

An Earlier Role for Intent in Children's Partner Choice Versus Punishment

Justin W. Martin, Kyleigh Leddy, Liane Young, and Katherine McAuliffe

Department of Psychology and Neuroscience, Boston College

Among the many factors that influence our moral judgments, two are especially important: whether the person caused a bad outcome and whether they intended for it to happen. Notably, the weight accorded to these factors in adulthood varies by the type of judgment being made. For punishment decisions, intentions and outcomes carry relatively equal weight; for partner choice decisions (i.e., deciding whether or not to interact with someone again), intentions are weighted much more heavily. These behavioral differences in punishment and partner choice judgments may also reflect more fundamental differences in the cognitive processes supporting these decisions. Exploring how punishment and partner choice emerge in development provides important and unique insight into these processes as they emerge and mature. Here, we explore the developmental emergence of punishment and partner choice decisions in 4- to 9-year-old children. Given the importance of intentions for partner choice decisions—from both theoretical and empirical perspectives—we targeted the sensitivity of these two responses to others' intentions as well as outcomes caused. Our punishment results replicate past work: Young children are more focused on outcomes caused and become increasingly sensitive to intentions with age. In contrast, partner choice judgments exhibit sensitivity to intentions at an earlier age than punishment judgments, manifesting as earlier partner choice in cases of attempted violations. These results reveal distinct developmental trajectories for punishment and partner choice judgments, with implications for our understanding of the processes underlying these two responses as well as the development of moral judgment more broadly.

Keywords: intentions, moral development, partner choice, punishment

Supplemental materials: <https://doi.org/10.1037/xge0001093.supp>

When deciding whether to morally condemn someone for a bad outcome—when someone drops our favorite coffee mug or pushes us from behind at a concert—myriad factors may influence how we respond. Two factors of principle importance are whether they intended this outcome and whether they were causally responsible for it (Berg-Cross, 1975; Cushman et al., 2009, 2008; Martin & Cushman, 2015; Nobes et al., 2016; for a review see Martin & Cushman, 2016). The unique influence of these two factors has been studied under the headings of moral luck (Lench et al., 2014; Nagel, 1979; Williams, 1981; Young et al., 2010), hindsight bias (Alicke & Davis, 1989; Baron & Hershey,

1988; Tostain & Lebreuilly, 2008), and outcome bias (Gerstenberg et al., 2010; Gino et al., 2009, 2010; Mazzocco & Cherubini, 2010; Mazzocco et al., 2004). Although both intentions and outcomes caused carry weight in our judgment, this work has largely found that intentions matter more (Cushman, 2008; Cushman et al., 2009; Young et al., 2006, 2007, 2010). Although variability exists, almost all cultures studied exhibit some degree of sensitivity to others' intent independent of outcomes caused (Barrett et al., 2016; McNamara et al., 2019). This privileged sensitivity is also reflected in the fact that a relatively circumscribed and dedicated network of brain regions is responsible for processing information about others' intentions (Castelli et al., 2000; Fletcher et al., 1995; Saxe & Kanwisher, 2003; Young et al., 2007, 2010). And, some understanding of other's mental states exerts an influence on our sociomoral decisions from relatively early in life, at least by 6–7 years of age (Armsby, 1971; Cushman et al., 2013; Farnill, 1974; Miller et al., 2014; Nobes et al., 2016; Yuill & Perner, 1988) and possibly earlier (Hamlin, 2013; Hamlin et al., 2011, 2013; Kiley Hamlin et al., 2013; Li & Tomasello, 2018; Van de Vondervoort & Hamlin, 2018).

Intentions' Influence on Punishment and Partner Choice Decisions in Adulthood

An implicit assumption in much prior work has been that the weight bad intentions are afforded is the same regardless of the

This article was published Online First August 30, 2021.

Justin W. Martin  <https://orcid.org/0000-0001-6100-8581>

We are grateful to Richard Ahl, Dorsa Amir, Hannah Bolotin, Michael Bogese, Fiery Cushman, Masoud Jasbi, Minjae Kim, Kate Lindsey, Laurent Pretot, Michael Henry Tessler, and Kara Weisman for feedback on earlier versions of this article. This work was supported by the Boston College Virtue Project, the John Templeton Foundation, and the Canadian Institute for Advanced Research (Azrieli Global Scholar award). Data and analysis code for this work can be found at <https://osf.io/7539h/> (Martin et al., 2020).

Correspondence concerning this article should be addressed to Justin W. Martin, Department of Psychology and Neuroscience, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467, United States. Email: justinwmartin@gmail.com

type of response we employ—whether we respond physically, by yelling at the person, by ending our relationship with them, and so forth. However, recent work has called this assumption into question (Liddell & Kruschke, 2014; Martin & Cushman, 2015). These studies have distinguished between two categories of responses: punishment (e.g., responding physically or verbally) and partner choice (e.g., ending the relationship and seeking out a new partner), sometimes studied under the headings of partner control and partner choice (Bshary & Grutter, 2002; Fruteau et al., 2011; Martin & Cushman, 2015; Schino & Aureli, 2015). These studies find that intentions exert a greater influence when making a partner choice decision than when making a punishment decision. That is, when deciding whether to end a relationship with someone, we care more about how they intended to behave—regardless of the outcome of their behavior—than when we are deciding whether to punish them. From a certain perspective, these results are surprising—they demonstrate that we care more or less about others' mental states depending upon how we choose to respond to their bad behavior.

However, a more provocative interpretation of these results is that distinct psychological processes support punishment and partner choice responses, with the processes supporting mental state inference being more integrated into the processes supporting partner choice decisions. This interpretation accords well with theoretical perspectives on the functional purposes of partner choice and punishment. Partner choice captures the idea that individuals exist in a “biological market” and can leave or reject bad partners (Barclay, 2013, 2016). From this perspective, the functional purpose of partner choice—which includes rejecting bad partners as well as attracting good partners by advertising our desirable traits and maintaining beneficial partnerships—is to find and maintain the best partnerships (Barclay, 2013, 2016; Misyak et al., 2014; Warneken, 2018). Indications that someone will be a good partner in the future can be gleaned from their past behavior: someone who has cooperated previously is likely to cooperate in the future. Importantly, past behavior has (at least) two components—the outcome the person was intending to bring about (i.e., their intention) and the outcome that actually occurred. When predicting whether someone will be a good partner in the future, someone who stably cooperates and reliably delivers benefits, what they previously intended is likely to be a superior cue than what they happened to have caused. That is, accidents can happen, and what someone was trying to do in the past predicts what they will try to do in the future better than whether they succeeded in their goal. Indeed, past work demonstrates that we prefer those who are more willing to deliver benefits over than those who are more able to deliver benefits (Bliege Bird & Power, 2015; Eisenbruch & Roney, 2017; Hackel et al., 2015; Raihani & Barclay, 2016). Of course, there may be exceptions to the greater predictive power of intentions: The intentions of someone who consistently intends to be on time but is always late may need to be discounted in favor of the outcome they consistently produce. Nevertheless, as a general rule, understanding others' intentions is a critical aspect of partner choice and these decisions may especially recruit processing for understanding others' mental states. In contrast, whereas many functional explanations for punishment have been proposed—including changing others' behavior (Clutton-Brock & Parker, 1995; Martin & Cushman, 2016; Trivers, 1971), signaling traits to others (Barclay, 2006; Jordan et al., 2016), or enforcing norms

(Boyd & Richerson, 1992; Buckholtz & Marois, 2012; Fehr & Gächter, 2002; Henrich et al., 2010; Yamagishi, 1986)—none of these explanations suggests that punishment is about selecting partners. Although obvious, the fact that punishment does not serve the functional purpose of selecting partners indicates that partner choice and punishment responses serve different functional purposes to at least some extent. Critically, although some of the explanations for punishment do implicate mental state processing in punishment decisions, a functional perspective implicates these processes more strongly in partner choice decisions. In sum, behavioral evidence demonstrating a greater reliance on information about intentions when making a partner choice decision than a punishment decision is consistent with the idea that a crucial aspect of partner choice is understanding other's mental states, as a cue to how they will behave in the future.

The Development of Punishment and Partner Choice Decisions in Childhood

Of course, the idea that punishment and partner choice decisions are supported by distinct underlying processes is limited by the lack of work exploring these processes. One important way of pulling apart the processes that uniquely support punishment and partner choice is to focus on how these responses emerge in development. This approach has two principal benefits. First, developmental data provide unique insight into how differences in these responses observed in adulthood first emerge and mature. At minimum, these data yield insights into the origins of an interesting pattern that has been documented in adults. In addition, by examining how these responses emerge, we gain insight into the nature of these processes, in the same way that a full understanding of any organism comes from understanding both adult and juvenile forms of the species. Second, if punishment and partner choice are supported by unique processes in adulthood, owing to their distinct functional purposes, we might expect these processes to emerge differently in childhood. More specifically, given the importance of understanding others' intentions for making partner choice decisions, we might expect this ability to emerge first in partner choice decisions before this information is incorporated into punishment. Alternatively, it is possible that punishment and partner choice decisions emerge similarly in ontogeny and remain similar throughout childhood, diverging only at adulthood. This pattern would suggest that differences in punishment and partner choice stem more from socialization than any fundamental difference in these processes. Thus, developmental data have important relevance for our understanding of these responses in adulthood, by providing a critical test of the idea that partner choice and punishment are functionally distinct: If this argument is true, we would expect partner choice decisions to emerge distinctly from punishment, and in particular to integrate information about intentions earlier than punishment. Indeed, recent work hints at this possibility.

Much is known about the developmental emergence of punishment. One major finding of this effort is the well-studied “outcome-to-intent” shift (Cushman et al., 2013; Imamoglu, 1975; 2016; Nobes et al., 2009; Zelazo et al., 1996). Specifically, young children's punishment decisions are principally sensitive to outcomes caused and only minimally sensitive to others' intentions. For instance, in one study (Cushman et al., 2013); 4-year-olds

judged cases of accidental harm (in which there is a bad outcome but no bad intent) morally worse than cases of attempted harm (in which there is a bad intent but no bad outcome). With age, the relative weighting of these two factors flips, such that by 5-years-old age, children judged cases of attempted harm morally worse than cases of accidental harm (Cushman et al., 2013). More generally, intentions exert a greater influence than outcomes by around 6–7 years of age, consistent with adults' focus on intentions, although we also note that some work has found signatures of intent-based punishment as young as 3 years of age (Van de Vondervoort & Hamlin, 2018).

Although much less is known about the developmental emergence of partner choice decisions, some recent work on partner choice–like decisions has begun to explore this question. In particular, some work finds that 8-month-olds prefer those who act positively toward prosocial individuals and act negatively toward antisocial individuals (Hamlin et al., 2011) and, holding outcome constant, prefer those who intend to help others over those who intend to harm others (Hamlin, 2013). Ten-month old infants choose those who accidentally cause harm over those who intentionally cause harm (Woo et al., 2017). By 13 months of age, children use information about distributions a third party makes to others to inform their social evaluations: after observing a third party distribute resources either equally or unequally, infants prefer to interact with (e.g., accept a toy from) someone who distributes fairly (Lucca et al., 2018). And, infants give up resources to interact with someone who helps others rather than harms others (Tasimi & Wynn, 2016). At 2 years of age, children prefer to help those who have helped others over those who have harmed others (Dahl et al., 2013); by 4 years, children preferentially choose and distribute resources to partners whose bad action benefited the child (compared with when it did not benefit the child; Myslinska Szarek et al., 2020). Combined, these results demonstrate early-emerging sensitivity to others' behavior in infants' and young children's social evaluations (especially choices between two agents). Furthermore, in some cases, these results suggest sensitivity to the mental states with which agents behave even when holding outcomes caused constant (Hamlin, 2013; Woo et al., 2017)—whether those mental states are goal-directed actions (Woodward, 2009) or epistemic states like knowledge or belief (Van de Vondervoort & Hamlin, 2016). Although these studies do not provide insight into the weight accorded to intentions versus outcomes in young children's social evaluations, they do demonstrate early-emerging sensitivity to intentions.

Overview of the Present Research

Our central research question is whether punishment and partner choice decisions emerge in development in the same way or in distinct ways. We approach this question by examining the sensitivity of these two responses to others' intentions and the outcomes they cause. As outlined above, there is good reason to think that the developmental emergence of punishment and partner choice decisions, and the processes supporting these decisions, may be distinct. In particular, partner choice decisions may exhibit sensitivity to others' intentions, and prioritize this information over outcomes caused, earlier in life than punishment decisions. However, no extant work has directly compared punishment and partner choice decisions in children. The goal of the present research is to

provide such a comparison, focusing on the sensitivity of these responses to intentions and outcomes. We focus on children between the ages of 4 and 9, the ages at which sensitivity to these factors undergoes the most change in past work (Cushman et al., 2013; Hebble, 1971; Imamoglu, 1975; Killen et al., 2011; Nobes et al., 2009, 2016). That is, we recruit children at the age at which punishment first begins to show sensitivity to intentions in past work (~6 years of age), as well as children immediately younger and immediately older than this. Thus, we capture the development stages just before and just after intentions begin to influence punishment, allowing us to investigate whether intentions influence partner choice before, at the same time as or after punishment. In line with past work (Cushman et al., 2013); we present children with vignettes describing situations involving a protagonist and a potential violation (e.g., tripping another child, destroying someone's property). We vary whether the protagonist intended to cause this violation or not and whether they actually do cause this violation or not. Children then rate the appropriateness of engaging in either punishment or partner choice. We predict that children's partner choice ratings will prioritize information about others' intentions earlier than children's punishment ratings, which will only begin to prioritize information about intentions around 5–6 years of age, consistent with past work (Cushman et al., 2013).

Method

Preregistration and Data Availability

The preregistration for this study is available at <http://aspredicted.org/blind.php?x=e7wj7g>. Data and analysis scripts are available at <https://osf.io/7539h/>

Participants

Participants ($N = 187$) were recruited into three predetermined age groups: 4- to 5-year-olds, 6- to 7-year-olds, and 8- to 9-year-olds. Children were grouped in this way so as to facilitate counterbalancing and to ensure relatively equal numbers of children across our entire age range; our main analyses include age as a continuous measure. This preregistered sample size was determined to yield >80% power using simulation-based power calculations, implemented in the R package *simr* (Green & Macleod, 2016); using pilot data from adults as well as an imputed small effect size (Log Odds Ratio = .36). Participants were tested one-on-one in public parks ($n = 150$), a lab space ($n = 22$), a local museum ($n = 11$), and a zoo ($n = 4$). Twenty-one subjects were excluded from the final dataset due to either experimental/technical error ($n = 10$), failure to fully complete the study ($n = 9$), parental interference ($n = 1$), or information from the guardian that the child was not typically developing ($n = 1$). Thus, the final sample size is 166¹ (4–5 years: $n = 57$, M [months] $\pm SD = 61.5 \pm 7.0$; 6–7 years: $n = 54$, 83.1 ± 6.6 ; 8–9 years: $n = 55$, 106.6 ± 6.5 ;

¹ We targeted a preregistered final sample of 162 participants. Four extra participants were collected because of an error in tracking our sample size: Four participants were incorrectly excluded, and thus four replacement participants were recruited. Data from these additional participants has not been discarded.

see Table S1 in the online supplemental materials for gender breakdown). Verbal assent and informed parental consent were obtained from all subjects as well as written assent for participants who were 8 years of age or older. All methods and procedures were approved by the Boston College Institutional Review Board. Information on our adult participant sample can be found in [online supplemental materials](#).

Stimuli: Familiarization

Each participant was shown a “familiarization” vignette, consisting of three cartoon images (adapted from Cushman et al., 2013). The vignette described a protagonist accidentally scooping up the wrong fish from an aquarium tank. The vignette was presented aloud by the experimenter and accompanied by images presented on an iPad. After the presentation of the story, participants were asked two comprehension questions, one targeted at the protagonist’s intention (“Did Veronica want to get the circle fish?”) and one targeted at the outcome produced (“Did Veronica actually get the circle fish?”). Participants responded by touching a cartoon puppy head nodding “YES” or shaking its head “NO” presented on an iPad. For all comprehension questions throughout the experiment, including the warmup, if the participant failed to answer a question correctly, the experimenter explained the answer and then asked again. If the participant still did not provide the correct answer, the experimenter presented the vignette once more and then asked the comprehension question again. If the participant failed to answer the question correctly again, the experimenter explained the correct answer and moved on. Almost all participants were able to answer these questions correctly (>95% provided the correct response either spontaneously or with additional questioning; see Table S2 in the online supplemental materials).

Stimuli: Main Study

A battery of 12 vignettes was constructed (four adapted from Cushman et al., 2013), each involving either potential property damage or minor violation to an individual. Each vignette included three conditions: intentional violation, attempted violation, and accidental violation. The intentional condition included a protagonist who wanted to bring about a bad outcome (e.g., to break a mirror, to push a child) and who succeeded in doing so. The attempted condition included a protagonist who wanted to bring about a bad outcome but failed. Finally, the accidental condition included a protagonist who does not want to do anything bad, but accidentally causes a bad outcome. Thus, protagonists in the intentional condition had a bad intention and produced a bad outcome, protagonists in the attempted condition had a bad intention and produced a neutral outcome, and protagonists in the accidental condition had a neutral intention and produced a bad outcome. Images depicting the events in the vignette were created using VYond software (<https://www.vyond.com>). A synopsis of one vignette in each of the three conditions is provided below (differences across conditions bolded; all vignette text can be found in the [online supplemental materials](#); all images accompanying the vignettes can be found in the [online supplemental materials](#)):

“Baseball,” Intentional: This is Derek. Derek sees a baseball on the floor of his dining room. Next to the box where the balls go, Derek

sees his Mom’s favorite mirror. Derek wants to break the mirror by throwing the baseball at the mirror. Derek throws the baseball at the mirror and the mirror breaks. Derek wanted to break his Mom’s favorite mirror, and he really did break the mirror.

“Baseball,” Attempt: This is Cliff. Cliff sees a baseball on the floor of his dining room. Next to the box where the balls go, Cliff sees his Mom’s favorite mirror. Cliff wants to break the mirror by throwing the baseball at the mirror. But just as Cliff is throwing the ball, he sneezes, so the ball doesn’t go where he wanted. Instead of hitting and breaking the mirror, the ball goes into the box. Cliff wanted to break his Mom’s favorite mirror, but really, he didn’t break anything at all.

“Baseball,” Accident: This is Jack. Jack sees a baseball on the floor of his dining room. Next to the box where the balls go, Jack sees his Mom’s favorite mirror. Jack wants to put the baseball away in the box by throwing it into the box. But just as Jack is throwing the ball, he sneezes, so the ball doesn’t go where he wanted. Instead of going into the box, the ball hits and breaks the mirror. Jack didn’t want to break anything, but really he broke his Mom’s favorite mirror.

Design

The study followed a 2×3 within-subjects design with two main conditions. The first condition was rating type: participants made a punishment judgment on half of trials and made a partner choice judgment on the other half (order counterbalanced). The second condition was protagonist behavior, with three levels: intentional violation, attempted violation, or accidental violation (described above). Every participant completed 12 trials, divided into two blocks of six trials. Within these blocks, participants assessed punishment or partner choice for all trials after being presented with two vignettes of each of our three protagonist behavior types. The order that the protagonist behavior condition was presented in was pseudorandomized, such that each participant was presented with an intentional violation, an attempted violation, and an accidental violation within the first three trials. Every participant was presented with all 12 vignette scenarios, each in one of the three violation conditions. Across participants, we counterbalanced which rating type block was presented first, the order of protagonist behaviors, and the order of comprehension questions displayed. There were two orders in which vignettes were presented, one that was randomly determined and one that was the reverse of the first.

Procedure

Participants were tested individually with a single experimenter. Participants were told they would hear a few stories and be asked questions about the story and the characters in it. Vignettes were presented aloud by the experimenter individually and accompanied by a series of three to four images presented on an iPad that depicted the scenario. For the first three vignettes, participants completed two comprehension questions per vignette. Our counterbalancing scheme (see Design) ensured that each level of protagonist behavior occurred once in these first three trials. Participants were asked the same comprehension questions as during the warm-up: “Did [protagonist] want [the negative outcome to occur]?” and “Did [protagonist] actually [produce the negative outcome]?” (order counterbalanced). Participants responded by touching a cartoon puppy head nodding “YES” or shaking its head

“NO” presented on an iPad. Almost all participants were able to pass these comprehension questions spontaneously or with additional questioning (correct answer provided >95% of the time; see Table S2 in the online supplemental materials). After these comprehension questions (for the first three trials) or after the story was finished (for all other trials), a partner choice or punishment question was then asked.

On partner choice trials, an image of a boy or girl was presented in the center of the screen, with the protagonist from the story and another gender- and race-matched character presented above (see Figure 1). Participants were told about an additional child who knew about the events just described and was deciding who to invite to an event (this event varied by vignette: the movies, a baseball game, a game of kickball, etc.; see the online supplemental materials for the text of all vignettes). This child could invite the protagonist or another child. For example, in the Baseball vignette above, participants would be told “Jake is starting a hockey team. Jake knows what happened earlier in the dining room. Jake is thinking about inviting Derek to join his new hockey team or inviting another kid instead. Should Jake invite Derek to join his hockey team? Click here if you think Jake should invite Derek and click here if you think Jake should invite another kid.” Participants then made their response by touching one of the two characters.

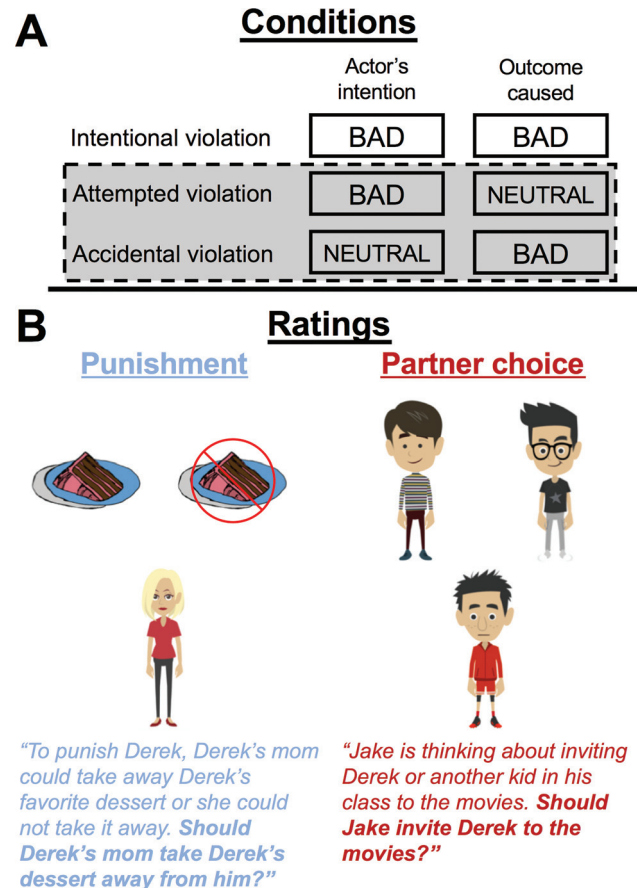
On punishment trials, an image of the protagonist’s parent was presented, with a desirable item (e.g., a dessert or a toy) and the desirable item with a red circle with a line through it (a “NO” symbol) on top of it presented above (see Figure 1). Participants were told that this parent learned about the events just described, and the possibility of punishment was described (this punishment varied by vignette: losing favorite candy, losing favorite baseball mitt for a week, losing a birthday present, etc.; see online supplemental materials for the text of all vignettes). For example, in the Pool example above, participants would be told “Derek’s mom realizes what happened in the dining room. To punish Derek, Derek’s mom could take Derek’s favorite dessert OR she could not take it away. Should Derek’s mom take Derek’s dessert away from him? Click here if you think Derek should get his dessert, or click here if you think Derek should not get his dessert.” Participants then made their choice by touching one of the two images on the iPad. For both partner choice and punishment ratings, the positioning (right vs. left) of the possible answers was randomized across trials and participants.

After the 12 trials, participants were asked a series of 14 additional questions. Twelve of these questions were designed to control for variability across participants in how severe particular punishment and partner choice options were perceived. These questions asked the participants to choose, for the options available for each vignette, whether the partner choice or punishment response was worse (e.g., whether it would be worse to have your favorite toy taken away for a week or to not be invited to the movies). The order of these questions corresponded to the order in which vignettes were presented. The final two questions asked participants to explain how they made their decisions (Table S3 in the online supplemental materials). The study was then concluded.

Analysis Approach

Data were downloaded from Qualtrics software and analyzed using R Version 3.5.0 (R Core Team, 2018). To analyze our results, we use mixed-effects logistic regression implemented in

Figure 1
Task Design



Note. (A) Participants ($n = 166$) were presented with vignettes that varied an actor’s intention and the outcome caused, across three conditions. In cases of intentional violation, the actor intended and successfully caused a violation (e.g., throwing a baseball at and breaking a mirror intentionally). These cases are a control condition, in that we do not expect ratings of punishment and partner choice to differ for them. Our critical cases (highlighted) are attempted and accidental violations. In cases of attempted violation, the actor intended but failed to cause a violation. In cases of accidental violation, the actor did not intend to but did cause a violation (e.g., throwing a baseball at a box near a mirror and accidentally hitting the mirror). (B) Participants made either a punishment or a partner choice judgment on each trial. Punishment involved assessing whether a third party (the actor’s parent) should take something (e.g., their favorite dessert) away from the actor or not. Partner choice involved assessing whether a third party (the child in red) should interact with (e.g., invite to go to the movies) the actor (the child in stripes) or a random other individual. See the online article for the color version of this figure.

package lme4 (Bates et al., 2014; with endorsement of engaging in punishment or partner choice coded as 1 and endorsement of not engaging in punishment or partner choice coded as 0). Our main predictors included rating type (punishment, partner choice), condition (intentional violation, attempted violation and accidental violation; accidental violation was the default reference level, with attempted violation used as a comparison reference level as noted),

and age (preregistered to be included in months, z-scored). For ease of data visualization, we plot data according to the three predetermined age groups (4–5, 6–7, and 8–9 years of age) that participants were recruited into (analyzing our data with age as a three-level factor does not change the overall pattern of results; Table S4 in the online supplemental materials).

To control for response variability across participants and vignettes, we include in all models a random intercept for each participant as well as each vignette. The overall pattern of statistical results is unchanged if a predictor for each individual participant's performance on comprehension questions is included in models (Table S5 in the online supplemental materials).

Because of the within-subjects nature of our design, it is possible that evaluating either punishment or partner choice first could influence subsequent ratings of the other type. Consistent with our preregistration, we first test for this possibility, examining the influence of the order in which the two ratings were made as a predictor. Indeed, we find that order matters: we find an interaction between rating type order, rating type and condition (LRT $\chi^2[3] = 22.04, p < .001$) and between rating type order, age and condition (LRT $\chi^2[3] = 21.20, p < .001$), (see Results in the online supplemental materials for further analysis of this order effect and Table 1, model 1 for full model results). Thus, to avoid this unwanted influence of rating order, we focus our analyses on only

Table 1

Estimate and Standard Error of Fixed Effects in Generalized Linear Mixed Models Predicting Children's Punishment and Partner Choice Ratings

Measure	Model 1 (Rating Type \times Condition \times Age \times Rating Type Order)	Model 2 (Rating Type \times Condition \times Age [First Half Data])	Model 3 (Rating Type \times Condition \times Age [First Half Data, Attempted Violation Reference Level])
Intercept	-0.16 (0.19)	0.09 (0.22)	0.96 (0.23)***
Rating type: Pun.	0.17 (0.24)	-0.07 (0.30)	-0.07 (0.31)
Cond.: Att.	0.45 (0.23)	0.87 (0.26)***	
Cond.: Int.	1.17 (0.24)***	1.66 (0.32)***	
Age	-0.40 (0.19)*	-0.03 (0.21)	0.29 (0.22)
Rating type order: Part. Choice first	0.24 (0.26)		
Rating Type: Pun \times Cond.: Att.	0.34 (0.34)	-0.00 (0.37)	
Rating Type: Pun. \times Cond.: Int	0.47 (0.36)	0.12 (0.42)	
Rating Type: Pun. \times Age	-0.28 (0.24)	-0.71 (0.30)*	0.39 (0.32)
Cond.: Att. \times Age	0.55 (0.24)*	0.32 (0.26)	
Cond.: Int. \times Age	0.36 (0.25)	1.22 (0.32)***	
Rating Type: Pun \times Rating Type Order: Part. Choice First	-0.76 (0.34)*		
Cond.: Att. \times Rating Type Order: Part. Choice First	0.35 (0.34)		
Cond.: Int. \times Rating Type Order: Part. Choice First	0.40 (0.39)		
Age \times Rating Type Order: Part. Choice First	0.37 (0.27)		
Rating Type: Pun. \times Cond.: Att. \times Age	0.73 (0.35)*	1.10 (0.38)**	
Rating Type: Pun. \times Cond.: Int. \times Age	0.65 (0.37)	-0.12 (0.43)	
Rating Type: Pun. \times Cond.: Att. \times Rating Type Order: Part. Choice First	-0.57 (0.49)		
Rating Type: Pun. \times Cond.: Int. \times Rating Type Order: Part. Choice First	-0.10 (0.55)		
Rating Type: Pun. \times Age \times Rating Type Order: Part. Choice First	0.27 (0.34)		
Cond.: Att \times Age \times Rating Type Order: Part. Choice First	-0.26 (0.34)		
Cond.: Int. \times Age \times Rating Type Order: Part. Choice First	0.80 (0.39)*		
Rating Type: Pun \times Cond.: Att. \times Age \times Rating Type Order: Part. Choice First	-0.79 (0.49)		
Rating Type: Pun \times Cond.: Int. \times Age \times Rating Type Order: Part. Choice First	-0.80 (0.56)		
Cond. (Att. ref.): Acc.			-0.87 (0.26)***
Cond. (Att. ref.): Int.			0.79 (0.32)*
Rating Type: Pun. \times Cond. (Att. Ref.): Acc.			0.00 (0.37)
Rating Type: Pun. \times Cond. (Att. Ref.): Int.			0.12 (0.43)
Cond. (Att. Ref.): Acc. \times Age			-0.32 (0.26)
Cond. (Att. Ref.): Int. \times Age			0.90 (0.33)**
Rating Type: Pun. \times Cond. (Att. Ref.): Acc. \times Age			-1.10 (0.38)**
Rating Type: Pun. \times Cond. (Att. Ref.): Int. \times Age			-1.22 (0.44)**
AIC	2,376.26	1,127.62	1,127.62
BIC	2,521.31	1,196.01	1,196.01
Log likelihood	-1,162.13	-549.81	-549.81
Number of observations	1,956	978	978
Number of subjects	163	163	163
Number of vignettes	12	12	12
Variance (subjects)	0.54	1.02	1.02
Variance (vignettes)	0.02	0.03	0.03

Note. AIC = Akaike information criterion; BIC = Bayesian information criterion. Baselines were set as follows (except where noted): condition = accidental, rating type = partner choice, rating type order = punishment first.

* $p < .05$. ** $p < .01$. *** $p < .001$.

the first rating that participants made (note that this includes the first six trials and not just the first trial). A sensitivity analysis of our main result (the significant 3-way interaction between rating type, age, and condition) indicated that we have 80% power to detect an $OR > 3.0$ or $< .33$, indicating we have reasonable power to detect medium-sized effects.

A large majority (88%) of the sessions were videotaped. For responses on the main trials, children's responses were recorded on an iPad (using Qualtrics online software). Children's free responses when asked to explain how they made their decisions were recorded on paper by the primary experimenter and later recoded by an independent paper coder as well as an independent video coder. These responses were coded as belonging to one of five categories (described below) by two independent coders in a nonmutually exclusive fashion (a response could thus belong to no categories or up to all five categories). These coders disagreed $< 5\%$ of the time as to whether a response did or did not belong to each of the five categories for each participant. Looking at each question and collapsing across participants, coders agreed $> 80\%$ of the time for each category within each question. Disagreements were resolved by consensus.

Results

Does Sensitivity to Intentions Emerge in Development Differently for Punishment and Partner Choice?

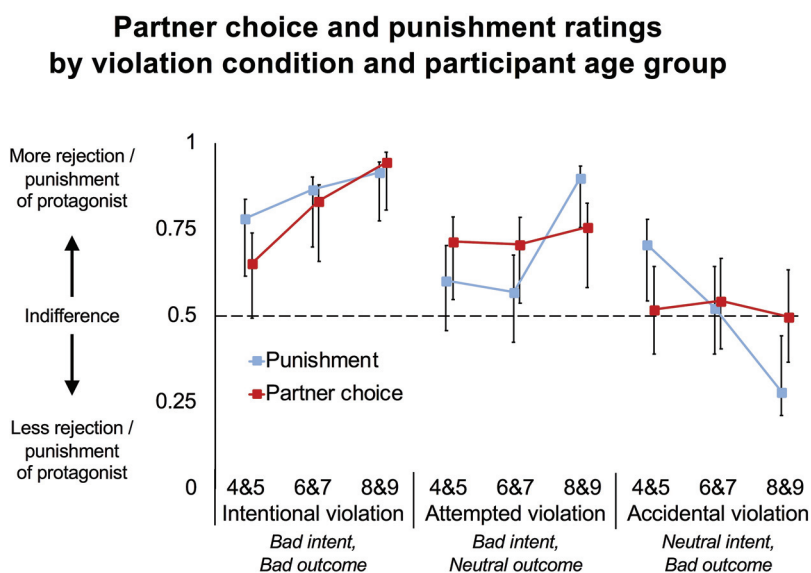
We first turn to our main question—does sensitivity to intentions emerge in development differently for ratings of punishment and partner choice? To answer this question, we ran models including regressors for condition (once with accidental violation as the reference level and once with attempted violation as the

reference level), age, and rating type, as well as all possible interactions between them (Models 2 and 3, Table 1). We found a significant three-way interaction (Figure 2; $LRT \chi^2[2] = 11.39, p < .005$). In particular, we find an interaction between rating type, age and ratings of accidental violation versus attempted violation ($OR = 3.01, 95\% CI [1.44, 6.36], z = 2.91, p = .004$) as well as between rating type, age and judgments of attempted violation versus intentional violation ($OR = .29, 95\% CI [.12, .69], z = -2.76, p = .006$).

How Does Punishment Emerge?

Looking at our punishment results alone, we find that they match past work (Cushman et al., 2013). In particular, we find that intentions come to exert a greater influence with age. We found two-way interactions between age and accidental versus attempted violations ($OR = 4.29, 95\% CI [2.50, 7.67], z = 5.10, p < .001$) and between age and accidental versus intentional violations ($OR = 2.98, 95\% CI [1.68, 5.46], z = 3.65, p < .001$)—suggesting that intentions exert an influence that is modulated by age—but not between age and attempted versus intentional violations ($OR = .69, 95\% CI [.38, 1.26], z = -1.21, p = .23$), cases that differ only by the absence or presence of a bad outcome. These results are further confirmed when comparing the judgment for each condition within each age group against .5, the point of indifference between judging punishment as appropriate and as not appropriate. We find that punishment is greater than indifference for intentional violations at all ages (all $p < .005$, all $t > 3.52$, all $d > .65$). More importantly, we find that punishment in cases of attempted violation is not significantly greater than indifference at either 4–5 ($p = .26, t = 1.15, d = .21, 95\% CI of M = .43-.74$) or 6–7 years of age ($p = .42, t = .83, d = .16, 95\% CI of mean .42-.69$), but is significantly greater at 8–9 years of age ($p < .001, t = 8.93, d = .87, 95\%$

Figure 2
Participants' Punishment and Partner Choice Ratings in Cases of Intentional, Attempted, and Accidental Violation Across Age



Note. Predicted effects from our final model are plotted. Error bars show 95% confidence intervals. See the online article for the color version of this figure.

CI of $M = .78-.95$). In contrast, we find that punishment in cases of accidental violation is significantly greater than indifference at by 4–5 years of age ($p = .03$, $t = 2.28$, $d = .67$, 95% CI of $M = .52-.83$), not significantly different at 6–7 years of age ($p = .82$, $t = .23$, $d = .52$, 95% CI of $M = .36-.68$) and significantly less than indifference at 8–9 years of age ($p = .02$, $t = -2.48$, $d = .32$, 95% CI of mean $.17-.47$). We note that any null findings from these comparisons should be interpreted with caution given the reduced power they have to detect small differences.

Looking at the pattern of results, we find that punishment ratings are relatively more sensitive to the presence of bad outcomes early in life, seen in higher punishment of accidental violations and lower punishment of attempted violations at 4–5 years of age. With age, intentions exert a greater influence, such that by 8–9 years of age both attempted and intentional violations are punished strongly, and punishment is not chosen in cases of accidental violation.

How Does Partner Choice Emerge?

For partner choice ratings, we find a different pattern of results. Here, we find 2-way interactions between age and accidental versus intentional violations ($OR = 3.51$, 95% CI [1.89, 6.94], $z = 3.80$, $p < .001$) and between age and attempted versus intentional violations ($OR = 2.62$, 95% CI [1.39, 5.22], $z = 2.87$, $p = .004$), but not between age and accidental versus attempted violations ($OR = 1.34$, 95% CI [.79, 2.28], $z = 1.10$, $p = .27$). Two features of this pattern of results are worth highlighting. First, partner choice in cases of intentional violation becomes harsher with age. Second, we do not observe an increasing difference in judgment between cases of attempted versus accidental violation (but do observe an overall difference, $OR = 2.28$, 95% CI [1.35, 3.93], $z = 3.03$, $p = .002$). One interpretation of these results is that partner choice ratings become increasingly sensitive to intentions with age when assessing intentional violations (especially relative to accidental violations). This interpretation is consistent with the fact that partner choice ratings are generally sensitive to intentions, seen in partner choice in cases of attempted violations. However, it is unclear why the increasing influence of intentions with age would manifest in cases of intentional violation but not cases of attempted violation, the judgment of which does not change with age. An alternative interpretation of these results is that partner choice ratings become increasingly sensitive to outcomes with age when assessing intentional violations, given that intentional violations involve a bad outcome as well as a bad intention. However, this second interpretation is inconsistent with the finding that partner choice ratings are never significantly different from indifference for accidental violations, as described below. Thus, the pattern of ratings in this condition is most consistent with the idea that intentions exert a greater influence with age on partner choice in the context of intentional violations, with little to no change in ratings of attempted or accidental violations, although we note that inconsistent data exists (specifically, the lack of harsher judgment of attempted violations with age) and more work on this developmental pattern is needed.

We can unpack these results further by again comparing the judgment for each condition within each age group against .5, the point of indifference between judging partner choice as appropriate and as not appropriate. We find that judgment of partner choice

is not greater than indifference for intentional violation at 4–5 years of age ($p = .15$, $t = 1.49$, $d = .28$, 95% CI of mean $.45-.80$; see the Discussion for more analysis of this null result) but is greater than indifference at 6–7 and 8–9 years of age (both $p < .001$, both $t > 4.56$, $d > .78$). In contrast, we find that partner choice judgments in cases of attempted violation are always greater than indifference (all $p < .04$, all $t > 2.21$, all $d > .43$) but never greater than indifference for cases of accidental violation (all $p > .63$, all $t < .49$, all $d < .10$). We again note that any null findings from these comparisons should be interpreted with caution given the reduced power they have to detect small differences.

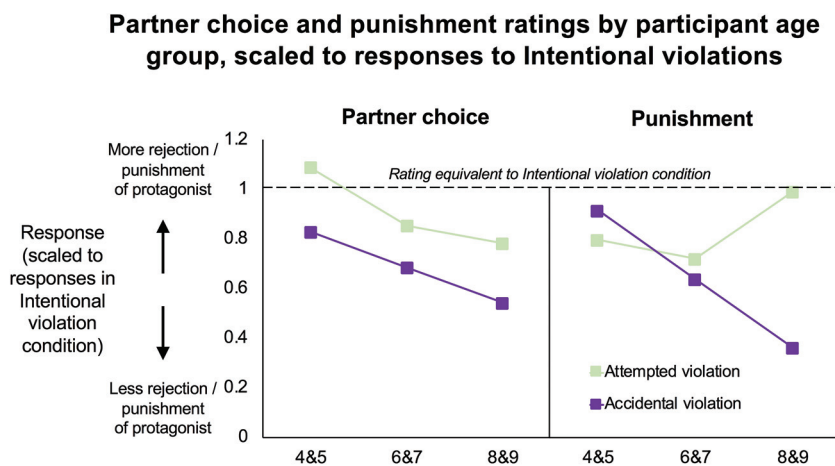
Exploring the Influence of Violation type

Another way of examining the 3-way interaction between rating type, age and protagonist behavior is by investigating the influence of rating type and age within each condition. First looking at accidental violations, we find a significant interaction between rating type and age ($OR = .44$, 95% CI [.23, .82], $z = -2.34$, $p = .02$), indicating that partner choice and punishment ratings of accidental violations shift with age. In contrast, we find no significant interaction between rating type and age when assessing attempted violations ($OR = 1.40$, 95% CI [.80, 2.59], $z = 1.20$, $p = .22$) or intentional violations ($OR = .30$, 95% CI [.01, 1.03], $z = -1.66$, $p = .10$). These results again highlight the role that development plays on the assessment of accidental violations, and in particular the difference between punishment and partner choice assessments of accidental violations with age. See the Results in the [online supplemental materials](#) for a similar analysis in which we examine the three-way interaction by examining the influence of rating type and condition for each age group.

Baselining Ratings to the Intentional Violation Condition

A final way of understanding these data is by plotting our data a second way. Specifically, given that our intentional violations can be thought of as a control condition (in that both punishment and partner choice should be endorsed in these cases), we can plot ratings for attempted and accidental violations relative to ratings for intentional violations. That is, if 75% of participants endorsed punishment of intentional violations and 60% of participants endorsed punishment of attempted violations, attempted violations would now be plotted at 80% ($60\%/75\% = 80\%$). This visualization (see [Figure 3](#)) treats intentional violation as a baseline, highlighting the difference in judgment of attempted and accidental violations across punishment and partner choice ratings. In particular, we can see larger change in punishment ratings than partner choice ratings for accidental violations across age, highlighting the increasing role that intentions play in punishment with development. In addition, whereas partner choice rating of attempted violations become less like partner choice ratings of intentional violations with age, the opposite is true for punishment ratings—whereas attempted violations are punished less often before 8 years of age, 8–9 year-olds punish attempted violations very similarly to intentional violations. Combined, this way of inspecting our data highlights the increasing influence that intentions have with age on punishment

Figure 3
Participants' Punishment and Partner Choice Ratings in Cases of Attempted and Accidental Violations Across Age Groups, Normalized With Respect to Ratings in Cases of Intentional Violations



Note. Predicted effects from our final model are plotted. See the online article for the color version of this figure.

ratings relative to partner choice ratings, especially for cases of accidental violation.

Together, these results suggest that sensitivity to intentions does emerge differently for punishment and partner choice, and in particular that partner choice decisions are sensitive to intentions earlier than punishment decisions.

What Do Children Say About How They Made Their Choices?

Following the main trials, children were asked how they made their punishment and partner choice ratings. These responses were coded as referencing five nonmutually exclusive categories: Intentions (e.g., “on purpose,” “didn’t mean to”), outcomes (e.g., “broke,” “hit”), actions (e.g., “did something bad”), dispositions/character (e.g., “bad boy,” “nice girl”), and deterrence or preventing future harm (e.g., “stop,” “do that again”). Because our investigation of these responses is exploratory, we have small cell sizes for these responses, and there are a large number of potential comparisons (increasing type II error rates), we report descriptive statistics only for these data.

Across questions, children tended to reference actions (18.5% of responses) and intentions (12.1%), followed by dispositions (9.2%) and outcomes (7.0%). While references were largely similar across age and punishment versus partner choice explanations (see Table 2), we note two interesting patterns in their responses. First, 8- to 9-year-old children tended to reference intentions more when explaining their punishment ratings (34.5%) than their partner choice ratings (14.5%). This was not true for younger children (6- to 7-year-olds: punishment [8.7%], partner choice [8.2%]; 4- to 5-year-olds: punishment [4.3%], partner choice [0%]). Second, and relatedly, children tended to reference dispositions and character more when explaining partner choice than punishment, especially at 6–7 years of age (10.2% vs. 4.3%) and 8–9 years of age (21.8% vs. 5.5%). Finally, this shift when explaining partner

choice decisions appears to be driven by a tradeoff between referencing intentions versus referencing dispositions. That is, when explaining punishment, if children referenced intentions they tended to do so exclusively (4–5: 2.1%; 6–7: 6.5%; 8–9: 34.5%) and not while also mentioning dispositions (4–5: 2.1%; 6–7: 2.2%; 8–9: 0%). They also rarely mentioned dispositions exclusively (4–5: 4.3%; 6–7: 2.2%; 8–9: 5.5%). In contrast, when explaining partner choice decisions, children tended to either reference intentions exclusively (4–5: 0%; 6–7: 8.2%; 8–9: 12.7%) or to reference dispositions exclusively (4–5: 8.2%; 6–7: 10.2%; 8–9: 20%) and almost never referenced both factors (4–5: 0%; 6–7: 0%; 8–9: 1.8%). In sum, when explaining their partner choice ratings, 8- to 9-year-old children tend to reference dispositions that underlie intentions more than the intentions themselves. However, we again note that these patterns should be interpreted cautiously, given the issues we outlined above and the fact that we do not perform statistical analyses on these data.

Is There Variability in Whether Punishment or Partner Choice Seems More Severe?

After all main trials were completed, each participant rated whether the punishment or partner choice option was more severe for each vignette. For instance, they were asked whether it is worse to have your favorite dessert taken away or to not be invited to join a hockey team. These ratings allow us to determine if either the punishment or partner choice options we included are perceived as more severe and to account for individual differences in perceptions of severity in participants’ responding.

First, we model these severity ratings as a function of the individual vignette they accompanied as well as participants’ age. We find no interaction between vignette and age (LRT $\chi^2[22] = 13.5$, $p = .92$) and no main effect of age (LRT $\chi^2[2] = 4.87$, $p = .09$). We do find a main effect of vignette (LRT $\chi^2[11] = 68.2$, $p < .001$), indicating that whether the punishment or partner choice option

Table 2

Percentage of Participants' Responses Falling Into One of Five Nonmutually Exclusive Categories When Asked to Explain How They Made Their Punishment and Partner Choice Ratings

Measure	Punishment			Partner choice		
	4–5 years old	6–7 years old	8–9 years old	4–5 years old	6–7 years old	8–9 years old
Intentions	4.3%	8.7%	34.5%	0.0%	8.2%	14.5%
Outcomes	2.1%	4.3%	18.1%	0.0%	2.0%	12.7%
Actions	6.4%	19.6%	36.4%	0.0%	12.2%	36.3%
Dispositions/character	6.4%	4.3%	5.5%	8.1%	10.2%	21.8%
Deterrence	0.0%	0.0%	3.6%	2.0%	0.0%	0.0%

was more severe varied across vignettes. Importantly, however, we find that the overall pattern of participants partner choice and punishment ratings is unchanged if a predictor for each participant's own rating of whether the punishment or partner choice option was more severe for each vignette is included in models (Table S6 in the online supplemental materials). Furthermore, we find that, coding the punishment option as 0 and the partner choice option as 1, the mean severity rating is not different from .5 ($p = .12$, 95% CI [.50, .54]), indicating no overall bias toward viewing one response as more severe than the other.

Do These Trends Continue Into Adulthood?

In a follow-up study, we investigate whether the trends found in children extend into adulthood. Specifically, we present a sample of adult participants with a text version of our vignettes and ask them to make punishment and partner choice judgments in the same manner as children (see Methods and Results in the online supplemental materials). We find two main patterns. First, adults' punishment is consistent with the trend found here and with past work (see Figure 2 in the online supplemental materials). Specifically, punishment is sensitive both to bad outcomes (seen in greater punishment of intentional and accidental violations compared with attempted violations) as well as bad intentions (seen in greater punishment of intentional and attempted violations than accidental violations). Interestingly, we find some evidence supporting a U-shaped pattern in the emergence of intentions in punishment judgments (Margoni & Surian, 2016). Specifically, intentions exert little influence on young children's punishment and come to exert the strongest influence by around 8–9 years of age, seen in no difference in punishment of intentional and attempted violations. In adulthood, this influence is tempered, such that intentional violations are punished more than attempted violations. Regardless, we see a pattern of results which is consistent with the trends observed in our own data as well as other work on the emergence of punishment (Cushman et al., 2013; Imamoglu, 1975; Nobes et al., 2009, 2016; Zelazo et al., 1996).

Second, we find that intentions are prioritized to an even greater extent in adult partner choice judgments compared with children's judgments (Figure 2 in the online supplemental materials). That is, adults now view partner choice as inappropriate in response to accidental violations (cases in which intent was benign) and show a smaller difference in partner choice between intentional and attempted violations (cases that differ in outcome but not intent). Thus, adults continue the pattern found in children in which partner choice judgments are influenced principally by intentions.

Discussion

We explored the developmental emergence of punishment and partner choice, focusing on the sensitivity of these two responses to others' intentions and the outcomes they cause. Our results are the first to demonstrate a dissociable pattern in how these two responses emerge. Four- to 9-year-old children rated how appropriate punishment and partner choice were in response to cases that varied whether the protagonist intended something bad or not and whether they caused something bad or not. In line with our hypotheses, we find very different developmental trajectories for punishment and partner choice ratings. First, our punishment results replicate past work (Cushman et al., 2013). Punishment is sensitive principally to outcomes caused early in life. Younger children judge that punishment should occur in cases of accidental violation (cases in which there is a bad outcome but no bad intent) but do not say that punishment should occur in cases of attempted violation (cases in which there is no bad outcome and a negative intent). With age, intentions begin to exert an increased influence, such that older children judge that punishment should not occur in cases of accidental violation and should occur in cases of attempted violation. In contrast, partner choice ratings are sensitive to the intent with which a protagonist behaved earlier than punishment decisions. At the earliest age we tested (4 years), we find signs of sensitivity to intentions: children report that partner choice should occur in the case of attempted violation. However, we find no sensitivity to outcomes, seen in the lack of partner choice in response to accidental violations. This pattern holds steady through development, with children judging that partner choice should occur in response to attempted violation at all ages and no judgment that partner choice should occur in response to accidental violation at any age.

This pattern of results—with partner choice prioritizing information about intentions over outcomes at an earlier age than punishment—has broad implications for our understanding of the development of the use of other's intentions, the development of moral judgment and the processes underlying partner choice and punishment.

The Development of the Use of Other's Intentions

A long history of work has focused on how sensitivity to other's mental states emerges in moral judgment (Cushman et al., 2013; Hebble, 1971; Imamoglu, 1975; Killen et al., 2011; Li & Tomasello, 2018; Nobes et al., 2009, 2016; Piaget, 1932; Van de Vondervoort & Hamlin, 2018; Zelazo et al., 1996). One fundamental

finding from this literature is an outcome-to-intent shift, in which early moral judgments (<4 years of age) are principally sensitive to outcomes caused and sensitivity to intentions emerges only with development, by around 5–6 years of age (Cushman et al., 2013; Hebble, 1971; Imamoglu, 1975; Nobes et al., 2009, 2016; Zelazo et al., 1996). However, recent work has begun to paint a more nuanced picture. First, the emergence of sensitivity to intentions does not occur simultaneously for all types of moral judgments. In particular, judgments of wrongness develop sensitivity to intentions before judgments of punishment (Cushman et al., 2013). Second, recent work with infants demonstrates that they can take agents' mental states (Hamlin, 2013; Woo et al., 2017) as well as their previous prosocial or antisocial actions (Dahl et al., 2013; Tasimi & Wynn, 2016) into account when making sociomoral evaluations. Although the nature of the mental states that infants take into account is under debate (Van de Vondervoort & Hamlin, 2016; Woodward, 2009); these results suggest that at least some degree of information about others' mental states or disposition can exert an influence on early forms of moral judgment (e.g., sociomoral evaluations) at a relatively young age. Moreover, some work has demonstrated that 3-year-olds can take intentions into account under certain conditions when making punishment decisions (Van de Vondervoort & Hamlin, 2018). In part owing to this evidence, some have begun to question the outcome-to-intent shift and have instead proposed a U-shaped developmental curve (Margoni & Surian, 2016). This proposal is motivated by the recently observation that intentions impact even infants' judgments (Dunfield & Kuhlmeier, 2010; Hamlin, 2013; Woo et al., 2017). Incorporating these results with prior work on the outcome-to-intent shift, intentions seem to exert an influence on judgment very early in life, a reduced influence on judgments by older children and a strong influence again on judgments by adults; thus, the influence of intentions through development is characterized by a U-shaped curve.

Our results are consistent with these new patterns in a few ways. First, we find that intentions can be used earlier for some types of judgments (i.e., partner choice judgments) than others. This is consistent with earlier-emerging sensitivity to intentions for wrongness judgments than for punishment judgments (Cushman et al., 2013). Combined, these results call into question accounts which suggest that the emergence of sensitivity to intentions requires the development of increased cognitive resources (e.g., Killen et al., 2011; Zelazo et al., 1996). Were that true, once this resource-dependent capacity emerged, it should impact all moral judgments simultaneously. Thus, our results suggest that the increased integration of intentions into moral judgments cannot fully be explained by the emergence of a resource-dependent capacity. By providing evidence against this possibility, our results are more consistent with arguments suggesting that this integration is facilitated by conceptual reorganization in the moral domain (Cushman et al., 2013). That is, around 6–7 years of age, a novel intent-based concept of wrongness emerges which subsequently constrains punishment judgments. Our results do not provide direct support for this hypothesis but are rather consistent with it by providing evidence against the alternative account, while also finding a pattern of development of punishment ratings that mirrors this prior work. The relationship between this concept of wrongness and judgments of partner choice is an intriguing area for future work.

Second, our results are consistent with results suggesting that infants can take agent's intentions into account when making sociomoral evaluations. Critically, these studies frequently ask children to choose between agents, making these judgments an early form of partner choice decision (Hamlin, 2013; Lucca et al., 2018; Tasimi & Wynn, 2016; Woo et al., 2017). For instance, 8-month old infants choose those who intend to help others over those who intend to harm others (Hamlin, 2013) and those who act positively toward prosocial individuals and act negatively toward antisocial individuals (Hamlin et al., 2011). In addition, 10-month-olds choose those who accidentally cause harm over those who intentionally cause harm (Woo et al., 2017). Of course, there are important differences in these preference decisions and the partner choice decisions used in our study. But the fact that even the youngest children in our study engaged in intent-sensitive partner choice judgment suggests continuity in this capacity through middle- to late-childhood.

Finally, we find that the influence of intentions on punishment judgments peaks around 8–9 years of age, relative to adults and younger children. That is, the influence of intent on punishment increases from 4 to 9 years of age, such that by 8–9 years of age punishment of attempted violations, in which there is a bad intent but no bad outcome, is as harsh as punishment of intentional violations, in which there is both a bad intent and a bad outcome. Thus, by this age, intentions dominate punishment. However, we find that the influence of intentions is moderated to some extent into adulthood. We find greater punishment of intentional violations than attempted violations, and greater punishment of attempted violations than accidental violations. This pattern reflects a contribution of both intentions, in that attempted violations are punished more than accidental violations, as well as outcomes, in that intentional violations are punished more than attempted violations. Of course, the influence of outcomes remains weaker than the influence of intentions, which is consistent with the general pattern suggested by the U-shaped emergence of intent-based judgment (Margoni & Surian, 2016).

In sum, we find that sensitivity to intentions when making a punishment judgment in our task reflects something like a U-shaped pattern: Weakest around 4–5 years of age, strongest around 8–9 years of age and continuing into adulthood. Interestingly, we find a more linear pattern in the case of partner choice, in which intentions influence judgment early in life and this influence only becomes stronger with development. Our results are therefore more consistent with a U-shaped development of intentions in the case of punishment than in the case of partner choice. Future work on the degree to which punishment is unique in this regard or not is needed.

We also note that our results are not consistent with all past work. Indeed, some work has found that punishment judgments are sensitive to intentions by 3 years of age (Van de Vondervoort & Hamlin, 2018). There are important differences between the procedure used in this study and our own. In particular, Van de Vondervoort & Hamlin used a live puppet show whereas we used vignettes described out loud and accompanied by images. Using a live puppet show may have increased the salience of the actor's intent or reduced the processing demands required to incorporate the actor's intent into punishment. Furthermore, children in their study were required to infer the actor's intent rather than being told what the actor intended in our study. Finally, children in our

study were asked to decide whether a third-party should or should not punish the protagonist, whereas Van de Vondervoort and Hamlin told children that punishment should occur (“I think that one of these guys should be in trouble.”) and asked children to choose who should get in trouble. These differences in our methodologies, or others not identified here, may account for the fact that Van de Vondervoort and Hamlin observed intent-based punishment at an earlier age than we did. Future work should continue to explore under what circumstances children’s punishment decisions are influenced by others’ intentions.

The Processes Underlying Punishment and Partner Choice

Our results also provide new evidence regarding the processes underlying punishment and partner choice decisions. In line with past work (Liddell & Kruschke, 2014; Martin & Cushman, 2015), we find a dissociation in how sensitive these two responses are to others’ intentions and outcomes caused. Whereas punishment is sensitive to both intentions and outcomes, partner choice is sensitive only to agent’s intentions. We find this both in children and in adults, replicating past work (Liddell & Kruschke, 2014; Martin & Cushman, 2015). Thus, these results add to a growing body of work that these decisions are made in different ways.

More importantly, we find distinct developmental trajectories for these two responses. Punishment judgments only develop sensitivity to intentions with age, by around 6–7 years of age in the case of accidental violation. In contrast, partner choice decisions are sensitive to intentions by 4 years or age, seen in condemnation of attempted violations. Although preliminary, these results are in line with other work showing very early-emerging sensitivity in infants’ choice of agents, as discussed above (Hamlin, 2013; Woo et al., 2017). These distinct developmental trajectories provide evidence that the processes underlying punishment and partner choice are dissociable, in that the processes underlying partner choice decisions begin incorporating information about other’s intentions earlier in life than the processes underlying punishment decisions. Future work should continue to explore this question by probing the nature of these processes more directly.

Finally, this pattern of emergence is consistent with the different functional purposes these responses serve. To the extent that the psychology underlying partner choice is centered around finding and maintaining the best partnerships (Barclay, 2013, 2016; Misyak et al., 2014; Warneken, 2018), understanding others’ intentions is a crucial capacity. How individuals have intended to behave in the past is a critical predictor of how they will intend to behave in the future. Given the importance of this capacity for partner choice decisions, it makes sense that intentions would begin to exert an influence on partner choice decisions earlier than other ways we can evaluate partners (e.g., by punishing them). Thus, these developmental data underscore the importance of the processes supporting our understanding of others’ intentions for partner choice decisions.

Implications for Choice of Dependent Variable in Future Work

Beyond their implications for our understanding of punishment and partner choice, our results have methodological implications

for the study of the development of moral judgment more broadly. We find that asking participants to make different kinds of judgment can yield very different results: whether one is asked to assess punishment or partner choice influences the degree to which intentions influence that judgment. Importantly, these are not the only judgments investigated within this field: Previous studies have variously used ratings of punishment (Berg-Cross, 1975; Cushman et al., 2013; Nobes et al., 2009, 2016; Shultz et al., 1986; Van de Vondervoort & Hamlin, 2017; Zelazo et al., 1996), naughtiness (Cushman et al., 2013), wrongness (Li & Tomasello, 2018), badness (Hebble, 1971; Imamoglu, 1975; Nobes et al., 2009, 2016; Zelazo et al., 1996), and liking (Imamoglu, 1975; Li & Tomasello, 2018; Van de Vondervoort & Hamlin, 2017) and have asked children to choose between two agents (Hamlin, 2013; Hamlin et al., 2011, 2013; Kiley Hamlin et al., 2013; Steckler et al., 2017; Tasimi & Wynn, 2016; Woo et al., 2017). Based on these results and others, researchers have formed divergent opinions about the nature of the development of moral judgment, including the existence and shape of the outcome-to-intent shift (e.g., Margoni & Surian, 2016; Nobes et al., 2016). Our results highlight that perhaps part of this disagreement stems from the fact that different judgments paint different pictures: When participants make punishment decisions, they appear less sensitive to intentions than when they make partner choice decisions, for instance. Similar differences likely exist across other types of judgments. However, the extent to which this is the case is unknown, as most studies have not systematically investigated these differences. From a methodological perspective, our results suggest that researchers should carefully consider the type of judgment participants are asked to make and consider including multiple judgments so as to begin to document these differences. From a theoretical perspective, our results help explain the nature of this long-standing disagreement and suggest that greater attention to the type of judgment made may help in resolving it.

Limitations and Future Directions

Our results raise a number of important questions and point to important directions for future work. First, we have interpreted our results as suggesting that children make punishment and partner choice judgments in distinct ways, and in particular that they were expressing moral disapproval of these individuals’ actions. However, an alternative interpretation is that participants were more focused on providing beneficial outcomes to those who had not acted poorly. This would be especially true in the case of partner choice, where participants had the opportunity to have one of the two children invited to a desirable event (e.g., a baseball game, a hockey game, etc.) or not. That is, our focus was on children’s judgment that the protagonist should not be invited to the event, but an alternative interpretation is that children were judging whether the alternative individual should be invited to the event. From this perspective, children’s partner choice judgments may have reflected a psychology focused more on rewarding individuals who had not done anything wrong rather than avoiding those who had done something wrong.

However, a few pieces of evidence make this interpretation less reliable. Most importantly, all children completed a series of questions after the main trials were finished in which they were asked to judge whether the punishment or partner choice outcome for

each vignette was more severe. For instance, one of the questions asked “whether it would be worse to have your favorite coloring set taken away for a week or to not be invited to go to the movies.” We do find that these ratings vary across vignettes, indicating that sometimes the punishment option was viewed as more severe and sometimes the partner choice option was viewed as more severe. However, we find no overall preference for choosing either option over the other and no interaction with age. Thus, the partner choice option was not viewed as depriving children of a more valuable reward than the punishment option. Critically, we also find that including individual participants’ ratings of which option was more severe for each vignette in our models did not change the pattern of results (Table S6 in the online supplemental materials). In other words, we find that differences in viewing either the punishment or partner choice options as more severe do not change how appropriate they are viewed in response to cases varying intentions and outcomes. In addition, the overall pattern of our results is not consistent with this interpretation of partner choice judgments. For instance, if partner choice were used as a reward, we might have expected children’s judgments in cases of accidental violation to favor not engaging in it, especially among older children. That is, by 8–9 years of age, children judge punishment as inappropriate (rather than being indifferent as to whether punishment should occur or not). In contrast, we find no significant preference either for or against partner choice at this age in cases of accidental violation. Finally, we attempted to probe this explanation in our sample of adult data. We included a second group of adult participants who made punishment and partner choice ratings in response to the same vignettes but now rating punishment and partner choice in the abstract (see Methods and Results in the online supplemental materials). That is, in the Punishment condition they rated “Should [protagonist] be punished?,” and in the Partner Choice condition they rated “If you were going to do something fun, would you pick [protagonist] to go with you?.” These ratings are devoid of the concrete elements present in our original ratings, allowing us to test whether these features influenced our pattern of results. We find largely consistent results with these abstract ratings. Partner choice decisions continue to be influenced by intentions to a greater extent than punishment decisions. One interesting result is that punishment decisions of attempted violations were less harsh using these abstract ratings, potentially reflecting a reduced influence of intentions when thinking about punishment in the abstract. Importantly, this was not found for partner choice judgments, meaning that our results—especially in the Partner Choice condition—are unlikely to be influenced by particular features of how these ratings were made. Combined, this pattern of results in children’s judgments, the lack of influence of individual children’s severity ratings on our main results and the lack of influence of changes to the rating made speaks against the idea that our results were influenced by the particular way we asked about partner choice. Nevertheless, we cannot completely rule out such an influence—future work should explore the degree to which the patterns we find here are influenced by the specific manner in which punishment and partner choice were assessed.

Relatedly, in constructing a tight contrast between punishment and partner choice in our paradigm, we removed much of the nuance that distinguishes partner choice and punishment in the real world. In some sense, this makes the differences observed

here all the more substantial—we find differences between these judgments when they are stripped of much of what typically makes them unique. Nevertheless, future work should investigate whether similar results are found when asking about partner choice and punishment in more naturalistic ways.

Second, an unexpected pattern in our data is worth further discussion: Looking at partner choice ratings, we find that at 4–5 years of age, partner choice is viewed as appropriate in the case of attempted violations but not viewed as appropriate in the case of intentional violations. We note that this latter lack of difference is approaching significance ($p = .15$) and that the means for attempted (.72) and intentional violations (.65) at this age are quite close, with mostly overlapping 95% confidence intervals (attempted: .55–.79; intentional: .54–.81). Furthermore, a sensitivity analysis of the partner choice judgments of intentional violations by 4- to 5-year-olds indicates that we have 80% power to detect an effect as large as $d = .56$ and we thus cannot rule out a small difference in this condition. Thus, we do not view this difference in the assessment of attempted and intentional violations at this age as meaningful.

Third, the decisions children made in our study were third-party, judging whether punishment or partner choice should occur in a situation not involving the child. Past work has shown important differences in second- and third-party punishment decisions (FeldmanHall et al., 2014; Gummerum & Chu, 2014; Raihani & Bshary, 2015; Zhou et al., 2016), and future work should explore the degree to which second- and third-party partner choice decisions are made similarly.

Fourth, the decisions made in our study were hypothetical moral judgments, consistent with much past work on how children’s moral judgments emerge (Cushman et al., 2009; Killen et al., 2011; Nobes et al., 2009, 2016). Nevertheless, a fuller picture of the emergence of partner choice, especially in contrast to punishment, would include having children make decisions in which their own payoff was on the line. Indeed, prior adult work showing a dissociation between punishment and partner choice has relied on these types of incentivized economic games (Liddell & Kruschke, 2014; Martin & Cushman, 2015). To the extent that our adult results with hypothetical vignettes mirror these results, we show convergence in conclusions across methods assessing adults’ judgments. Recent work has shown that whether or not punishment is costly influences the age at which children will first engage in it (McAuliffe et al., 2015); future work should explore whether engagement in partner choice is similarly lower when doing so costs the child resources.

Finally, our focus in this article has been on how punishment and partner choice are used in response to others’ transgressions, therefore viewing them from a moral perspective. Of course, as we note in the introduction, both of these responses can be used for a variety of reasons, including ones which are nonmoral. Thus, it is possible that punishment and partner choice may show different patterns of sensitivity in nonmoral contexts; future work should explore this issue. Moreover, partner choice itself is a varied construct (Barclay, 2013, 2016) and is used in contexts beyond responding to transgressions. Although intentions are likely to be prioritized broadly when making a partner choice decision, the relative weight given to intentions might vary depending upon the context or immediate goal. For instance, intentions may be weighted most strongly when choosing someone for a long-term

partnership and less strongly when choosing someone for a single-shot interaction or for a particular skill-dependent need. Future work should explore the degree to which intentions matter for partner choice in contexts beyond responding to transgressions.

Conclusion

When making a moral judgment, much work demonstrates a primary influence for whether the protagonist was intending something bad or not. Recent work has begun to demonstrate that the strength of this influence varies by the response one is making, with partner choice decisions more sensitive to intentions than punishment decisions. Here, we extend this difference to children as young as 4 years of age. We find that punishment and partner choice have distinct developmental trajectories, with partner choice exhibiting sensitivity to other's intentions earlier than punishment. Combined with other recent work, these results suggest that punishment and partner choice decisions are made in different ways, emerge in development with distinct trajectories, and are supported by distinct psychological processes.

References

- Alicke, M., & Davis, T. (1989). The role of a posteriori victim information in judgments of blame and sanction. *Journal of Experimental Social Psychology, 25*(4), 362–377. [https://doi.org/10.1016/0022-1031\(89\)90028-0](https://doi.org/10.1016/0022-1031(89)90028-0)
- Armsby, R. E. (1971). The reexamination of the development of moral judgments in children. *Child Development, 42*(4), 1241–1248. <https://doi.org/10.2307/1127807>
- 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>
- Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans. *Evolution and Human Behavior, 34*(3), 164–175. <https://doi.org/10.1016/j.evolhumbehav.2013.02.002>
- Barclay, P. (2016). Biological markets and the effects of partner choice on cooperation and friendship. *Current Opinion in Psychology, 7*, 33–38. <https://doi.org/10.1016/j.copsyc.2015.07.012>
- Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology, 54*(4), 569–579. <https://doi.org/10.1037/0022-3514.54.4.569>
- Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M. T., Fitzpatrick, S., Gurven, M., Henrich, J., Kanovsky, M., Kushnick, G., Pisor, A., Scelza, B. A., Stich, S., von Rueden, C., Zhao, W., & Laurence, S. (2016). Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proceedings of the National Academy of Sciences of the United States of America, 113*(17), 4688–4693. <https://doi.org/10.1073/pnas.1522070113>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4 (R Package Version, 1, 1–7).
- Berg-Cross, L. G. (1975). Intentionality, degree of damage, and moral judgments. *Child Development, 46*(4), 970–974. <https://doi.org/10.2307/1128406>
- Bliege Bird, R., & Power, E. A. (2015). Prosocial signaling and cooperation among Martu hunters. *Evolution and Human Behavior, 36*(5), 389–397. <https://doi.org/10.1016/j.evolhumbehav.2015.02.003>
- Boyd, R., & Richerson, P. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology, 195*(13), 171–195. [https://doi.org/10.1016/0162-3095\(92\)90032-Y](https://doi.org/10.1016/0162-3095(92)90032-Y)
- Bshary, R., & Grutter, A. S. (2002). Asymmetric cheating opportunities and partner control in a cleaner fish mutualism. *Animal Behaviour, 63*(3), 547–555. <https://doi.org/10.1006/anbe.2001.1937>
- Buckholtz, J. W., & Marois, R. (2012). The roots of modern justice: Cognitive and neural foundations of social norms and their enforcement. *Nature Neuroscience, 15*(5), 655–661. <https://doi.org/10.1038/nn.3087>
- Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage, 12*(3), 314–325. <https://doi.org/10.1006/nimg.2000.0612>
- Clutton-Brock, T. H., & Parker, G. A. (1995). Punishment in animal societies. *Nature, 373*(6511), 209–216. <https://doi.org/10.1038/373209a0>
- Cushman, F., Dreber, A., Wang, Y., & Costa, J. (2009). Accidental outcomes guide punishment in a “trembling hand” game. *PLoS ONE, 4*(8), e6699. <https://doi.org/10.1371/journal.pone.0006699>
- Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition, 108*(2), 353–380. <https://doi.org/10.1016/j.cognition.2008.03.006>
- Cushman, F., Knobe, J., & Sinnott-Armstrong, W. (2008). Moral appraisals affect doing/allowing judgments. *Cognition, 108*(1), 281–289. <https://doi.org/10.1016/j.cognition.2008.02.005>
- Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition, 127*(1), 6–21. <https://doi.org/10.1016/j.cognition.2012.11.008>
- Dahl, A., Schuck, R. K., & Campos, J. J. (2013). Do young toddlers act on their social preferences? *Developmental Psychology, 49*(10), 1964–1970. <https://doi.org/10.1037/a0031460>
- Dunfield, K. A., & Kuhlmeier, V. A. (2010). Intention-mediated selective helping in infancy. *Psychological Science, 21*(4), 523–527. <https://doi.org/10.1177/0956797610364119>
- Eisenbruch, A. B., & Roney, J. R. (2017). The skillful and the stingy: partner choice decisions and fairness intuitions suggest human adaptation for a biological market of cooperators. *Evolutionary Psychological Science, 3*(4), 364–378. <https://doi.org/10.1007/s40806-017-0107-7>
- Farnill, D. (1974). The effects of social-judgment set on children's use of intent information. *Journal of Personality, 42*(2), 276–289. <https://doi.org/10.1111/j.1467-6494.1974.tb00674.x>
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*(6868), 137–140. <https://doi.org/10.1038/415137a>
- 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), 5306. <https://doi.org/10.1038/ncomms6306>
- Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of “theory of mind” in story comprehension. *Cognition, 57*(2), 109–128. [https://doi.org/10.1016/0010-0277\(95\)00692-R](https://doi.org/10.1016/0010-0277(95)00692-R)
- Fruteau, C., Lemoine, S., Hellard, E., van Damme, E., & Noë, R. (2011). When females trade grooming for grooming: Testing partner control and partner choice models of cooperation in two primate species. *Animal Behaviour, 81*(6), 1223–1230. <https://doi.org/10.1016/j.anbehav.2011.03.008>
- Gerstenberg, T., Lagnado, D. A., & Kareev, Y. (2010, January). *The dice are cast: The role of intended versus actual contributions in responsibility attribution*. Proceedings of the 32nd Annual Meeting of the Cognitive Science Society. [http://web.mit.edu/tger/www/papers/TheDiceareCast\(Gerstenberg,Lagnado,Kareev,2010\).pdf](http://web.mit.edu/tger/www/papers/TheDiceareCast(Gerstenberg,Lagnado,Kareev,2010).pdf)
- Gino, F., Moore, D., & Bazerman, M. (2009). *No harm, no foul: The outcome bias in ethical judgments*. Harvard Business School (NOM Working Paper No. 08-080).
- Gino, F., Shu, L. L., & Bazerman, M. H. (2010). Nameless+harmless=blameless: When seemingly irrelevant factors influence judgment of (un)ethical behavior. *Organizational Behavior and Human*

- Decision Processes*, 111(2), 93–101. <https://doi.org/10.1016/j.obhdp.2009.11.001>
- Green, P., & Macleod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498. <https://doi.org/10.1111/2041-210X.12504>
- Gummerum, M., & Chu, M. T. (2014). Outcomes and intentions in children's, adolescents', and adults' second- and third-party punishment behavior. *Cognition*, 133(1), 97–103. <https://doi.org/10.1016/j.cognition.2014.06.001>
- Hackel, L. M., Doll, B. B., & Amodio, D. M. (2015). Instrumental learning of traits versus rewards: Dissociable neural correlates and effects on choice. *Nature Neuroscience*, 18(9), 1233–1235. <https://doi.org/10.1038/nn.4080>
- Hamlin, J. K. (2013). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition*, 128(3), 451–474. <https://doi.org/10.1016/j.cognition.2013.04.004>
- Hamlin, J. K., Mahajan, N., Liberman, Z., & Wynn, K. (2013). Not like me = bad: Infants prefer those who harm dissimilar others. *Psychological Science*, 24(4), 589–594. <https://doi.org/10.1177/0956797612457785>
- Hamlin, J. K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences of the United States of America*, 108(50), 19931–19936. <https://doi.org/10.1073/pnas.1110306108>
- Hebble, P. W. (1971). The development of elementary school children's judgment of intent. *Child Development*, 42(4), 1203–1215. <https://doi.org/10.2307/1127804>
- 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>
- Imamoglu, E. O. (1975). Children's awareness and usage of intention cues. *Child Development*, 46(1), 39–45. <https://doi.org/10.2307/1128831>
- Jordan, J. J., Hoffman, M., Bloom, P., & Rand, D. G. (2016). Third-party punishment as a costly signal of trustworthiness. *Nature*, 530(7591), 473–476. <https://doi.org/10.1038/nature16981>
- Kiley Hamlin, J., Ullman, T., Tenenbaum, J., Goodman, N., & Baker, C. (2013). The mentalistic basis of core social cognition: Experiments in preverbal infants and a computational model. *Developmental Science*, 16(2), 209–226. <https://doi.org/10.1111/desc.12017>
- 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>
- Lench, H. C., Domskey, D., Smallman, R., & Darbor, K. E. (2014). Beliefs in moral luck: When and why blame hinges on luck. *British Journal of Psychology*, 106(2), 1–16. <https://doi.org/10.1111/bjop.12072>
- Li, J., & Tomasello, M. (2018). The development of intention-based socio-moral judgment and distribution behavior from a third-party stance. *Journal of Experimental Child Psychology*, 167, 78–92. <https://doi.org/10.1016/j.jecp.2017.09.021>
- Liddell, T. M., & Kruschke, J. K. (2014). Ostracism and fines in a public goods game with accidental contributions: The importance of punishment type. *Judgement and Decision Making*, 9(6), 523–547. <http://econpapers.repec.org/RePEc:jdm:journl:v:9:y:2014:i:6:p:523-547>
- Lucca, K., Pospisil, J., & Sommerville, J. A. (2018). Fairness informs social decision making in infancy. *PLoS ONE*, 13(2), e0192848. <https://doi.org/10.1371/journal.pone.0192848>
- Margoni, F., & Surian, L. (2016). Explaining the U-shaped development of intent-based moral judgments. *Frontiers in Psychology*, 7, 219. <https://doi.org/10.3389/fpsyg.2016.00219>
- 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), e0125193. <https://doi.org/10.1371/journal.pone.0125193>
- Martin, J. W., & Cushman, F. (2016). The adaptive logic of moral luck. In J. Sytsma & W. Buckwalter (Eds.), *A companion to experimental philosophy* (pp. 190–202). Wiley. <https://doi.org/10.1002/9781118661666.ch12>
- Martin, J. W., Leddy, K., Young, L., & McAuliffe, K. (2020, February 26). An earlier role for intent in children's partner choice versus punishment. <https://osf.io/7539h/>
- Mazzocco, K., & Cherubini, P. (2010). The effect of outcome information on health professionals' spontaneous learning. *Medical Education*, 44(10), 962–968. <https://doi.org/10.1111/j.1365-2923.2010.03744.x>
- Mazzocco, P. J., Alicke, M., & Davis, T. L. (2004). On the robustness of outcome bias: No constraint by prior culpability. *Basic and Applied Social Psychology*, 26(2–3), 131–146. <https://doi.org/10.1207/s15324834basps2602>
- 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>
- McNamara, R. A., Willard, A. K., Norenzayan, A., & Henrich, J. (2019, January). Weighing outcome vs. intent across societies: How cultural models of mind shape moral reasoning. *Cognition*, 182, 95–108. <https://doi.org/10.1016/j.cognition.2018.09.008>
- Miller, R. M., Hannikainen, I. A., & Cushman, F. A. (2014). Bad actions or bad outcomes? Differentiating affective contributions to the moral condemnation of harm. *Emotion*, 14(3), 573–587. <https://doi.org/10.1037/a0035361>
- Misyak, J. B., Melkonyan, T., Zeitoun, H., & Chater, N. (2014). Unwritten rules: Virtual bargaining underpins social interaction, culture, and society. *Trends in Cognitive Sciences*, 18(10), 512–519. <https://doi.org/10.1016/j.tics.2014.05.010>
- Myslinska Szarek, K., Bocian, K., Baryla, W., & Wojciszke, B. (2020). Partner in crime: Beneficial cooperation overcomes children's aversion to antisocial others. *Developmental Science*, 24(2), e13038. <https://doi.org/10.1111/desc.13038>
- Nagel, T. (1979). *Mortal questions*. Cambridge University Press.
- Nobes, G., Panagiotaki, G., & Bartholomew, K. J. (2016). The influence of intention, outcome and question-wording on children's and adults' moral judgments. *Cognition*, 157, 190–204. <https://doi.org/10.1016/j.cognition.2016.08.019>
- Nobes, G., Panagiotaki, G., & Pawson, C. (2009). The influence of negligence, intention, and outcome on children's moral judgments. *Journal of Experimental Child Psychology*, 104(4), 382–397. <https://doi.org/10.1016/j.jecp.2009.08.001>
- Piaget, J. (1932). *The moral judgment of the child*. Kegan Paul. <http://library.brown.edu/find/Record/b1155458>
- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Raihani, N. J., & Barclay, P. (2016). Exploring the trade-off between quality and fairness in human partner choice. *Royal Society Open Science*, 3(11), 160510. <https://doi.org/10.1098/rsos.160510>
- Raihani, N. J., & Bshary, R. (2015). The reputation of punishers. *Trends in Ecology & Evolution*, 30(2), 98–103. <https://doi.org/10.1016/j.tree.2014.12.003>
- Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *NeuroImage*, 19(4), 1835–1842. [https://doi.org/10.1016/S1053-8119\(03\)00230-1](https://doi.org/10.1016/S1053-8119(03)00230-1)
- Schino, G., & Aureli, F. (2015). Reciprocity in group-living animals: Partner control versus partner choice. *Biological Reviews of the Cambridge Philosophical Society*, 92(2), 665–672. <https://doi.org/10.1111/brv.12248>

- Shultz, T., Wright, K., & Schleifer, M. (1986). Assignment of moral responsibility and punishment. *Child Development*, 392, 177–184. <https://doi.org/10.2307/1130649>
- Steckler, C. M., Woo, B. M., & Hamlin, J. K. (2017). The limits of early social evaluation: 9-month-olds fail to generate social evaluations of individuals who behave inconsistently. *Cognition*, 167, 255–265. <https://doi.org/10.1016/j.cognition.2017.03.018>
- Tasimi, A., & Wynn, K. (2016). Costly rejection of wrongdoers by infants and children. *Cognition*, 151, 76–79. <https://doi.org/10.1016/j.cognition.2016.03.004>
- Tostain, M., & Lebreuilly, J. (2008). Rational model and justification model in 'outcome bias.' *European Journal of Social Psychology*, 279, 272–279. <https://doi.org/10.1002/ejsp>
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46(1), 35–57. <https://doi.org/10.1086/406755>
- Van de Vondervoort, J. W., & Hamlin, J. K. (2016). Evidence for intuitive morality: Preverbal infants make sociomoral evaluations. *Child Development Perspectives*, 10(3), 143–148. <https://doi.org/10.1111/cdep.12175>
- Van de Vondervoort, J. W., & Hamlin, J. K. (2017). Preschoolers' social and moral judgments of third-party helpers and hinderers align with infants' social evaluations. *Journal of Experimental Child Psychology*, 164, 136–151. <https://doi.org/10.1016/j.jecp.2017.07.004>
- Van de Vondervoort, J. W., & Hamlin, J. K. (2018). Preschoolers focus on others' intentions when forming sociomoral judgments. *Frontiers in Psychology*, 9, 1851. <https://doi.org/10.3389/fpsyg.2018.01851>
- Warneken, F. (2018). How children solve the two challenges of cooperation. *Annual Review of Psychology*, 69(1), 205–229. <https://doi.org/10.1146/annurev-psych-122216-011813>
- Williams, B. (1981). *Moral luck: Philosophical papers 1973–1980*. Cambridge University Press.
- Woo, B. M., Steckler, C. M., Le, D. T., & Hamlin, J. K. (2017). Social evaluation of intentional, truly accidental, and negligently accidental helpers and harmers by 10-month-old infants. *Cognition*, 168, 154–163. <https://doi.org/10.1016/j.cognition.2017.06.029>
- Woodward, A. L. (2009). Infants' grasp of others' intentions. *Current Directions in Psychological Science*, 18(1), 53–57. <https://doi.org/10.1111/j.1467-8721.2009.01605.x>
- Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, 51(1), 110–116. <https://doi.org/10.1037/0022-3514.51.1.110>
- Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., & Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences of the United States of America*, 107(15), 6753–6758. <https://doi.org/10.1073/pnas.0914826107>
- Young, L., Cushman, F., Adolphs, R., Tranel, D., & Hauser, M. (2006). Does emotion mediate the relationship between an action's moral status and its intentional status? Neuropsychological evidence. *Journal of Cognition and Culture*, 6(1), 1–2. <https://doi.org/10.1163/156853706776931312>
- Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences of the United States of America*, 104(20), 8235–8240. <https://doi.org/10.1073/pnas.0701408104>
- Young, L., Nichols, S., & Saxe, R. (2010). Investigating the neural and cognitive basis of moral luck: It's not what you do but what you know. *Review of Philosophy and Psychology*, 1(3), 333–349. <https://doi.org/10.1007/s13164-010-0027-y>
- Yuill, N., & Perner, J. (1988). Intentionality and knowledge in children's judgments of actor's responsibility and recipient's emotional reaction. *Developmental Psychology*, 24(3), 358–365. <https://doi.org/10.1037/0012-1649.24.3.358>
- Zelazo, P. D., Helwig, C. C., & Lau, A. (1996). Intention, act, and outcome in behavioral prediction and moral judgment. *Child Development*, 67(5), 2478–2492. <https://doi.org/10.2307/1131635>
- Zhou, Y., Jiao, P., & Zhang, Q. (2016). Second-party and third-party punishment in a public goods experiment. *Applied Economics Letters*, 24(1), 54–57. <https://doi.org/10.2139/ssrn.2835884>

Received August 18, 2020

Revision received April 18, 2021

Accepted April 19, 2021 ■