this topic, the current research aimed to provide a more complete understanding of punishment's social consequences, including how these consequences may differ depending on the motives underlying the punishment. It is particularly important to examine whether children infer (dis)liking between punishers and their targets differently depending on external versus internal punishment motives. In everyday life, children often observe or receive punishment from parents or teachers. Even though the punitive act is the same, its underlying motive could differ, and different punishment motives might imply different social relationships. For example, in some cases, adults may punish children in order to teach children a lesson. In other cases, however, their punishment might seek to inflict suffering on children or to alter a behavior that many might not see as a transgression. The current work asked how children interpret social relationships between the punisher and the target across these varying types of motives.

One possibility is that children would infer more positive social relationships between punishers and targets when an external reason that they perceive as legitimate, such as the occurrence of a transgression, motivates the punishment. Children prefer people who punish transgressors rather than non-transgressors (Lee & Warneken, 2020), and they may expect people who are receiving punishment to be similar to themselves. That is, they may conclude

that targets of punishment would have more positive responses to people who punish them when they have committed a transgression than when they have not. Further, children might be more likely to infer positive social relationships between punishers and targets when punishers punish with consequentialist rather than retributivist motives because consequentialist motives might imply more prosociality (Dhaliwal et al., 2022) and prosociality underlies positive social relationships (Twenge et al., 2007).

To learn how inferences about social relationships change with age, we compared data from 5- to 10-year-olds and adults. Because both children and adults consider external motives in their own punishment behavior and social preferences for punishers (Arini et al., 2021; Barclay, 2006; Fehr & Fischbacher, 2004; Gummerum & Chu, 2014; Henrich et al., 2006; House et al., 2020; Lee & Warneken, 2020, 2022; Marshall et al., 2021; McAuliffe et al., 2015; Nelissen & Zeelenberg, 2009; Raihani & Bshary, 2015; Yamagishi et al., 2017; Yudkin et al., 2020), participants in both age groups may respond similarly when judging punishments driven by *external* factors. However, children may show less sensitivity than adults to *internal* factors. Within the developmental window we sampled (i.e., between 5 and 10 years old), children's moral judgments become increasingly sensitive to intent. That is, younger children tend to base their moral evaluations on outcomes—judging, for example, that it is worse to commit a relatively severe transgression accidentally than to commit a relatively minor transgression purposefully—whereas older children show more lenient judgments of accidental than purposeful transgressions (e.g., Cushman et al., 2013; Killen et al., 2011). Because intent is an internal state, it is possible that with increasing age, internal motivations for punishment would play an increasingly prominent role in participants' responses. Further, elementary schoolers—like adults—are more likely to punish transgressions than non-transgressions (i.e., their own

behavior shows sensitivity to external motivations; Gummerum & Chu, 2014; House et al., 2020; McAuliffe et al., 2015). Elementary-schoolers also prefer those who punish unfair individuals over those who punish fair individuals (Lee & Warneken, 2020) and punish other people based on consequentialist as well as retributivist motives (Marshall et al., 2021; Twardawski & Hilbig, 2020). Thus, prior work has shown that participants in our age range show sensitivity to both external and internal motives when making their own punishment decisions. We asked about the extent to which these types of motives shape inferences about *others'* punitive behaviors.

Overview of Current Research

The current work examined understanding of social preferences among punishers, targets and third-party observers (i.e., participants) by testing 5- to 10-year-old children (Study 1) and adults (Study 2). We assessed participants' own preferences for punishers and targets. Importantly, we also measured participants' perceptions about the extent to which punishers and targets liked each other to examine if participants expect bidirectional liking or disliking between punishers and targets, which could reveal potential boundary conditions of the effect of motive on social inferences. For instance, people might expect punishers with consequentialist motives to like their targets but might not expect the targets to like the punishers back irrespective of the punishers' motives. By probing multiple aspects of social inferences about punishers and targets, the present research aimed to understand the extent which external and internal punishment motives might shape children's and adults' inferences about social relationships.

Study 1

Study 1 investigated if and how punishment motives shape children's preferences and their inferences about social relationships. Five to 10-year-old children heard about a person who had a consequentialist or retributivist motive for punishing a target. Half of the punishments

occurred in response to a transgression, while the remaining punishments occurred in response to a non-transgression. Thus, Study 1 used a 2 (External Motive: Transgression versus Non-transgression) X 2 (Internal Motive: Retributivist versus Consequentialist) within-subject experiment design to test the effects of external and internal motives underlying punishment.

Participants

Our final sample included 128 5- to 10-year-old children (M_{age}=7.51 years, SD_{age}=1.66 years; n_{age5}=19, n_{age6}=21, n_{age7}=25, n_{age8}=22, n_{age9}=21, n_{age10}=20; 66 male, 59 female, 2 non-binary, 1 preferred not to say). Parents identified their child as White (57%), Asian (15%), Multiracial (14%), Black (6%), Native American or Pacific Islander (<1%) or some other race (<1%); the remaining children's parents did not report their race. Additionally, 9% of parents identified their child as Hispanic or Latinx. An a priori power analysis indicated that this sample size was sufficient to detect a medium-sized effect (d=.5) with an alpha of .05 and 80% power (see our pre-registration at https://aspredicted.org/VT2_1HC).

We recruited participants from a US-based laboratory database. We invited families to the current online research by sending a link, and anyone who landed on the page with a child in our age range was eligible to start the testing session. After we verified webcam recordings from testing sessions, parents of the child participants received a $5 Amazon gift card. Based on pre-registered criteria, we excluded 24 additional children because their parents did not provide a recording indicating their consent to participate (9), the participant completed the same study before (7), they were inattentive (4), they stopped the study before the test phase (2), they were out of our age range (1), or a sibling interfered with the study (1).[1]

[1] In this and the subsequent study, the patterns of results were similar even after including participants who were excluded from the main analyses. We did not analyze the data from participants who were excluded for no recording of consent because they did not provide permission for us to do so.

Experimental Design and Procedure

We used unmoderated remote research methods that have been established for data collection with children (Leshin et al., 2021; Rhodes et al., 2020). To begin the study, parents selected the study link, which allowed them to activate their webcam on their laptop or desktop computer. At the beginning of the study, parents provided verbal informed consent and children provided verbal assent. To confirm that participants could see and hear the stimuli presented on their screen, we played a sound of an animal and asked participants to choose an animal picture that matched the sound they heard. This procedure confirmed that all participants were able to see and hear the stimuli. Importantly, we instructed parents not to interfere during the study. After the instructions, children began the study, which consisted of a series of short videos and questions presented via Qualtrics.

Each child watched a total of four trials with a 2 (External motive: Transgression versus Non-transgression) X 2 (Internal Motive: Retributivist versus Consequentialist) within-subject experiment design. Each trial began with either a transgression or a non-transgression. The Transgression trial involved one character (i.e., target) destroying another character's belonging (e.g., ripping apart a drawing). The Non-transgression trial began with the target looking at another person's belonging, instead of destroying the belonging. Subsequently, children learned about a punisher with a retributivist motive or a consequentialist motive. Prior work has shown that elementary school-aged children are sensitive to different punishment motives and distinguish consequentialist motives from retributivist motives (Marshall et al., 2021, 2022). In a version with a retributivist motive, children heard that the punisher wanted the target to *pay for her behavior*. The punisher took the target's belonging (e.g., crayons) and told the target, e.g., "I'm going to lock your crayons in a lockbox so you *pay for what you did* to [victim's name].

You should *feel bad* because of this." Based on prior research with similar punishment (e.g., Marshall et al., 2021), we used locking the target's belonging in a lockbox as punishment. The version with a consequentialist motive was identical to the retributivist motive version except that the punisher held a consequentialist motive. For example, children heard that the punisher wanted the target to *learn from her behavior*. The punisher took the target's belonging and told the target, e.g., "I'm going to lock your crayons in a lockbox so you *stop doing this* to [victim's name]. You should *learn a lesson* because of this." At no point in the story did children hear the word "target", "victim" or "punisher" (see Materials for sample script at https://osf.io/gf7bn/?view_only=05903f5be0a34a0a8b5f4cb1289ff317). All characters in the stories looked like children.

At the end of each trial, participants answered four test questions. To assess children's own social preference, in a *participant-to-punisher* question, children answered the following questions: "Do you like [punisher's name] or not like [punisher's name]? Do you [like/not like] [punisher's name] a lot, kind of, or a little?" In a *participant-to-target* question, children answered the following questions: "Do you like [target's name] or not like [target's name]? Do you [like/not like] [target's name] a lot, kind of, or a little?" To assess children's inferences on social relationships, in a *punisher-to-target* question, children answered the following questions: "Does [punisher's name] like [target's name] or not like [target's name]? Does [punisher's name] [like/not like] [target's name] a lot, kind of, or a little?" In a *target-to-punisher* question, children answered the following questions: "Does [target's name] like [punisher's name] or not like [punisher's name]? Does [target's name] [like/not like] [punisher's name] a lot, kind of, or a little?" Across all four test questions, we coded children's responses on a scale ranging from 0 ("not like a lot") to 5 ("like a lot").

In addition, we asked several comprehension check questions. The results from these questions confirmed that a majority of children correctly identified the targets (91%), victims (91%), and punishers (97%). In a minority of trials in which children answered incorrectly, we provided corrective feedback by playing a video with a correct answer. In line with our pre-registration, we did not exclude children based on incorrect initial answers to the comprehension check questions. We also asked children whether it is okay to destroy (transgression trials) or look at (non-transgression trials) another person's belonging. Children reported that while it is okay to look at another person's belonging (88%), it is not okay to destroy another person's belonging (97%). This finding shows that most children considered the transgression scenario, but not non-transgression scenario, as unacceptable.

In this and the subsequent study, we counterbalanced the punishers and the type of transgressions across participants. For instance, a character named Daisy was a punisher in the Transgression-Retributivist scenario for some participants and in the Non-transgression-Consequentialist scenario for other participants. We counterbalanced four scenarios across trials: drawing (ripping apart or looking at another person's drawing), sandcastle (knocking down or looking at another person's sandcastle), teddy bear (ripping apart or looking at another person's teddy bear) and book (ripping apart or looking at another person's book). Additionally, some participants watched each scenario as a transgression (i.e., ripping apart another person's drawing), while other participants watched that same scenario as a non-transgression (i.e., taking a look at another person's drawing). We also coded our survey to randomize the order of the four scenarios and test questions, with one block consisting of liking questions between punishers and targets and another block consisting of children's own liking for punishers and targets. We randomized the order of the two blocks as well as the order of questions within each block.

Data Coding and Analyses

Qualtrics recorded children's responses. We conducted all statistical analyses with R statistical software (R version 4.1.2; R Core Team, 2021).

As pre-registered, we analyzed the four dependent measures (responses to participant-to-punisher, participant-to-target, punisher-to-target, and target-to-punisher questions) using Linear Mixed Models (LMM) with the R package "lme4" (Bates et al., 2015). Our full models included pre-registered predictors of interest such as children's age in months as a continuous variable, external motive (transgression versus non-transgression), internal motive (retributivist versus consequentialist), and interactions among these variables as fixed effects, plus subject identity (ID) as a random effect to account for the within-participant nature of the data. We first compared the full model with a null model that included only subject ID as a random intercept. If the full model provided a significantly better fit to the data than the null model, we then conducted hypothesis-driven tests to examine the role of individual predictors by sequentially dropping them from the full model and assessed changes in model fit using Likelihood Ratio Tests (LRTs). All the analyses in Studies 1 and 2 were pre-registered unless mentioned otherwise.

Results

Participant-to-Punisher

The full model provided a significantly better fit to the data than the null model, $\chi^2(7)=296.12$, $p<.001$. We found a significant interaction between external motive and age, $\chi^2(1)=11.04$, $p<.001$, which qualified a main effect of age, $b=-0.02$, $SE=0.01$, $p<.001$. The interaction (see Figure 1A) occurred because with increasing age, children reported less liking of punishers in non-transgressions scenarios, $\chi^2(1)=15.61$, $b=-0.02$, $SE=0.01$, $p<.001$, while they

reported relatively high liking of punishers across ages in transgression scenarios, χ^2 (1)<1, *b*=-0.01, *SE*=0.01, *p*=.337. Other main effects and interactions did not reach significance, *p*s>.438.

Participant-to-Target

Our full model provided a significantly better fit to the data than the null model, χ^2 (7)=324.64, *p*<.001. We found a significant interaction between external motive and age, χ^2 (1)=5.39, *p*=.020, which qualified a main effect of external motive, *b*=-1.24, *SE*=0.58, *p*=.033. The interaction (see Figure 1B) occurred because the effect of age on children's liking of targets varied by external motive in opposite directions: With age, children tended to like targets in the non-transgression scenario more, χ^2 (1)=1.49, *b*=0.01, *SE*=0.01, *p*=.223, while disliking targets in the transgression scenario more, χ^2 (1)=2.44, *b*=-0.01, *SE*=0.004, *p*=.119. In other words, the difference in children's liking between non-transgression and transgression became larger with age. Other main effects and interactions did not reach significance, *p*s>.133.

Punisher-to-Target

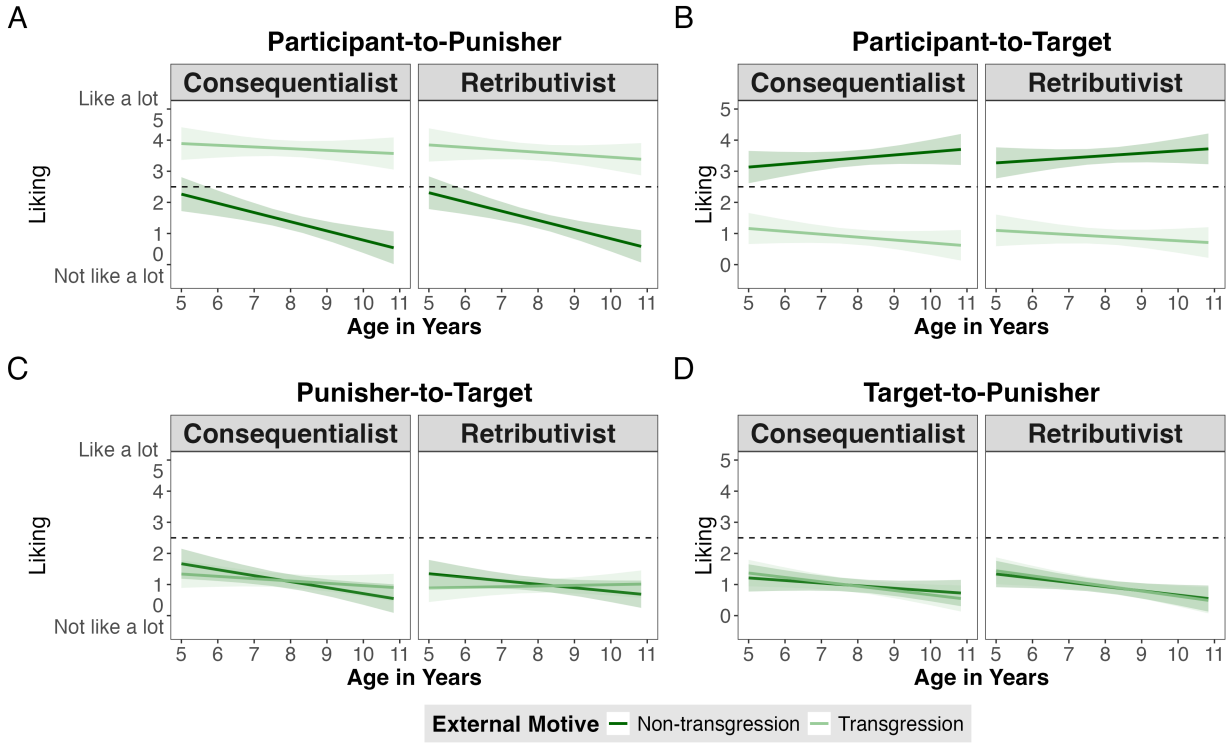
In this measure, we first examined whether children's age in months, external motive, and internal motive shaped children's inference on punishers' liking for targets. However, the full model did not provide a significantly better fit to the data than the null model, χ^2 (7)=13.24, *p*=.067. This result suggests that age, external motive, and internal motive did not explain the data significantly better above and beyond the random intercept. Although not pre-registered, to understand children's responses, we ran a *t*-test to examine if children's ratings differed significantly from the neutral valence of 2.5. Children overall expected punishers to dislike targets (i.e., children's responses fell significantly below the neutral rating; *M*=1.04, *SD*=1.25, *t*(486)=-25.76, *p*<.001; see Figure 1C).

Target-to-Punisher

We examined whether children's age in months, external motive, and internal motive shaped children's inference regarding targets' liking for punishers. The full model with the predictors illustrated above did not provide a significantly better fit to the data compared to the null model with only a random intercept, χ^2 (7)=8.42, p=.297. Although not pre-registered, we ran a *t*-test to examine if children's ratings differed significantly from the neutral valence of 2.5. Overall, children expected targets to dislike punishers (i.e., children's responses fell significantly below the neutral rating; M=0.96, SD=1.20, t(499)=-28.65, p<.001; see Figure 1D).

Figure 1

Children's Liking Evaluations by Age and Motive With 95% CIs in Study 1



Note. The dashed line indicates a neutral valence of 2.5. All estimations were based on the full models. (A) Estimated responses to the degree to which children liked punishers by age and motive. (B) Estimated responses to the degree to which children liked targets by age and motive. (C) Estimated responses to the degree to which punishers like targets by age and motive. (D) Estimated responses to the degree to which targets like punishers by age and motive.

Discussion

The goal of Study 1 was to examine children's social preferences and their inferences about social relationships based on external and internal punishment motives. With increasing age, children were more likely to prefer individuals who punish when a transgression occurred (versus when a transgression did not occur). They also liked targets who received punishment when a transgression did not occur more than those who received punishment when a transgression occurred, and this differentiation became more pronounced with age. However, neither children's preference for punishers nor that for targets differed by internal motives (retributivist versus consequentialist). Children as young as five years of age already demonstrated their own preference for punishers and targets based on external motives, but not on internal motives. These findings add to the literature in social cognitive development by showing that in a context involving punishment, children consider external motives, but not internal motives, which might be less observable and require more advanced social-cognitive skills.

We also found that children across ages expected both punishers and targets to dislike one another regardless of external or internal motives. It is noteworthy that children took external motives into account in their own preference for punishers and targets, but not in inferring social relationships between punishers and targets. This result implies that even when the punisher is motivated by relatively prosocial goals, such as consequentialist motives to teach a lesson, children primarily perceive negativity in social relationships between punishers and targets. This finding highlights potential negative consequences of punishment, which have received relatively less attention compared to its positive consequences (e.g., maintaining social norms and preventing future transgressions; Raihani & Bshary, 2019).

Study 2

Study 2 investigated the degree to which adults' responses are similar to those from 5- to 10-year-olds by using the same experimental design and stimuli from Study 1. By testing adults, Study 2 aimed to adjudicate between the developmental continuity versus discontinuity hypothesis. Specifically, we examined if adults share similar sensitivity to external motives and insensitivity to internal motives that we observed among children.

Method

Participants

Our final sample included 159 adults (M_{age}=44.28 years, SD_{age}=14.57 years; 77 female, 76 male, 5 non-binary, 1 preferred not to say). Participants self-identified as White (78%), Black (8%), Asian or Asian American (6%), Multiracial (6%), Native American or Pacific Islander (<1%), or some other race (2%). Additionally, 6% of adults self-identified as Hispanic or Latinx.

We recruited participants online via Prolific. We configured our study so that only United States residents whose approval rating was at least 98% with at least 100 submissions could participate. Additionally, only those who passed two tasks designed to screen out bots could participate. Participants received $2 if they summarized one of the vignettes correctly; otherwise, they received no payment. Based on pre-registered criteria, we excluded data from 21 additional participants who failed to provide a correct summary of one of the vignettes in the study.

As in Study 1, we pre-registered to collect data from n=128 participants (see our pre-registration at https://aspredicted.org/J43_XKL). However, to account for attrition (expected to be around 20%), we intentionally oversampled. Because the actual attrition rate was lower than expected, we retained all usable data as long as responses did not meet the pre-registered exclusion criteria.

Experimental Design and Procedure

The experimental stimuli and design in Study 2 were identical to those in Study 1 except that adults read the vignettes presented by Qualtrics. Adult participants read the same stories that we presented to children (all characters looked like children) to allow for a straightforward comparison between children and adults.

As in Study 1, the results from comprehension check questions confirmed that almost all adult participants correctly identified the targets (99.8%), the victims (99.8%), and the punishers (99.5%). Adults reported that it was okay to look at another person's belonging (99.4%) and not okay to destroy another person's belonging (99.7%).

Results

Participant-to-Punisher

The full model provided a significantly better fit to the data than the null model, χ^2 (3)=578.41, p<.001. We found a significant main effect of external motive, χ^2 (1)=575.86, b=2.93, SE=0.09, p<.001, revealing that adults liked punishers in the transgression scenario (M=3.91, SD=1.31) more than those in the non-transgression scenario (M=0.97, SD=1.27). These results are consistent with those from children in Study 1. Other main effects and interactions did not reach significance, ps>.151.

Additionally, we examined whether age group (children from Study 1 versus adults), external motive, and internal motive shaped participants' liking for punishers. The full model provided a significantly better fit to the data than the null model, χ^2 (7)=846.54, p<.001. We found a significant interaction between age group and external motive, χ^2 (1)=22.44, p<.001, which qualified significant main effects of external motive, b=2.93, SE=0.09, p<.001, and age group, b=0.44, SE=0.13, p<.001. The interaction occurred because in the non-transgression

scenario, adults (M=0.97, SD=1.27) disliked punishers more than did children (M=1.41, SD=1.54), χ^2 (1)=8.39, b=0.44, SE=0.15, p=.004 (see Figure 2A). In contrast, in the transgression scenario, adults' (M=3.91, SD=1.31) and children's (M=3.67, SD=1.45) liking of the punisher did not differ significantly from each other, χ^2 (1)=2.48, b=-0.23, SE=0.15, p=.117. Other main effects and interactions did not reach significance, all ps>.230.

Participant-to-Target

Our full model provided a significantly better fit to the data than the null model, χ^2 (3)=651.30, p<.001. We found a significant main effect of external motive, χ^2 (1)=650.74, b=-3.20, SE=0.09, p<.001. That is, adults liked targets in the non-transgression scenario (M=3.98, SD=1.23) more than those in the transgression scenario (M=0.78, SD=1.19), which is similar to the finding from children in Study 1. Other main effects and interactions did not reach significance, ps>.535.

We also examined whether age group (children from Study 1 versus adults from Study 2), external motive, and internal motive shaped participants' liking for targets. The full model provided a significantly better fit to the data than the null model, χ^2 (7)=965.46, p<.001. We found a significant interaction between age group and external motive, χ^2 (1)=18.69, p<.001, which qualified significant main effects of external motive, b=-3.20, SE=0.10, p<.001, and age group, b=-0.57, SE=0.14, p<.001. The interaction occurred because in the non-transgression scenario, adults (M=3.98, SD=1.23) liked targets more than did children (M=3.47, SD=1.53), χ^2 (1)=13.77, b=-0.52, SE=0.14, p<.001 (see Figure 2B). In contrast, in the transgression scenario, adults' (M=0.78, SD=1.19) and children's (M=0.90, SD=1.23) liking of targets did not differ significantly from each other, χ^2 (1)<1, b=0.12, SE=0.12, p=.325. Other main effects and interactions did not reach significance, all ps>.501.

Punisher-to-Target

In this measure, we examined whether external motive and internal motive shaped adults' inference on punishers' liking for targets. The full model provided a significantly better fit to the data than the null model, χ^2 (3)=15.64, p=.001. Further analysis indicated a significant main effect of external motive, χ^2 (1)=12.08, b=0.24, SE=0.07, p<.001. That is, adults inferred that punishers in the transgression scenario (M=0.98, SD=1.21) would like targets more compared to punishers in the non-transgression scenario (M=0.75, SD=1.15). However, although adults tended to expect punishers with consequentialist motives (M=0.93, SD=1.20) to like targets more compared to punishers with retributivist motives (M=0.80, SD=1.17), the main effect of internal motive did not reach significance, b=-0.13, SE=0.07, p=.056. The interaction between external motive and internal motive did not reach significance, χ^2 (1)<1, p=.963.

Additionally, we examined whether age group (children from Study 1 versus adults from Study 2), external motive, and internal motive shaped participants' inference on punishers' liking for targets. The full model provided a significantly better fit to the data than the null model, χ^2 (7)=19.65, p=.006. We found a significant interaction effect between external motive and age group, χ^2 (1)=5.65, p=.018, which qualified significant main effects of external motive, b=0.24, SE=0.07, p<.001, and age group, b=0.31, SE=0.12, p=.013. The interaction occurred because in the non-transgression scenario, adults (M=0.75, SD=1.15) were more likely to predict that punishers disliked targets than were children (M=1.04, SD=1.25), χ^2 (1)=5.24, b=0.29, SE=0.13, p=.022 (see Figure 2C). In contrast, in the transgression scenario, adults' (M=0.98, SD=1.21) and children's (M=1.04, SD=1.25) predictions did not differ significantly from each other, χ^2 (1)<1, b=0.06, SE=0.13, p=.629.

We also found a significant main effect of internal motive, b=-0.13, SE=0.05, p=.018, showing that overall, participants perceived that punishers with a consequentialist motive (M=1.00, SD=1.23) liked targets more compared to punishers with a retributivist motive (M=0.88, SD=1.21). This result is consistent with our prediction that participants would infer more positive social relationships following punishment with consequentialist, versus retributivist, motives. Although both children and adults showed a tendency to infer that the punishers with consequentialist motives would like their targets more than those with retributivist motives, this effect did not reach significance within either age group. We speculate that the sample size in each age group (ns=128 children and 159 adults) might have been too small to detect the effect of internal motives when they were analyzed separately. When the data from children and adults were combined, the main effect of internal motives was small (d=0.10). A post-hoc power analysis indicated that we would need at least 787 participants to detect an effect of this size using standard parameters (power=.80, alpha=.05). Because our sample lacked this statistical power and the effect of internal motives did not seem robust, interpreting this finding requires caution. Other main effects and interactions did not reach significance, ps>.659.

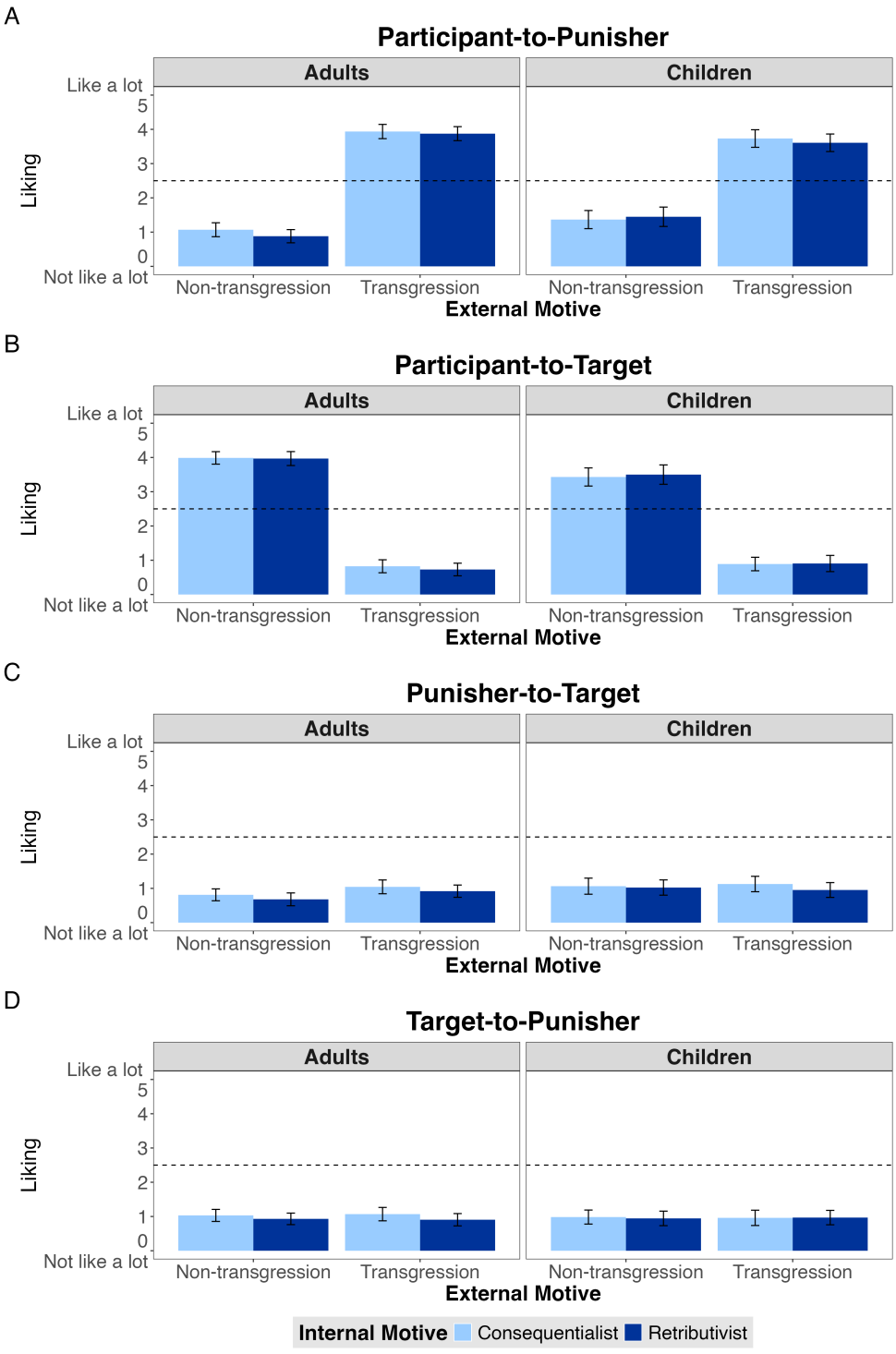
Target-to-Punisher

We examined whether external motive and internal motive shaped adults' inferences regarding targets' liking for punishers. However, the full model with the predictors did not provide a significantly better fit to the data compared to the null model with only a random intercept, χ^2 (3)=4.11, p=.250. This finding indicates that external motive and internal motive did not explain the data significantly better above and beyond the random intercept. Like children in Study 1, adults overall predicted that targets would dislike punishers (M=0.98, SD=1.15, t(635)=-33.28, p<.001; see Figure 2D).

Additionally, we examined whether age group (children from Study 1 versus adults from Study 2), external motive, and internal motive shaped participants' inferences regarding targets' liking for punishers. The full model did not provide better fit to the data than the null model, $\chi^2(7)=4.08$, $p=.770$, suggesting that children and adults did not differ from each other when predicting that targets would dislike punishers across motives.

Figure 2

Average Liking Evaluations by Age Group, External Motive and Internal Motive With 95% CIs in Studies 1 and 2



Note. The dashed line indicates a neutral valence of 2.5. (A) Average responses to the degree to which participants liked punishers by age group, internal motive and external motive. (B) Average responses to the degree to which participants liked targets by age group, internal motive and external motive. (C) Average responses to the degree to which punishers like targets by age

group, internal motive and external motive. (D) Estimated responses to the degree to which targets like punishers by age group, internal motive and external motive.

Discussion

The goal of Study 2 was to examine adults' social judgments based on internal and external punishment motives and to compare these judgments with those we observed among children in Study 1. Overall, our results are consistent with the developmental continuity hypothesis. Specifically, both children and adults (a) preferred people who punished transgressions rather than non-transgressions, (b) preferred targets who did not transgress over those who transgressed, and (c) expected targets to dislike punishers irrespective of internal or external motives. These results suggest that from as early as 5 years old, children show adult-like responses to interpreting social relationships involved in punishment, even though children might have relatively limited social experiences involving punishment compared to adults.

General Discussion

The present work examined 5- to 10-year-old children's and adults' social judgments in scenarios involving different punishment motives. Specifically, we probed participants' own responses to punishers and their targets as well as the responses participants expected punishers and targets to have toward each other. Our findings support the developmental continuity hypothesis, which posits that children and adults do not differ in terms of their attitudes toward those involved in punishment and in how they perceive social relationships that punishment could affect. Three key findings emerged: first and most importantly, children and adults generally expected bidirectional disliking between the punisher and the target, indicating that they infer negative social relationships in punishment-related contexts. We also found that

preferences of both children and adults showed sensitivity to external motives. Whereas, neither children nor adults considered internal motives in a reliable manner.

Our work makes several contributions to developmental science and moral psychology. First, the current findings extend scientific understanding of moral cognition by revealing potential negative consequences of punishment, which have received relatively less attention compared to its positive consequences. Specifically, existing literature has focused primarily on the positive role of punishment in maintaining cooperation and preventing future harms (see Raihani & Bshary, 2019). However, our findings show that both children and adults generally infer mutual disliking between the punisher and the target even when punishment is driven by reasons that people commonly view as appropriate and prosocial, such as a transgression occurrence or a desire to help transgressors learn a lesson. This result highlights the possibility that punishment might introduce or strengthen negativity within social relationships.

Second, the current research contributes to knowledge regarding social cognitive development by showing that children and adults exhibit less strong sensitivity to internal motives compared to external motives. In their inferences about punishers' liking for targets, both adults and children tended to infer that punishers with consequentialist motives would like targets better than those with retributivist motives. However, the effect reached significance only when we combined the responses from adults and children, indicating that the effect may not be strong or robust compared to external motives and therefore requires caution in interpretation. Further, these nuanced results contrast with earlier studies suggesting that sensitivity to internal factors such as intent increases across age (e.g., Cushman et al., 2013; Killen et al., 2011). Our findings may imply that when participants are third-party observers of punishment, its aggressive

nature may override an appreciation of punishers' internal states, which are less observable than the punishment itself.

Third, our work has translational implications for children's experience of punishment in their daily lives. Children in the present work inferred bidirectional disliking between punishers and targets regardless of internal or external motives. This finding implies that children may perceive negative messages when they receive punishment even if the punisher's intention is prosocial (e.g., consequentialist motive to teach a lesson). Therefore, our result may inform educators and caregivers about how they could mitigate negative messages punishment sends to children; for instance, they may wish to explicitly affirm that they like the child and are trying to help the child learn.

While the current research extends scientific understanding of children's and adults' moral cognition, like all research programs, it cannot possibly answer all questions. One potential direction for future research is to investigate other motivations for punishment. We focused on two broad motives—external (transgression versus non-transgression) and internal (retributivist versus consequentialist)—to build on existing literature focusing on these goals (e.g., Carlsmith et al., 2002; Crockett et al., 2014; Fehr & Fischbacher, 2004; Henrich et al., 2006; Marshall et al., 2021; McAuliffe et al., 2015; Yudkin et al., 2020). However, these motives are neither mutually exclusive nor an exhaustive list of all possible motives underlying punishment. For example, people might use punishment to create equality or to compete against other people (see Raihani & Bshary, 2019, for a review). Participants might be more likely to infer positive social relationships between punishers and targets when punishers have consequentialist motives than when they have motives that we did not test here, such as competitive motives. Future research should address this possibility by including other motives.

While our work has shown that the effect of internal motives may be less reliable or strong than external motives, there may be more effective ways to convey internal motives to participants. The current work manipulated punishers' internal motives by using verbal communications ("I'm going to lock your crayons in a lockbox so you pay for what you did to [victim's name]. You should feel bad because of this" versus "I'm going to lock your crayons in a lockbox so you stop doing this to [victim's name]. You should learn a lesson because of this"), which built on prior studies that used verbal communications to convey different internal punishment motives (e.g., Dhaliwal et al., 2022; Marshall et al., 2022). However, this manipulation might not have been sufficient to shift participants' preference or their inferences about social relationships in the current context. Future research can manipulate internal motives differently by varying whether or not the punisher punishes the transgressor when they could learn about the punishment (e.g., Crockett et al., 2014; Marshall et al., 2021; Twardawski & Hilbig, 2021).

Another future direction involves investigating cross-cultural differences in perceptions of punishment motives. Our work tested participants in the United States, a country with strong norms about civic cooperation (Herrmann et al., 2008). This decision might have contributed to pronounced differentiation of punishment between transgressions and non-transgressions. Our sample allowed systematic comparisons between studies by building on the existing literature with samples from the United States (Lee & Warneken, 2020, 2022; Marshall et al., 2021; McAuliffe et al., 2015; Yudkin et al., 2020). However, because societies differ in the degree to which they punish transgressions and non-transgressions (Henrich et al., 2006; Herrmann et al., 2008), there may be cross-cultural variations in the degree to which people's perceptions are sensitive to punishment following transgressions versus non-transgressions. For instance, in a

society where people tend to punish prosocial acts (e.g., Herrmann et al., 2008), participants might not distinguish between individuals who punish transgressions versus non-transgressions. Further, while participants in the current context liked those who punished transgressors, this finding may not generalize across different contexts and cultures, given that there are variations in the degree to which people prefer peer punishment (e.g., Eriksson et al., 2017). Future research can address these possibilities.

Conclusion

The current work examined how children and adults develop and perceive social preferences in situations involving punishment. Generally, children and adults were sensitive to the external causes of punishment. However, children and adults did not show strong or consistent sensitivity to punishment's internal causes. Most importantly, our results also suggest that children as young as five years old show adult-like patterns in the way they consider motives when inferring others' social relationships as well as forming their own social preferences. Our research contributes to work on moral cognition and social cognitive development by revealing potential negative consequences of punishment in social relationships.

References

Alfonseca, K. (2024, May 1). Georgia governor signs controversial bail fund restrictions, expands cash bail. *ABC News*. https://abcnews.go.com/Politics/georgia-governor-signs-controversial-bail-fund-restrictions-expands/story?id=108744778

Arini, R. L., Wiggs, L., & Kenward, B. (2021). Moral duty and equalization concerns motivate children's third-party punishment. *Developmental Psychology*, *57*(8), 1325–1341. https://doi.org/10.1037/dev0001191

Barclay, P. (2006). Reputational benefits for altruistic punishment. *Evolution and Human Behavior*, *27*(5), 325–344. https://doi.org/10.1016/j.evolhumbehav.2006.01.003

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish?: Deterrence and just deserts as motives for punishment. *Journal of Personality and Social Psychology*, *83*(2), 284–299. https://doi.org/10.1037//0022-3514.83.2.284

Crockett, M. J., Özdemir, Y., & Fehr, E. (2014). The value of vengeance and the demand for deterrence. *Journal of Experimental Psychology: General*, *143*(6), 2279–2286. https://doi.org/10.1037/xge0000018

Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intend-based moral judgment. *Cognition*, *127*, 6–21.

Denant-Boemont, L., Masclet, D., & Noussair, C. N. (2007). Punishment, counterpunishment and sanction enforcement in a social dilemma experiment. *Economic Theory*, *33*(1), 145–167. https://doi.org/10.1007/s00199-007-0212-0

Dhaliwal, N. A., Skarlicki, D. P., Hoegg, J. A., & Daniels, M. A. (2022). Consequentialist motives for punishment signal trustworthiness. *Journal of Business Ethics*, *176*(3), 451–466. https://doi.org/10.1007/s10551-020-04664-5

Eriksson, K., Andersson, P. A., & Strimling, P. (2016). Moderators of the disapproval of peer punishment. *Group Processes & Intergroup Relations*, *19*(2), 152–168. https://doi.org/10.1177/1368430215583519

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

Fehl, K., Sommerfeld, R. D., Semmann, D., Krambeck, H. J., & Milinski, M. (2012). I dare you to punish me – vendettas in games of cooperation. *PLoS ONE, 7*, e45093. https://doi.org/10.1371/journal.pone.0045093

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

Gordon, D. S., Madden, J. R., & Lea, S. E. G. (2014). Both loved and feared: Third party punishers are viewed as formidable and likeable, but these reputational benefits may only be open to dominant individuals. *PLoS ONE*, *9*(10), 1–10. https://doi.org/10.1371/journal.pone.0110045

Gummerum, M., & Chu, M. T. (2014). Outcomes and intentions in children's, adolescents', and adults' second- and third-party punishment behavior. *Cognition*, *133*, 97–103. doi:10.1016/j.cognition.2014.06.001

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

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, *319*, 1362–1367.

Hoskins, Z. and Duff, A., "Legal Punishment", *The Stanford Encyclopedia of Philosophy* (Spring 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), forthcoming URL = <https://plato.stanford.edu/archives/spr2024/entries/legal-punishment/>.

House, B. R., Kanngiesser, P., Barrett, H. C., Yilmaz, S., Smith, A. M., Sebastian-Enesco, C., … 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*, 20192794. https://doi.org/10.1098/rspb.2019.2794

Killen, M., Lynn Mulvey, K., Richardson, C., Jampol, N., & Woodward, A. (2011). The accidental transgressor: Morally-relevant theory of mind. *Cognition*, *119*(2), 197–215. https://doi.org/10.1016/j.cognition.2011.01.006

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

Lee, Y., & Warneken, F. (2022). Does third-party punishment in children aim at equality? *Developmental Psychology*, *58*(5), 866–873. https://doi.org/10.1037/dev0001331

Lee, Y., Solomon, L. H. (2024). Not all punishment is equal: The effect of punishment severity on children's social evaluations. *Developmental Psychology*. 10.1037/dev0001845

Leshin, R. A., Leslie, S. J., & Rhodes, M. (2021). Does it matter how we speak about social kinds? A large, preregistered, online experimental study of how language shapes the development of essentialist beliefs. *Child Development*, *92*(4), e531–e547. https://doi.org/10.1111/cdev.13527

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

Marshall, J., Gollwitzer, A., & Bloom, P. (2022). Why do children and adults think other people punish? *Developmental Psychology*. https://doi.org/10.1037/dev0001378

Marshall, J., Yudkin, D. A., & Crockett, M. J. (2021). Children punish third parties to satisfy both consequentialist and retributive motives. *Nature Human Behaviour, 5*, 361-368. https://doi.org/10.1038/s41562-020-00975-9

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

Nelissen, R. M. a., & Zeelenberg, M. (2009). Moral emotions as determinants of third-party punishment: Anger, guilt, and the functions of altruistic sanctions. *Judgment and Decision Making*, *4*(7), 543–553.

Petersen, M.B., Sell, A., Tooby, J., & Cosmides, L. Evolutionary psychology and criminal Justice: A recalibrational theory of punishment and reconciliation. In: Høgh-Olesen, Henrik

(Ed.): Human morality and sociality: Evolutionary and comparative perspectives, Palgrave Macmillan, 2010.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raihani, N. J., & Bshary, R. (2015). Third-party punishers are rewarded, but third-party helpers even more so. *Evolution*, *69*(4), 993–1003. https://doi.org/10.1111/evo.12637

Raihani, N. J., & Bshary, R. (2019). Punishment: one tool, many uses. *Evolutionary Human Sciences*, *1*, e12, 1–26. https://doi.org/10.1017/ehs.2019.12

Rhodes, M., Rizzo, M. T., Foster-Hanson, E., Moty, K., Leshin, R. A., Wang, M., Benitez, J., & Ocampo, J. D. (2020). Advancing developmental science via unmoderated remote research with children. *Journal of Cognition and Development*, *21*(4), 477–493. https://doi.org/10.1080/15248372.2020.1797751

Simon, C. (2018, October 28). Weirdest laws passed in every state. *USA Today*. https://www.usatoday.com/list/news/nation-now/weirdest-laws-every-state/53ad0541-3518-4432-adc4-0fec193d389e/

Twardawski, M., & Hilbig, B. E. (2020). The motivational basis of third-party punishment in children. *PLoS ONE*, *15*(11 November), 1–15. https://doi.org/10.1371/journal.pone.0241919

Twenge, J. M., Baumeister, R. F., DeWall, C. N., Ciarocco, N. J., & Bartels, J. M. (2007). Social exclusion decreases prosocial behavior. *Journal of Personality and Social Psychology*, *92*(1), 56.

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

Yamagishi, T., Li, Y., Fermin, A. S. R., Kanai, R., Takagishi, H., Matsumoto, Y., Kiyonari, T., & Sakagami, M. (2017). Behavioural differences and neural substrates of altruistic and spiteful punishment. *Scientific Reports*, *7*(1), 1–8. https://doi.org/10.1038/s41598-017-15188-w

Yucel, M., & Vaish, A. (2018). Young children tattle to enforce moral norms. *Social Development*, *27*(4), 924–936. https://doi.org/10.1111/sode.12290

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

Zheng, X., & Nie, P. (2013). Effective punishment needs legitimacy. *Economic Record*, *89*(287), 522–544. https://doi.org/10.1111/1475-4932.12073