



# Toddlers draw broad negative inferences from wrongdoers' moral violations

Fransisca Ting<sup>a,1,2</sup> and Renée Baillargeon<sup>a,2</sup>

<sup>a</sup>Department of Psychology, University of Illinois at Urbana–Champaign, Champaign, IL 61820

Contributed by Renée Baillargeon, August 10, 2021 (sent for review May 15, 2021; reviewed by Andrew Scott Baron and Jean Decety)

**By 2 y of age, children possess expectations about several different moral principles. Building on these results, we asked whether children who observed a wrongdoer violate a principle would draw negative inferences from this violation about how the wrongdoer was likely to behave in other contexts. In four experiments, 25-mo-old toddlers ( $n = 152$ ) first saw a wrongdoer harm a protagonist. When toddlers judged the wrongdoer's behavior to violate the principle of ingroup support or harm avoidance, they did not find it unexpected if the wrongdoer next violated the principle of fairness by dividing resources unfairly between two other protagonists (Exps. 2 and 3), but they did find it unexpected if the wrongdoer next acted generously by giving another protagonist most of a resource to be shared between them (Exp. 4). When toddlers did not construe the wrongdoer's harmful behavior as a moral violation, these responses reversed: They found it unexpected if the wrongdoer next acted unfairly (Exp. 1) but not if the wrongdoer next acted generously (Exp. 4). Detecting a moral violation thus lowered toddlers' assessment of the wrongdoer's moral character and brought down their expectations concerning the likelihood that the wrongdoer would perform: 1) obligatory actions required by other principles and 2) supererogatory or virtuous actions not required by the principles. Together, these findings expand our understanding of how young children evaluate others' moral characters, and they reveal how these evaluations, in turn, enable children to form sophisticated expectations about others' behavior in new contexts.**

infancy | moral violations | moral character | obligatory actions | supererogatory actions

Social scientists have long been interested in uncovering the basic structure of human moral cognition. An influential proposal has been that a small set of universal moral principles contributes to this structure (1–12). Although researchers disagree widely on the nature, number, and contents of these principles, common candidates include fairness, harm avoidance, ingroup support, and authority. General assumptions about these principles are that they evolved during the millions of years our ancestors lived in small groups of hunter-gatherers, where survival depended on cooperation; that they are normative and determine what actions are obligatory, forbidden, and permissible in specific contexts; and that they are implemented, stressed, and rank-ordered differently in different cultures, resulting in the diverse moral landscape that exists in the world today.

Over the past 10 y, developmental psychology has produced new evidence for this principle-based approach to moral cognition by showing that children under 2.5 y of age (henceforth “young children”) already possess abstract expectations related to each of the four candidate principles (13–26; for reviews, see refs. 27 and 28). With respect to fairness, for example, 4- to 28-mo-olds expected an individual to divide windfall resources equally between two similar potential recipients (13–15, 18, 22, 23); 21-mo-olds expected an individual to reward a worker who had done an assigned chore but not a slacker who had done no work (23); and 17-mo-olds expected a resource acquired by two workers to be shared according to the amount of effort each had

exerted (26). Similarly, with respect to ingroup support, 13- to 29-mo-olds expected an individual to help an ingroup member in need of instrumental assistance (16, 25), but to refrain from helping a wrongdoer who had harmed an ingroup victim (21, 25); during an episode of conflict between two protagonists from different groups, 9- to 18-mo-olds found it unexpected if a member of one group chose to help the protagonist from the other group, instead of the protagonist from the same group (20); and after two groups performed distinct novel conventional actions, 7- to 12-mo-olds found it unexpected if a member of one group chose to imitate the other group's conventional action (19).

Building on these and other findings of early moral expectations, the present research asked the following questions: How closely linked are the different moral principles in the first years of life? Do young children construe the principles as separate, unconnected constraints on individuals' actions, or do they view them, at least to some extent, as interconnected? To get at these questions, we first showed 25-mo-old toddlers a scene in which a wrongdoer violated a principle when interacting with a protagonist. Next, the same wrongdoer interacted with new protagonists in a new scene that called for the application of a different principle. At issue was whether toddlers would expect the wrongdoer to adhere to this second principle. On the one hand, evidence that they did would suggest that they construed the different principles as independent, unconnected constraints on social interactions. After observing a wrongdoer violate one principle, they still expected the wrongdoer to adhere to another principle, as though the transgression they had witnessed had no

## Significance

**According to prior research, 2-y-old toddlers possess expectations about several moral principles. Here, we show that when a wrongdoer violates one of these principles, toddlers draw broad negative inferences from this violation. After a wrongdoer violated ingroup support or harm avoidance when interacting with one individual, toddlers did not find it unexpected if the wrongdoer next violated fairness when interacting with other individuals. This was not because of toddlers' mere reluctance to form expectations about the wrongdoer's behavior, as they found it unexpected if the wrongdoer next acted generously toward another individual. Toddlers who observe a moral violation thus draw negative inferences about the wrongdoer's moral character, which then guide their expectations about the wrongdoer's likely behavior in other situations.**

Author contributions: F.T. and R.B. designed research; F.T. performed research; F.T. analyzed data; and F.T. and R.B. wrote the paper.

Reviewers: A.S.B., University of British Columbia; J.D., University of Chicago.

The authors declare no competing interest.

Published under the [PNAS license](#).

<sup>1</sup>Present address: Department of Psychological and Brain Sciences, Boston University, Boston, MA 02215.

<sup>2</sup>To whom correspondence may be addressed. Email: [fting@bu.edu](mailto:fting@bu.edu) or [rbailiar@illinois.edu](mailto:rbailiar@illinois.edu).

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2109045118/-DCSupplemental>.

Published September 20, 2021.

bearing on their expectations concerning other principles. On the other hand, evidence that toddlers did not find it unexpected if the wrongdoer, having violated the first principle, now violated the second one as well, would suggest that they did view the different principles as interconnected and drew broad negative inferences from the initial transgression they witnessed.

What mechanism might support such inferences? One hypothesis was suggested by evidence that beginning at a young age, children not only attempt to predict how individuals will act toward others but also evaluate their actions, as though assessing their moral characters (29–34; for relevant research with older children and adults, see refs. 35–40). For example, when faced with two distributors, one who had divided resources fairly and one who had done so unfairly, 13- to 17-mo-olds preferred the fair distributor (29–31), 13- to 15-mo-olds associated admonishment (e.g., “She’s a bad girl!”) with the unfair distributor (32), and 10- to 25-mo-olds either gave or expected others to give a treat to the fair as opposed to the unfair distributor (33, 34). These results suggested that upon observing a fairness violation, children evaluated the wrongdoer’s moral character unfavorably, and this evaluation drove their affiliative and punitive attitudes toward the wrongdoer: They were less likely to select the wrongdoer as a target for an affiliative action, they were more likely to select the wrongdoer as a target for a punitive action, and they expected others to do the same. We reasoned that if children were able to use their evaluation of a wrongdoer’s moral character to also consider how the wrongdoer might behave in a new situation, then observing a moral violation might lead them to draw broad negative inferences from this violation.

To flesh out this idea, we speculated that when encountering an unfamiliar individual, young children typically endow the individual with a good moral character, by default [i.e., everyone is assumed to have a good character until proven otherwise (41–43)]; they view it as highly likely that the individual will perform obligatory actions (or avoid forbidden actions) specified by the various moral principles; and they find it unexpected if the individual violates one of the principles, as noted earlier (13–26). In turn, observing such a violation causes children to lower their assessment of the individual’s character and brings down their estimate of the likelihood that the individual will perform other obligatory actions, from highly likely to only somewhat likely. In this scheme, all of the principles thus bear on the common evaluative dimension of good/bad character, making possible broad negative inferences from any one violation. Specifically, children should not find it unexpected if an individual who has violated one principle—and thus seems to be of questionable character—next violates another principle.

To evaluate these speculations, we conducted a series of experiments in which 25-mo-old toddlers first saw a wrongdoer harm a protagonist. Next, the wrongdoer divided resources either fairly or unfairly between two new protagonists. At issue was whether toddlers, when they perceived the harm committed by the wrongdoer as a moral violation, would look about equally at the fair and unfair distributions, suggesting that they did not find it unexpected when the wrongdoer chose to act unfairly. Such evidence would suggest that by 2 y of age, some degree of interconnection is already present among the different moral principles: Children realize that a wrongdoer who violates one principle may violate others as well.

### The Present Research

According to a recent proposal (28), at least two principles, harm avoidance and ingroup support, contribute to young children’s judgments about whether a harmful action constitutes a moral violation. Harm avoidance applies to interactions among members of the same moral circle (e.g., humans) and sets broad limits on the amount of harm a member can inflict on other members. Ingroup support applies to interactions among members of the

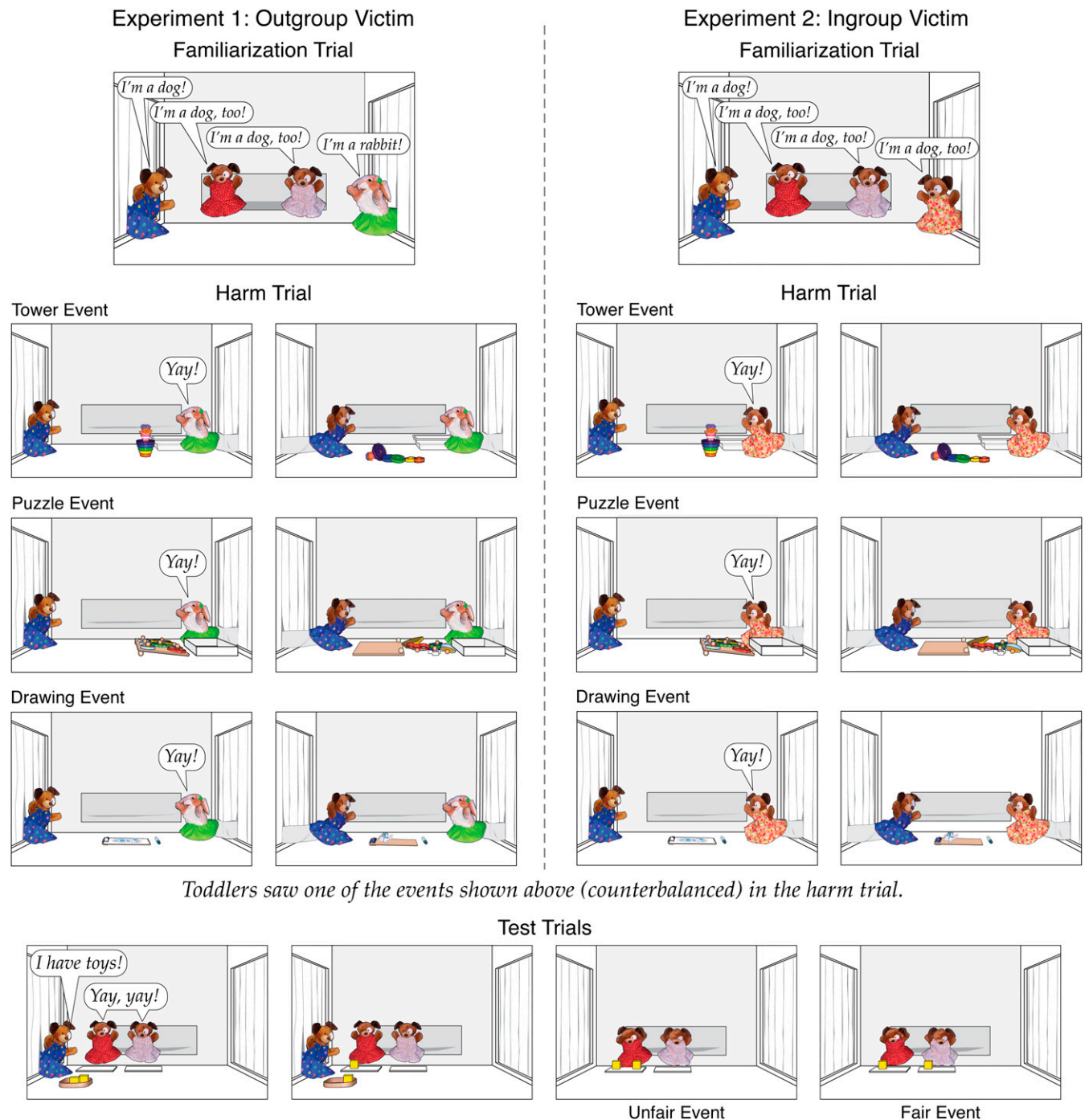
same social group within a moral circle (e.g., sports team) and sets much stricter limits on the amount of harm that can be inflicted on ingroup members. Recent findings are consistent with this proposal. First, when individual A harms another individual B, and there is nothing to suggest that A and B belong to the same moral circle (e.g., they are different nonverbal, non-human characters), young children give no evidence that they view A’s behavior as a moral violation, even if it involves substantial unprovoked harm; for example, 6- to 12-mo-olds detected no violation when A repeatedly hit B or knocked B down a steep hill, causing it to roll end-over-end to the bottom of the hill (44–47). Second, when there are unambiguous cues that A and B belong to the same social group, even slight unprovoked harm is perceived as a violation; for example, 13- to 29-mo-olds detected a violation when A ignored B’s need for instrumental assistance or threw an object B needed on the floor (16, 25, 28). Finally, when there are cues that A and B belong to the same moral circle, but no cues that they belong to the same social group, results are intermediate between those described above, with substantial but not slight unprovoked harm being perceived as a violation (16, 25, 28). We built on these results in deciding what harmful actions to use in our research.

In Exp. 1, which served as a baseline, toddlers were first introduced to two groups of puppets marked by different kinds (dogs vs. rabbits), different labels spoken in female voices by the puppets (“I’m a dog!” vs. “I’m a rabbit!”), and different outfits (dresses vs. skirts). We expected that toddlers would view the dogs and the rabbits 1) as members of the same moral circle, based on their shared capacity for human speech (much like Bert, Ernie, Big Bird, and other verbal puppets on the television show *Sesame Street*, who are akin to “honorary humans”), and also 2) as members of two distinct social groups, based on the multiple cues provided (i.e., different kinds, labels, and outfits) (13, 16, 21, 25). Next, toddlers saw a puppet from one group (the wrongdoer; e.g., a dog) direct a single slightly harmful action at a puppet from the other group (the victim; a rabbit). More specifically, the wrongdoer knocked down a tower built by the victim, flipped over a puzzle filled by the victim, or crumpled a drawing made by the victim. Finally, toddlers saw the wrongdoer divide two items either equally (“fair event”) or unequally (“unfair event”) between two other puppets (recipients). Both recipients belonged to either the same group as the wrongdoer (two dogs; ingroup-recipients condition) or the other group (two rabbits; outgroup-recipients condition). Based on the findings reviewed above, toddlers should perceive the wrongdoer’s slightly harmful action as permissible, because it did not violate the restrictions imposed by the harm-avoidance principle on unprovoked harm to outgroup members. Consequently, toddlers should still expect the wrongdoer to divide the two items equally between the two ingroup or outgroup recipients, in accordance with the fairness principle, and they should find it unexpected when she did not, as in prior research on early sensitivity to fairness (13–15, 18, 22, 23). [Prior findings also suggest that infants who do not construe a wrongdoer’s harmful actions as a moral violation may still draw negative inferences about the wrongdoer’s moral character if these actions are repeatedly contrasted with the helpful, compassionate actions of an altruist (48); we return to these findings in the discussion of Exp. 1.]

Exp. 2 was identical except that the wrongdoer now harmed an ingroup puppet, instead of an outgroup puppet. Based on the findings reviewed above, toddlers should now perceive the wrongdoer’s slightly harmful action as forbidden, because it violated the strict limits imposed by the ingroup-support principle on unprovoked harm to ingroup members (16, 25, 28). We reasoned that if by 2 y of age children already construe the different moral principles as interconnected and recognize that a wrongdoer who fails to adhere to one principle may also fail to adhere to others, then toddlers should no longer find it unexpected when the

wrongdoer divided the two items unequally between the two recipients, in violation of the fairness principle. Moreover, finding the same results whether the recipients were ingroup (ingroup-recipients condition) or outgroup (outgroup-recipients condition) members not only would support the conclusion that toddlers drew broad negative inferences about the wrongdoer, but also would rule out narrower inferences (e.g., perhaps the wrongdoer had a conflictual relationship with her ingroup but otherwise treated others as morally expected).

Exp. 3 was identical to the ingroup-recipient conditions of Exps. 1 and 2 with one exception: Instead of directing a single slightly harmful action at the victim, the wrongdoer now directed three such actions at the victim, in successive trials (tower, puzzle, and drawing). If toddlers judged that all three actions together amounted to substantial harm, then predictions for when the wrongdoer harmed an outgroup victim (outgroup-victim condition) should differ from those in Exp. 1. Because substantial harm to an outgroup member violated the harm-avoidance principle,



**Fig. 1.** Familiarization, harm, and test trials in the ingroup-recipient conditions of Exp. 1 (Left) and Exp. 2 (Right); half the children saw a dog wrongdoer (shown here), and half saw a rabbit wrongdoer. Toddlers in the outgroup-recipients conditions received identical test trials except that the potential recipients were outgroup members. In Exp. 3, toddlers in the outgroup-victim and ingroup-victim conditions were tested using the same procedure as in the ingroup-recipients conditions of Exps. 1 and 2, respectively, except that they received all three harm trials, in the order depicted.

toddlers should no longer find it unexpected if a wrongdoer who had committed such harm subsequently treated ingroup recipients unfairly. In contrast, predictions for when the wrongdoer harmed an ingroup victim (ingroup-victim condition) should be the same as in Exp. 2. If one slightly harmful action was sufficient to violate the ingroup-support principle, then three such actions should do so as well, with the same consequences for toddlers' expectations; in either case, toddlers should not find it unexpected when the wrongdoer next treated ingroup recipients unfairly.

Finally, Exp. 4, which we introduce later, sought to address a possible alternative interpretation of the results of Exps. 2 and 3, and also began to examine additional expectations toddlers might form about individuals with a bad moral character (see *SI Appendix*, Table S1 for a summary of all experiments and conditions).

## Experiment 1

Exp. 1 served as a baseline and tested whether 25-mo-olds who first saw a wrongdoer direct a slightly harmful action at an outgroup victim (an action toddlers should not view as a moral violation) would still expect the wrongdoer to act fairly when dividing resources between two ingroup (ingroup-recipients condition) or outgroup (outgroup-recipients condition) recipients. In each condition, English-speaking toddlers sat on a parent's lap facing a puppet-stage apparatus and received one familiarization trial, one harm trial, and two test trials (Fig. 1). In total, eight puppets were used in Exps. 1 to 3: There were four identical brown dogs who wore dresses in different colors and patterns (D1–D4), and four identical beige rabbits who wore skirts in different solid colors (R1–R4). Each toddler saw a subset of four puppets that included a wrongdoer (D1 or R1), a victim (D2 or R2) and two recipients (D3 and D4 or R3 and R4). In the familiarization trial, the four puppets occupied windows around the three sides of the apparatus: The wrongdoer stood at the left window, the victim stood at the right window, and the two recipients stood at the ends of a wide window in the back wall. All of the windows were filled with curtains, to hide the assistants who operated the puppets. Whether D1 or R1 served as the wrongdoer was counterbalanced within each experiment; for ease of communication, we use D1 in our descriptions.

In the ingroup-recipients condition of Exp. 1, the familiarization trial involved wrongdoer D1, outgroup victim R2, and ingroup recipients D3 and D4. The trial was computer-controlled, followed a second-by-second script, and lasted 12 s. The four puppets labeled themselves in turn, one group at a time, starting from the left or the right (counterbalanced; e.g., D1: "I'm a dog!"; D3: "I'm a dog, too!"; D4: "I'm a dog, too!"; R2, "I'm a rabbit!"). Each puppet tilted from side to side as she spoke, to make clear who was speaking. The trial ended after the last puppet spoke.

The harm trial was computer-controlled, lasted 40 s, and depicted a slightly harmful action in which D1 destroyed a tower, puzzle, or drawing (counterbalanced) produced by R2. In the tower event, R2 was initially alone and stood next to a tray filled with five discs of different colors and sizes; the discs were designed to be stacked by increasing sizes, and the largest disk had a toy bear attached to its top. After R2 stacked the first two discs, one at a time, D1 entered, and the two puppets greeted each other ("Hi!"). Next, while D1 watched, R2 stacked the remaining discs, returned to her window, and clapped twice while admiring her tower ("Yay!"). D1 then approached the tower and hit it, causing it to fall over. She then returned to her window, and the trial ended. The puzzle event was similar, with the following exceptions: R2 had a tray filled with six puzzle pieces, each with a round knob, and she inserted them one by one into a wooden frame; after she completed her puzzle and admired it ("Yay!"), D1 grabbed its top edge, pulled it toward herself, and flipped it over, causing all of the pieces to fall out (the top edge of the puzzle rested on small supports, making it easy for D1 to grab it). Finally, the drawing event was similar to the preceding

events, with the following exceptions: R2 had a marker and a coloring page (the page rested on a clipboard for easy coloring); after R2 completed her drawing and admired it ("Yay!"), D1 crumpled it and tossed it back on the clipboard.

Each test trial had an initial phase and a final phase. The initial phase was computer-controlled, lasted 26 s, and depicted the fair or unfair event, ending with a paused scene; during the final phase, which was infant-controlled, toddlers watched this scene until the trial ended (*Methods*). At the start of the initial phase, only D3 and D4 were present and tilted from side to side in unison. D3 stood at the left end of the back window, as before, and D4 now stood next to her; in front of each puppet was a placemat. Next, D1 entered, carrying a tray with two toys (yellow blocks or purple balls, counterbalanced). She set down the tray and announced, "I have toys!", and D3 and D4 responded enthusiastically, "Yay, yay!" in different voices. D1 placed one toy on one of the placemats (counterbalanced), and then she placed the other toy either on the same placemat (unfair event) or on the other placemat (fair event). Next, D1 picked up her empty tray and left the scene, and D3 and D4 looked down at their placemats and paused until the trial ended. The order of the two test events was counterbalanced.

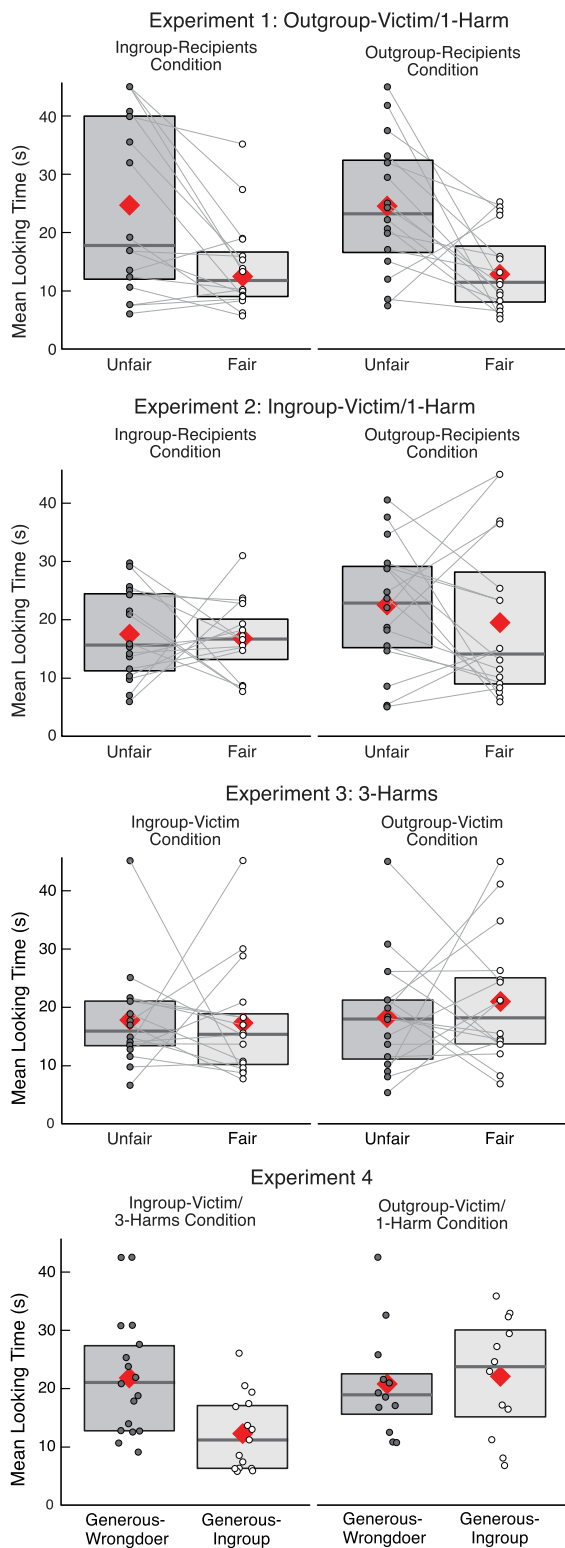
The outgroup-recipients condition was identical except that the two puppets in the back window in the familiarization and test trials were R3 and R4, instead of D3 and D4.

We reasoned that if toddlers 1) viewed the four puppets as members of the same moral circle, based on their shared capacity for human speech, 2) assigned the dog puppets and the rabbit puppets to two distinct social groups, based on the multiple cues provided, and 3) viewed the wrongdoer's single slightly harmful action as permissible (i.e., it was not substantial enough to violate the principle of harm avoidance, and it was directed at an outgroup victim so the principle of ingroup support did not apply), then toddlers should still expect the wrongdoer to adhere to fairness when distributing the toys. Toddlers in both conditions should thus look significantly longer at the unfair than at the fair event, as in prior research on early sensitivity to fairness.

Toddlers ( $n = 32$ ) were highly attentive during the familiarization and harm trials and the initial phases of the test trials (which were all computer-controlled); across conditions, they looked, on average, for 99% of the trials. Looking times during the final phases of the test trials (Fig. 2) were analyzed using an ANOVA with Condition (ingroup- or outgroup-recipients) as a between-subjects factor and Event (unfair or fair) as a within-subject factor. The main effect of Event was significant [ $F(1, 30) = 21.55, P < 0.001, \eta_p^2 = 0.42$ ], but the other effects were not [both  $F_s(1, 30) \leq 0.02, P \geq 0.904, \eta_p^2 = 0.00$ ]. Planned comparisons revealed that: 1) toddlers in the ingroup-recipients condition looked significantly longer at the unfair [mean ( $M$ ) = 24.58,  $SD = 15.20$ ] than at the fair ( $M = 14.16, SD = 7.95$ ) event [ $F(1, 30) = 10.23, P = 0.003, \text{Cohen's } d = 0.86$ ], with 12 of 16 toddlers showing this pattern; and 2) toddlers in the outgroup-recipients condition also looked significantly longer at the unfair ( $M = 24.48, SD = 11.34$ ) than at the fair ( $M = 13.51, SD = 7.06$ ) event [ $F(1, 30) = 11.34, P = 0.002, d = 1.16$ ], with 14 of 16 toddlers showing this pattern. Nonparametric Wilcoxon signed-rank tests confirmed the results of both conditions (with  $z = 2.56, P = 0.010$ , in each condition).

After observing a wrongdoer direct a slightly harmful action at an outgroup victim (i.e., destroy the victim's tower, puzzle, or drawing), toddlers still expected the wrongdoer to divide windfall resources fairly between two ingroup or outgroup recipients. These results suggested that toddlers did not perceive the wrongdoer's harmful action as a moral violation, and hence that they drew no particular inferences from this action about her moral character. As such, these results provided an important baseline: They confirmed, with the present stimuli, events, and procedure, that young children tend to view a slightly harmful





**Fig. 2.** Mean test looking times ( $n = 152$ ), separately by condition and event. In Exps. 1 to 3, which used a within-subject design, connected dots represent individual toddlers, with 16 per condition. Red diamonds indicate means, boxes represent the interquartile ranges between the first and third quartiles, and horizontal lines inside boxes indicate medians. In Exp. 4, which used a between-subjects design, each dot represents an individual toddler; dots are jittered horizontally to improve visibility. In the ingroup-victim/3-harms condition, 17 toddlers saw the generous-wrongdoer event and 15 saw the generous-ingroup event; in the outgroup-victim/1-harm condition, 12 toddlers saw each event.

action directed at an outgroup victim as permissible (16, 25, 28), and that they expect a distributor (who to their knowledge has performed no moral violation) to treat two similar recipients fairly, whether they belong to the distributor's group or not (13–15, 18, 22, 23).

Nevertheless, the main finding of Exp. 1—that toddlers still expected a wrongdoer to treat two recipients fairly after she had inflicted permissible harm on a victim—might, at first sight, appear inconsistent with prior findings from infant experiments conducted with the help/hinder paradigm (44, 46, 48–53). In these experiments, 3- to 21-mo-olds watched events involving three different nonverbal, nonhuman protagonists, A, B, and C; no cues signaled that they belonged to the same social group or the same moral circle. The events included a "help" and a "hinder" event, which were repeated across trials. In the help event, A gave C the instrumental assistance it needed to achieve its goal (e.g., helped C reach the top of a steep hill); in the hinder event, B interfered with C's goal (e.g., knocked C down the hill). Across ages, infants looked equally at the two events, indicating that they did not construe the hinder event as a moral violation (had they done so, they would have looked significantly longer at that event than at the help event). In the absence of cues signaling that B and C belonged to the same social group or the same moral circle, no harm restrictions applied, so infants detected no violation in the hinder event.

Nevertheless, infants provided robust evidence that they evaluated A's action favorably and B's action unfavorably and assessed their moral characters accordingly. First, these evaluations influenced infants' affiliative attitudes: When asked to choose between A and B, 3- to 11-mo-olds preferred A (44, 49–51), and 10- to 12-mo-olds expected C to do the same (44, 46, 52). Second, these evaluations influenced infants' punitive attitudes: 21-mo-olds chose B when asked to take away a treat, but they chose A when asked to give a treat (53). Third, and most relevant for present purposes, these evaluations led infants to draw broad inferences about A's and B's moral characters. In an experiment by Surian et al. (48) with 15-mo-olds, the help and hinder events were followed by a single test event in which A or B divided two items either equally (fair event) or unequally (unfair event) between two potential recipients. When A was the distributor, infants looked significantly longer if shown the unfair as opposed to the fair event; when B was the distributor, however, infants looked equally at the two events, as though they did not find the unfair event unexpected given B's bad character.

How can we reconcile these last results (infants did not find it unexpected when a wrongdoer who had committed permissible harm acted unfairly) with the results of Exp. 1 (toddlers found it unexpected when a wrongdoer who had committed permissible harm acted unfairly)? Salient differences between the two experiments included the participants' ages (15 vs. 25 mo) and the number and nature of the wrongdoer's harmful actions (the hinderer bumped the climber down the hill twice per trial for a total of four bumps; the wrongdoer destroyed one item produced by the victim). However, we suspect that the key difference between the two experiments has to do with the contrast between the hinderer's harmful actions and the helper's compassionate, altruistic actions. Across trials, as in previous experiments using the help/hinder paradigm (44, 46, 49, 50, 52), the hinderer consistently chose to hinder, whereas the helper consistently chose to help, even though this action was not obligatory (there were no cues to indicate that the helper and the climber belonged to the same group). This repeated contrast may have induced infants to compare the helper's and hinderer's actions and to draw inferences about their respective moral characters, even though they detected no moral violation in the hinderer's actions. In the present research, however, no such contrast was available: Toddlers saw only a wrongdoer who directed one or more harmful actions at an ingroup or an outgroup member. Our research examined

whether toddlers would draw broad negative inferences about the wrongdoer's moral character when they construed her behavior as a moral violation, but would draw no such inferences otherwise. The results of Exp. 1 supported this last prediction (toddlers still expected the wrongdoer to act fairly following a harmful action they deemed permissible), setting the scene for the following experiments.

## Experiment 2

Exp. 2 was identical to Exp. 1, with one exception: R2 was replaced with D2, so that D1's slightly harmful action was directed at an ingroup as opposed to an outgroup victim (Fig. 1). We reasoned that if 1) toddlers now perceived D1's action as an ingroup-support violation, as in prior research (16, 25, 28), and 2) this violation lowered their assessment of D1's moral character and their estimate of the likelihood that she would perform other obligatory actions, then toddlers should no longer find it unexpected when D1 acted unfairly in the test trials, and they should therefore look equally at the unfair and fair events. Moreover, obtaining this negative result in both the ingroup- and outgroup-recipients conditions would rule out alternative interpretations of the former condition (e.g., perhaps D1 had a conflictual relationship with her ingroup members and ignored moral principles when interacting with them but still treated others as morally expected). Finding that following D1's slightly harmful action toward D2, toddlers no longer expected D1 to act fairly when dealing with either ingroup members D3 and D4 or outgroup members R3 and R3, would suggest that they drew broad negative inferences about D1's moral character upon observing her initial violation.

Toddlers ( $n = 32$ ) were highly attentive during the familiarization and harm trials and the initial phases of the test trials, looking, on average, for 99% of the trials. Looking times during the final phases of the test trials (Fig. 2) were analyzed as before. The main effect of Event was not significant [ $F(1, 30) = 0.50, P = 0.484, \eta_p^2 = 0.02$ ], nor were the other effects [both  $F_s(1, 30) \leq 2.12, P \geq 0.156, \eta_p^2 \leq 0.07$ ]. Planned comparisons revealed that: 1) toddlers in the ingroup-recipients condition looked equally at the unfair ( $M = 17.49, SD = 7.79$ ) and fair ( $M = 16.66, SD = 6.55$ ) events [ $F(1, 30) = 0.06, P = 0.810, d = 0.12$ ], with 8 of 16 toddlers looking longer at the unfair event; and 2) toddlers in the outgroup-recipients condition also looked equally at the unfair ( $M = 22.26, SD = 10.91$ ) and fair ( $M = 19.65, SD = 13.90$ ) events [ $F(1, 30) = 0.58, P = 0.453, d = 0.21$ ], with 9 of 16 toddlers looking longer at the unfair event. Wilcoxon signed-rank tests confirmed the results of the ingroup-recipients ( $z = 0.31, P = 0.756$ ) and outgroup-recipients ( $z = 0.52, P = 0.605$ ) conditions.

In another ANOVA with Victim as an additional between-subjects factor, we compared test responses in Exp. 1 (outgroup victim) and Exp. 2 (ingroup victim). This analysis yielded a significant main effect of Event [ $F(1, 60) = 13.77, P < 0.001, \eta_p^2 = 0.19$ ] and a significant Victim  $\times$  Event interaction [ $F(1, 60) = 7.20, P = 0.009, \eta_p^2 = 0.11$ ], confirming that toddlers drew different inferences from D1's slightly harmful action depending on whether it was directed at an outgroup or ingroup victim.

Exp. 1 showed that after observing a wrongdoer direct a slightly harmful action at an outgroup victim, toddlers still expected the wrongdoer to divide windfall resources fairly between two ingroup or outgroup recipients. In contrast, Exp. 2 showed that when this same action was directed at an ingroup victim, thereby violating the ingroup-support principle, toddlers no longer found it unexpected if the wrongdoer next treated ingroup or outgroup recipients unfairly. These negative results support a characterization of early moral cognition that posits some degree of interconnection among the different moral principles: Toddlers assumed that a wrongdoer who violated the ingroup-support principle might also violate the fairness principle, suggesting that her initial violation led them to draw broad negative inferences about her moral

character and about the likelihood that she would perform obligatory actions dictated by other principles.

## Experiment 3

In Exps. 1 and 2, the wrongdoer (D1) directed a single slightly harmful action at a victim; in Exp. 3, the wrongdoer committed three such actions, in successive harm trials (i.e., destroyed the victim's tower, puzzle, and drawing, in that order). For half of the toddlers, these actions were directed at an outgroup victim (R2), as in Exp. 1 (outgroup-victim condition); for the other toddlers, these actions were directed at an ingroup victim (D2), as in Exp. 2 (ingroup-victim condition). In both conditions, the three harm trials were followed by the same test trials as in the ingroup-recipient conditions of Exps. 1 and 2, with the wrongdoer dividing two toys between two ingroup members (D3 and D4).

We reasoned that in the ingroup-victim condition, results should be the same as in Exp. 2: If a single slightly harmful action directed at an ingroup victim was sufficient for toddlers to detect an ingroup-support violation and, consequently, to no longer expect the wrongdoer to act fairly, then three such actions should, a fortiori, have the same effect. In the outgroup-victim condition, however, results might differ from those in Exp. 1. If toddlers took all three harmful actions to amount to substantial harm, then they should judge these actions to violate the harm-avoidance principle (recall that this principle sets limits on the amount of harm that can be inflicted on other members of a moral circle). Furthermore, if, as in Exp. 2, this violation lowered toddlers' evaluation of the wrongdoer's moral character, thereby bringing down their estimate of the likelihood that she would perform other obligatory actions, then toddlers should no longer find it unexpected when the wrongdoer acted unfairly in the test trials. Toddlers in both the ingroup- and outgroup-victim conditions were thus expected to look equally at the fair and unfair events.

As before, toddlers ( $n = 32$ ) were highly attentive during the familiarization and harm trials and the initial phases of the test trials, looking, on average, for 99% of the trials. Looking times during the final phases of the test trials (Fig. 2) were analyzed using an ANOVA with Condition (outgroup- or ingroup-victim) as a between-subjects factor and Event (unfair or fair) as a within-subject factor. The main effect of Event was not significant [ $F(1, 30) = 0.26, P = 0.613, \eta_p^2 = 0.01$ ], nor were the other effects [both  $F_s(1, 30) \leq 1.49, P \geq 0.231, \eta_p^2 \leq 0.05$ ]. Planned comparisons confirmed that: 1) toddlers in the outgroup-victim condition looked equally at the unfair ( $M = 18.27, SD = 9.88$ ) and fair ( $M = 21.04, SD = 11.24$ ) events [ $F(1, 30) = 0.55, P = 0.466, d = -0.26$ ], with 6 of 16 toddlers looking longer at the unfair event; and 2) toddlers in the ingroup-victim condition also looked equally at the unfair ( $M = 16.73, SD = 9.09$ ) and fair ( $M = 16.66, SD = 10.28$ ) events [ $F(1, 30) = 0.00, P = 0.987, d = 0.01$ ], with 9 of 16 toddlers looking longer at the unfair event. Wilcoxon signed-rank tests confirmed the results of the outgroup-victim ( $z = 0.63, P = 0.532$ ) and ingroup-victim ( $z = 0.41, P = 0.679$ ) conditions.

In additional ANOVAs, we compared test responses across Exps. 1 to 3, focusing on harm to outgroup and ingroup victims separately. For conditions involving harm to an outgroup victim, we first compared the outgroup-victim condition of Exp. 3 (three harmful actions) to the ingroup-recipients condition of Exp. 1 (one harmful action; recall that both of these conditions involved ingroup recipients). This ANOVA ( $n = 32$ ) used Harm (3 or 1) as a between-subjects factor and Event (fair or unfair) as a within-subject factor. The Harm  $\times$  Event interaction was significant [ $F(1, 30) = 7.43, P = 0.011, \eta_p^2 = 0.20$ ], and the same result was found when the outgroup-recipients condition in Exp. 1 was substituted for the ingroup-recipients condition [ $F(1, 30) = 7.24, P = 0.012, \eta_p^2 = 0.19$ ]. Turning to conditions involving harm to an ingroup victim, we first compared the ingroup-victim condition of Exp. 3 (three harmful actions) to the ingroup-recipients

condition of Exp. 2 (one harmful action). The Harm  $\times$  Event interaction was not significant [ $F(1, 30) = 0.03, P = 0.873, \eta_p^2 = 0.00$ ], and the same result was found when the outgroup-recipients condition in Exp. 2 was substituted for the ingroup-recipients condition [ $F(1, 30) = 0.23, P = 0.638, \eta_p^2 = 0.01$ ].

Together, the data of Exps. 1 to 3 supported three conclusions. First, when the wrongdoer directed one or more slightly harmful actions at an ingroup victim, toddlers judged the wrongdoer's behavior to violate the ingroup-support principle, and they did not find it unexpected if she next violated the fairness principle when dealing with other protagonists. Second, when the wrongdoer directed three slightly harmful actions at an outgroup victim, toddlers judged the wrongdoer's behavior to violate the harm-avoidance principle, and they again did not find it unexpected if she next acted unfairly. Finally, when the wrongdoer directed a single slightly harmful action at an outgroup victim, toddlers did not perceive this action as a moral violation, and they found it unexpected if the wrongdoer next acted unfairly. These results indicate that by 2 y of age, toddlers already draw broad negative inferences from the moral violations they observe, suggesting that some degree of interconnection is already present among the different principles.

#### Experiment 4

We have argued that toddlers in Exps. 2 and 3 looked equally at the fair and unfair events they were shown because: 1) they perceived the wrongdoer's harmful actions as a violation of the ingroup-support or harm-avoidance principle; 2) this violation lowered their assessment of her moral character; and 3) this unfavorable assessment, in turn, led them to not find it unexpected when she chose to violate the fairness principle. However, our results were open to another, subtly different interpretation. It could be that upon detecting the wrongdoer's moral violation in the harm trials, toddlers simply concluded that her actions were unpredictable, causing them to refrain from forming any further expectations about her behavior. Exp. 4 sought to rule out this alternative interpretation.

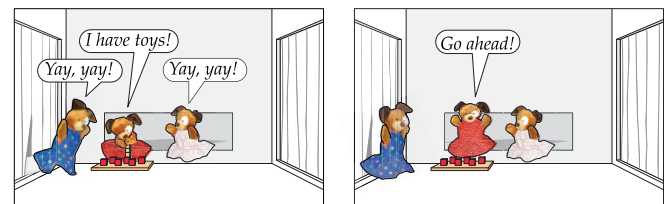
To do so, we needed a situation in which toddlers would find it unexpected if a wrongdoer who first violated a moral principle next performed a particular action. We turned to supererogatory actions, which are virtuous actions that go beyond what is required by the moral principles and that signal excellences of character, such as compassion, generosity, and bravery (28, 54, 55). Previous research suggests that whereas young children view individuals of good character as highly likely to perform obligatory actions, as was discussed previously, they view them as only somewhat likely to perform supererogatory actions (16, 24, 25, 28). For example (16), 17-mo-old infants found it unexpected when an unfamiliar individual chose to ignore an ingroup member in need of instrumental assistance, suggesting that they judged helping the ingroup member to be an obligatory action (dictated by the ingroup-support principle) and assumed the individual was highly likely to perform that action. In contrast, infants did not find it unexpected when the individual chose to ignore an outgroup member with the same need, suggesting that they judged helping the outgroup member to be a supererogatory, compassionate action and assumed the individual was only somewhat likely to perform that action. We speculated that if observing a wrongdoer violate a moral principle brought down toddlers' expectations not only about the likelihood that the wrongdoer would perform other obligatory actions (from highly likely to somewhat likely), as shown in Exps. 2 and 3, but also about the likelihood that the wrongdoer would perform supererogatory actions (from somewhat likely to very unlikely), then toddlers should find it unexpected if the wrongdoer next performed a supererogatory action.

In line with these speculations, Exp. 4 asked whether toddlers who first saw a wrongdoer direct three harmful actions at an

ingroup member (ingroup-victim/3-harms condition) would find it unexpected if the wrongdoer next gave another ingroup member most of a resource to be shared between them (a supererogatory, generous action). Evidence that toddlers found the wrongdoer's generosity unexpected would rule out the suggestion that children in Exps. 2 and 3 merely viewed the wrongdoer as unpredictable and refrained from forming expectations about her behavior.

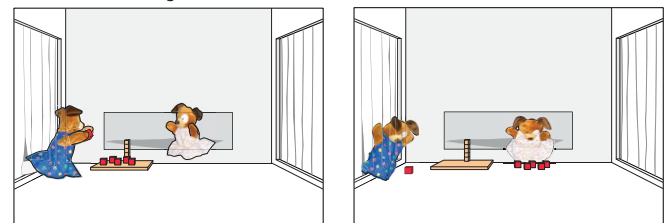
Toddlers first received the same three harm trials as in the ingroup-victim condition of Exp. 3. Next, they received a single test trial in which they saw either a generous-wrongdoer or a generous-ingroup event (Fig. 3). Each event had a computer-controlled initial phase that lasted 22 s, followed by an infant-controlled final phase. At the start of the initial phase in the generous-wrongdoer event, only wrongdoer D1 and ingroup member D4 were present, in the same places as in the familiarization trial. They tilted left and right until they were joined by D3, who entered the apparatus (in her usual place, between D1 and D4) carrying a tray with six identical toys (red blocks). D3 set the tray down and announced, "I have toys!", and D1 and D4 both answered excitedly, "Yay, yay!" Next, D3 let go of the tray, opened her arms, and said, "Go ahead!", as though inviting D1 and D4 to share the toys. She then exited the apparatus. As soon as she left, D1 quickly lunged over the tray, picked up the toy in the farthest corner, placed it in front of herself, and then pushed the remaining five toys in front of D4. The two puppets then looked down and paused, and toddlers watched this scene until the trial ended. The generous-ingroup event was identical except

#### Experiment 4 Test Trial



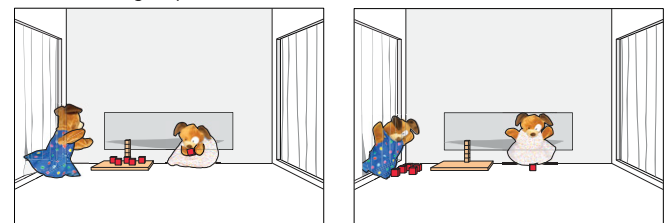
Each event began as shown above and continued as shown below.

#### Generous-Wrongdoer Event



The wrongdoer takes one block then slides the other five to her ingroup member.

#### Generous-Ingroup Event



The ingroup member takes one block then slides the other five to the wrongdoer.

**Fig. 3.** Test trial in Exp. 4. Toddlers in the ingroup-victim/3-harms condition first received the same familiarization and harm trials as in the ingroup-victim condition of Exp. 3; toddlers in the outgroup-victim/1-harm condition first received the same familiarization and harm trials as in Exp. 1.

that D4 lunged over the tray, took the toy in the farthest corner, and gave the remaining five toys to D1.

Before we outline our predictions, three comments are in order. First, because D3 and D4 were absent in the three harm trials when D1 destroyed D2's tower, puzzle, and drawing, they should be ignorant about D1's harmful actions; toddlers should therefore not expect D3 and D4 to adopt a punitive attitude toward D1 (23, 56). Second, while young children generally expect a distributor to divide a resource equally between two similar recipients (13–15, 18, 22, 23), they recognize that an individual who is sharing a resource with an ingroup member may elect to give more than an equal share (24, 26); toddlers should thus view generosity to an ingroup member as morally permissible. Third, Exp. 4 used a between-subjects design, instead of a within-subject design, to avoid interpretive issues that might arise across test trials. We were concerned that after toddlers saw one puppet give away five of the six toys in the first trial, seeing the other puppet do the same in the second trial would be open to multiple interpretations (e.g., she might want to imitate her ingroup member, she might want to achieve an equal distribution of the toys across trials, or she might want to reciprocate by also acting generously). To avoid these ambiguities, each toddler received a single test trial. Moving on to our predictions, we reasoned that if toddlers viewed ingroup member D4 as somewhat likely to act generously, but wrongdoer D1 as very unlikely to do so, then they should look significantly longer if shown the generous-wrongdoer as opposed to the generous-ingroup event.

Toddlers ( $n = 32$ ) were highly attentive during the familiarization and harm trials and the initial phase of the test trial, looking, on average, for 98% of the trials. Looking times in the final phase of the test trial were subjected to an ANOVA with Event (generous-wrongdoer or generous-ingroup) as a between-subjects factor. This effect was significant [ $F(1, 30) = 9.85, P = 0.004, d = 1.13$ ], indicating that toddlers looked significantly longer if shown the generous-wrongdoer event ( $M = 23.16, SD = 10.89$ ) as opposed to the generous-ingroup event ( $M = 12.89, SD = 6.89$ ). A nonparametric Wilcoxon rank-sum test confirmed this result ( $z = 2.85, P = 0.004$ ).

After they saw the wrongdoer harm an ingroup victim three times, toddlers found it unexpected if the wrongdoer next acted generously toward another ingroup member. To provide evidence that this response arose only when the wrongdoer's generosity came on the heels of a moral violation, additional toddlers were tested in a final, identical condition except that the wrongdoer harmed an outgroup victim once, as in Exp. 1 (outgroup-victim/1-harm condition). If toddlers again viewed this single harmful action as permissible and drew no broad negative inferences from it, then they should not find it unexpected when the wrongdoer next chose to act generously. Toddlers ( $n = 24$ ) were highly attentive during the familiarization and harm trials and the initial phase of the test trial, looking, on average, for 99% of the trials. Looking times during the final phase of the test trial were analyzed as above. The main effect of Event was not significant [ $F(1, 22) = 0.11, P = 0.741, d = -0.14$ ], indicating that toddlers looked about equally whether they were shown the generous-wrongdoer event ( $M = 21.87, SD = 9.85$ ) or the generous-ingroup event ( $M = 23.27, SD = 10.65$ ). A Wilcoxon rank-sum test confirmed this result ( $z = 0.46, P = 0.644$ ). Finally, to compare the two conditions of Exp. 4, we conducted an ANOVA similar to that above with Condition (ingroup-victim/3-harms or outgroup-victim/1-harm) as an additional between-subjects factor. The Condition  $\times$  Event interaction was significant [ $F(1, 52) = 4.97, P = 0.030, \eta_p^2 = 0.09$ ], indicating that toddlers found the wrongdoer's generosity unexpected when it was preceded by harm they construed as a moral violation (ingroup-victim/3-harms condition), but not otherwise (outgroup-victim/1-harm condition).

The results of Exp. 4 thus complemented those of Exps. 1 to 3 and supported their interpretations. When toddlers judged a

wrongdoer's harmful behavior to be a moral violation, this caused them to lower their assessment of her moral character and their expectations concerning the likelihood that she would perform other obligatory actions (from highly likely to somewhat likely) or supererogatory actions (from somewhat likely to very unlikely). As a result, toddlers did not find it unexpected if she next acted unfairly (Exps. 2 and 3), but they did find it unexpected if she next acted generously (Exp. 4). However, when toddlers did not view the wrongdoer's harmful behavior as a moral violation, they apparently drew no negative inferences from her behavior and showed the reverse responses: They found it unexpected if she next acted unfairly (Exp. 1), but not if she next acted generously (Exp. 4).

## General Discussion

Research over the past decade has revealed that young children hold expectations about several different moral principles, including ingroup support, harm avoidance, and fairness (13–16, 18–23, 25, 26, 29–34, 48). The present research asked whether children who observe a wrongdoer violate one of these principles draw broad negative inferences about the wrongdoer's moral character, which then guide their expectations about the wrongdoer's likely behavior in other social situations.

Exp. 1 served as a baseline and showed that after a wrongdoer directed a single slightly harmful action at an outgroup victim (an action young children typically do not perceive as a moral violation), 25-mo-old toddlers still expected the wrongdoer to act fairly when dividing windfall resources between two ingroup or outgroup recipients. In Exp. 2, the same harmful action was directed at an ingroup victim, violating the principle of ingroup support, and toddlers no longer found it unexpected when the wrongdoer next acted unfairly toward ingroup or outgroup recipients. In Exp. 3, three harmful actions were shown. When these were directed at an ingroup victim, toddlers again detected an ingroup-support violation and, consistent with the results of Exp. 2, did not find it unexpected if the wrongdoer next treated ingroup recipients unfairly. When these harmful actions were directed at an outgroup victim, toddlers judged these actions to amount to substantial harm, thereby violating the harm-avoidance principle, and they no longer found it unexpected if the wrongdoer next treated ingroup recipients unfairly. Exps. 2 and 3 thus provided converging evidence that toddlers did not find it unexpected if a wrongdoer who violated one principle next violated another. Complementing these results, Exp. 4 showed that following three harmful actions to an ingroup victim, toddlers again detected an ingroup-support violation and found it unexpected if the wrongdoer next generously shared a resource with another ingroup member, by giving her most of the resource. Finally, in line with the results of Exp. 1, toddlers did not find the wrongdoer's generosity unexpected if she first directed a single slightly harmful action at an outgroup victim. Together, the results of Exps. 1 to 4 thus indicated that after observing a wrongdoer violate the ingroup-support or harm-avoidance principle when interacting with one protagonist, toddlers did not find it unexpected if she next violated the fairness principle when interacting with other protagonists, but they did find it unexpected if she next behaved generously toward another protagonist.

Our results provide further evidence for several prior findings in early moral cognition. In particular, they confirm that: 1) toddlers use available cues to assign unfamiliar individuals to social groups (13, 16, 19–21, 25, 29); 2) all other things being equal, toddlers expect individuals to divide windfall resources fairly between ingroup or outgroup recipients (in accordance with fairness) (13–15, 18, 22, 23), to refrain from harming ingroup members (in accordance with ingroup support) (16, 25, 28), and to comply with restrictions on harm to outgroup members (in accordance with harm avoidance) (28); and 3) toddlers do not expect individuals who are unaware of harm inflicted by a wrongdoer



(because they were absent when it occurred) to adopt a punitive attitude toward the wrongdoer (25, 56).

In addition, our results break new ground in two important ways: by showing that by 2 y of age, young children already draw broad negative inferences from the moral violations they observe, and by suggesting a mechanism that explains the basis of these inferences. When young children encounter an unfamiliar individual, they assume, by default, that this individual possesses a good moral character; as such, they view the individual as highly likely to perform obligatory actions dictated by the moral principles. At the same time, children do not assume (without evidence to this effect) that the individual possesses excellences of character, such as compassion, generosity, and bravery; they therefore view the individual as only somewhat likely to perform supererogatory actions that go beyond what is required by the principles. If the individual then fails to perform an obligatory action, this violation causes children to lower their evaluation of the individual's moral character, which in turn brings down their expectations concerning the likelihood that the individual will perform other obligatory actions (now viewed as only somewhat likely) as well as supererogatory actions (now viewed as very unlikely). Our results thus make clear that a complete account of early moral cognition must include not only the different moral principles that help children predict individuals' actions, but also a basic understanding of moral character that includes concepts of obligatory and supererogatory actions and considers the likelihood of these actions for good vs. bad individuals.

Future research can build on our results in several directions. First, it will be important to confirm our results using a variety of scenarios that weave together different obligatory and supererogatory actions. For example, recall that young children look significantly longer if an individual ignores as opposed to helps an ingroup member in need of assistance, but look about equally if the individual ignores or helps an outgroup member with the same need (16). The present research suggests that if children first saw the individual divide resources unfairly between two similar recipients, both looking patterns would change. Children should now look about equally whether the individual helped or ignored an ingroup member in need, and they should look significantly longer if the individual helped as opposed to ignored an outgroup member in need. In other words, following a wrongdoer's fairness violation, children should view an obligatory action dictated by the ingroup-support principle as only somewhat likely, and they should view a supererogatory, compassionate action toward an outgroup member as very unlikely.

Second, future research can explore whether observing a wrongdoer commit more or more varied moral violations would lead children to lower their expectations concerning the likelihood that the wrongdoer would perform other obligatory actions from very likely to very unlikely (as opposed to somewhat likely). In Exp. 3, for example, after toddlers saw the wrongdoer violate the principle of harm avoidance by directing three harmful actions at an outgroup victim, they tended to look equally whether she acted fairly or unfairly toward ingroup members. What if the wrongdoer first committed a more severe violation, by directing more harmful actions at the victim or by harming a larger number of victims (35)? Would children now expect her to act unfairly in the test trials, and hence would they look significantly longer if she acted fairly instead? Another approach would be to show children a wrongdoer who violated two different principles. For example, what if a wrongdoer failed to distribute resources fairly between two outgroup members and then harmed another outgroup member by destroying her tower, puzzle, and drawing? Would children infer that a wrongdoer who violated two principles (fairness and harm avoidance) was very likely to also violate a third (e.g., ingroup support or authority)? For example, would children look significantly longer if the wrongdoer chose to help, as opposed to ignore, an ingroup member in need of assistance?

Third, another research direction will be to explore what information might lead children to reverse their negative evaluations of wrongdoers. For example, would genuine signs of contrition by the wrongdoer, perhaps accompanied by explicit forgiveness from the victim, lead children to revert to baseline expectations about the wrongdoer's likely behavior in the same context or in new contexts (57)?

Finally, and more generally, the present experimental approach may be used to investigate whether young children who observe a moral violation restrict their negative inferences about the wrongdoer to the moral domain or extend these inferences to others areas of social cognition. For example, there is substantial evidence that a principle of rationality (together with corollaries of consistency and efficiency) guides early psychological reasoning about agents' actions (58, 59). Would young children who first observed a wrongdoer commit a moral violation no longer find it unexpected if the wrongdoer next acted irrationally when pursuing an object-directed goal (e.g., searched for an object in a visibly empty container or retrieved an object using an inefficient route)? Such evidence would suggest that, at least initially, young children draw far-reaching inferences from the moral violations they observe, inferences that go well beyond the moral domain. On the other hand, evidence that children still expected wrongdoers to act rationally when pursuing goals would provide tantalizing evidence that the moral and psychological domains are conceptually distinct from an early age.

## Conclusion

The present research showed that by 2 y of age, children draw broad negative inferences from the moral violations they observe. After seeing a wrongdoer violate the principle of ingroup support or harm avoidance when interacting with a protagonist, toddlers did not find it unexpected if the wrongdoer next acted unfairly toward other protagonists, but they did find it unexpected if the wrongdoer next acted generously toward another protagonist. These findings expand our understanding of how young children evaluate others' moral characters and demonstrate how these evaluations, in turn, enable children to form sophisticated and nuanced expectations about others' behavior in new contexts.

## Methods

**Power Analyses.** Exps. 1 to 3 each had a 2 (Conditions)  $\times$  2 (Events) within-subject design, with Event as a repeated factor. To estimate the appropriate sample size of each experiment, we relied on a prior report on early sensitivity to fairness that also used a within-subject design (23). A 2 (Conditions)  $\times$  2 (Events) ANOVA, with a sample of 32, yielded a Condition  $\times$  Event effect size ( $\eta_p^2$ ) of 0.14. A G\*Power analysis (60) based on this value, with  $\alpha$  set at 0.05 and power set at 0.80, suggested that the minimum total sample size per experiment was at least 22. We tested 32 participants in each experiment, with 16 per condition, as in this prior report. Exp. 4 had a 2 (Conditions)  $\times$  2 (Events) between-subjects design, and here we relied on a prior report on early sensitivity to ingroup support that also used a between-subjects design (16). A series of 2 (Conditions)  $\times$  2 (Events) ANOVAs, each with a sample of 32, yielded an average Condition  $\times$  Event effect size ( $\eta_p^2$ ) of 0.19. A G\*Power analysis based on this value suggested that the minimum total sample size in Exp. 4 was at least 36. We tested 56 participants, 32 in the ingroup-victim/3-harms condition and 24 in the final, outgroup-victim/1-harm condition (we had planned to again test 32 but data collection was shut down by the COVID-19 pandemic).

**Participants.** Participants were 152 English-speaking toddlers (75 male;  $M = 25$  mo, 5 d, range = 21 mo, 18 d to 29 mo, 10 d). Another nine toddlers were excluded, five because they were fussy (four) or distracted (one), two because they experienced parental interference, and two because their test looking time was over 3 SDs from the condition mean (both were in the ingroup-victim/3-harms condition of Exp. 4 and saw the generous-ingroup event). In each condition of Exps. 1 to 3, about half of the toddlers saw the unfair event first, and the others saw the fair event first. In the ingroup-victim/3-harms condition of Exp. 4, 17 toddlers saw the generous-wrongdoer

event and 15 saw the generous-ingroup event; in the outgroup-victim/1-harm condition, 12 toddlers saw each event. In all four experiments, each toddler's parent gave written informed consent, and the protocol was approved by the Institutional Review Board of the University of Illinois at Urbana-Champaign.

**Apparatus.** The apparatus consisted of a brightly lit display booth (201-cm high × 101-cm wide × 58-cm deep) with a large opening (56 × 95 cm) in its front wall; between trials, a supervisor lowered a curtain in front of this opening. Inside the apparatus, the floor and side walls were white and the back wall was blue. The wrongdoer and the victim were introduced into the apparatus through side windows (each 51 × 38 cm and filled with a white fringe curtain), and the other protagonists were introduced through a back window (18 × 61 cm and filled with stretchy blue fabric divided by a horizontal slit). Four dog puppets and four rabbit puppets were used across experiments. The dogs (each about 24 × 24 × 18 cm at the largest points) were brown with a white circular patch around the left eye, and they wore different patterned dresses; the wrongdoer and the victim wore a blue dress with scattered dots and a cream dress with overlapping dots (counterbalanced), and the other two protagonists wore a red dress with small flowers (left dog) and a white dress with small lines (right dog). The rabbits (each about 24 × 20 × 18 cm) were beige with a white lower face and stomach, and they wore solid color skirts and matching hair clips; the wrongdoer and the victim wore purple and green skirts (counterbalanced), and the other two protagonists wore yellow (left rabbit) and pink (right rabbit) skirts. The puppets were operated by three female assistants; one (in a white shirt) knelt at the left window and operated the wrongdoer, one (also in a white shirt) knelt at the right window and operated the victim, and one (in a blue shirt) sat at the back window and operated the other two protagonists. Behind the assistants, floor-to-ceiling white (side assistants) or blue (back assistant) curtains hid the testing room from view. In the familiarization trial, when the two protagonists at the back spoke in turn, the back assistant used a high voice for one protagonist and a normal voice for the other; in the test trials of Exps. 1 to 3, when the two protagonists cheered together, the back assistant and the supervisor spoke in unison.

Stimuli in the harm trials included a tower (consisting of five colorful discs) and a clear rectangular tray; a wooden puzzle depicting a street scene (containing six puzzle pieces) and a white rectangular tray; and a green marker and a beige rectangular clipboard with a coloring page depicting a flower (the clipboard was affixed to the apparatus floor by Velcro strips, to make it easier for the wrongdoer to take the drawing). Stimuli in the test trials of Exps. 1 to 3 included two rectangular placemats with a granite-pattern adhesive paper, a beige rectangular tray, and either two yellow blocks or two purple balls (with flat bottoms to prevent rolling); the blocks were used with the dog recipients and the balls with the rabbit recipients, to stand out better against their clothes. Stimuli in Exp. 4 included six small red blocks and an edgeless rectangular tray that was covered with a wood-

pattern adhesive paper and had a thin vertical handle at the back. For better control as events unfolded, the tray had a thin metal plate under its adhesive paper, the blocks had magnets that kept them in place as the tray was moved, and the tray was deposited on Velcro strips on the apparatus floor. The blocks were placed on the tray in two staggered rows of three, to make it easy for toddlers to see how many there were.

During each testing session, two cameras captured images of the toddler and events; the two images were combined, projected onto a monitor located behind the apparatus, and checked by the supervisor to confirm that the events followed the prescribed scripts. Recorded sessions were also checked off-line for experimenter accuracy.

**Procedure.** Toddlers sat on a parent's lap in front of the apparatus; parents were instructed to close their eyes in the test trials and to remain silent and neutral in all trials. During a test trial, the toddler's looking behavior was monitored by two hidden observers who were naïve about which event was shown in the trial; looking times were computed using the primary observer's responses. During familiarization and harm trials, the primary observer was absent from the testing room and was thus also naïve about which puppets were present and how many harmful actions the wrongdoer inflicted on the victim.

The final phase of a test trial ended when toddlers either: 1) looked away for 1 consecutive second after having looked for at least 5 cumulative seconds or 2) looked for a maximum of 45 cumulative seconds. The 5-s minimum-look criterion gave toddlers more time to process the events shown in the initial phase before the final phase could end (when actions stop and a paused scene begins, children sometimes look away briefly, and a minimum-look criterion allows them to return to the scene and continue processing it). Across all test trials in the present research, toddlers took 5.22 s (SD = 0.78), on average, to complete the 5-s minimum-look. Finally, interobserver agreement during the final phase of a test trial was calculated by dividing the number of 100-ms intervals in which the two observers agreed by the total number of intervals in the final phase. Agreement in the two test trials of Exps. 1 to 3 was calculated for 91 of 96 toddlers (only one observer was present for the other toddlers) and averaged 96% per trial; agreement in the test trial of Exp. 4 was calculated for 55 of 56 toddlers and averaged 95% per trial.

**Data Availability and Preliminary Analyses.** For the data from all four experiments, see [Dataset S1](#); for the preliminary analyses of the test data in each experiment, see [SI Appendix, Table S2](#).

**ACKNOWLEDGMENTS.** We thank the University of Illinois at Urbana-Champaign Infant Cognition Laboratory for their help with the data collection; graphic artist Steve Holland for producing the figures; and the families who participated in the experiments. This research was supported by a grant from the John Templeton Foundation (to R.B.).

1. M. B. Brewer, The psychology of prejudice: Ingroup love and outgroup hate? *J. Soc. Issues* 55, 429–444 (1999).
2. C. T. Dawes, J. H. Fowler, T. Johnson, R. McElreath, O. Smirnov, Egalitarian motives in humans. *Nature* 446, 794–796 (2007).
3. E. Dupoux, P. Jacob, Universal moral grammar: A critical appraisal. *Trends Cogn. Sci.* 11, 373–378 (2007).
4. S. Dwyer, Moral dumbfounding and the linguistic analogy: Methodological implications for the study of moral judgment. *Mind Lang.* 24, 274–296 (2009).
5. E. Fehr, I. Schurtenberger, Normative foundations of human cooperation. *Nat. Hum. Behav.* 2, 458–468 (2018).
6. J. Graham et al., Moral foundations theory: The pragmatic validity of moral pluralism. *Adv. Exp. Soc. Psychol.* 47, 55–130 (2013).
7. R. Jackendoff, *Language, Consciousness, Culture: Essays on Mental Structure* (MIT Press, Cambridge, MA, 2007).
8. T. S. Rai, A. P. Fiske, Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychol. Rev.* 118, 57–75 (2011).
9. R. A. Shweder, N. C. Much, M. Mahapatra, L. Park, "The 'big three' of morality (autonomy, community, and divinity) and the 'big three' explanations of suffering" in *Morality and Health*, A. Brandt, P. Rozin, Eds. (Routledge, New York, NY, 1997), pp. 119–169.
10. H. Tajfel, M. G. Billig, R. P. Bundy, C. Flament, Social categorization and intergroup behaviour. *Eur. J. Soc. Psychol.* 1, 149–178 (1971).
11. J. Tooby, L. Cosmides, M. E. Price, Cognitive adaptations for *n*-person exchange: The evolutionary roots of organizational behavior. *MDE. Manage. Decis. Econ.* 27, 103–129 (2006).
12. M. Van Vugt, R. Hogan, R. B. Kaiser, Leadership, followership, and evolution: Some lessons from the past. *Am. Psychol.* 63, 182–196 (2008).
13. L. Bian, S. Sloane, R. Baillargeon, Infants expect ingroup support to override fairness when resources are limited. *Proc. Natl. Acad. Sci. U.S.A.* 115, 2705–2710 (2018).
14. M. Buyukozer Dawkins, S. Sloane, R. Baillargeon, Do infants in the first year of life expect equal resource allocations? *Front. Psychol.* 10, 116 (2019).
15. E. A. Enright, H. Gweon, J. A. Sommerville, 'To the victor go the spoils': Infants expect resources to align with dominance structures. *Cognition* 164, 8–21 (2017).
16. K. S. Jin, R. Baillargeon, Infants possess an abstract expectation of ingroup support. *Proc. Natl. Acad. Sci. U.S.A.* 114, 8199–8204 (2017).
17. F. Margoni, R. Baillargeon, L. Surian, Infants distinguish between leaders and bullies. *Proc. Natl. Acad. Sci. U.S.A.* 115, E8835–E8843 (2018).
18. M. Meristo, K. Strid, L. Surian, Preverbal infants' ability to encode the outcome of distributive actions. *Infancy* 21, 353–372 (2016).
19. L. J. Powell, E. S. Spelke, Preverbal infants expect members of social groups to act alike. *Proc. Natl. Acad. Sci. U.S.A.* 110, E3965–E3972 (2013).
20. A. Pun, S. A. J. Birch, A. S. Baron, The power of allies: Infants' expectations of social obligations during intergroup conflict. *Cognition* 211, 104630 (2021).
21. M. Rhodes, C. Hetherington, K. Brink, H. M. Wellman, Infants' use of social partnerships to predict behavior. *Dev. Sci.* 18, 909–916 (2015).
22. M. F. H. Schmidt, J. A. Sommerville, Fairness expectations and altruistic sharing in 15-month-old human infants. *PLoS One* 6, e23223 (2011).
23. S. Sloane, R. Baillargeon, D. Premack, Do infants have a sense of fairness? *Psychol. Sci.* 23, 196–204 (2012).
24. M. Stavans, R. Baillargeon, Infants expect leaders to right wrongs. *Proc. Natl. Acad. Sci. U.S.A.* 116, 16292–16301 (2019).
25. F. Ting, Z. He, R. Baillargeon, Toddlers and infants expect individuals to refrain from helping an ingroup victim's aggressor. *Proc. Natl. Acad. Sci. U.S.A.* 116, 6025–6034 (2019).
26. Y. Wang, A. M. E. Henderson, Just rewards: 17-month-old infants expect agents to take resources according to the principles of distributive justice. *J. Exp. Child Psychol.* 172, 25–40 (2018).
27. J. Decety, N. Steinbeis, J. M. Cowell, The neurodevelopment of social preferences in early childhood. *Curr. Opin. Neurobiol.* 68, 23–28 (2021).

28. F. Ting, M. Buyukozer Dawkins, M. Stavans, R. Baillargeon, "Principles and concepts in early moral cognition" in *The Social Brain*, J. Decety, Ed. (MIT Press, Cambridge, MA, 2020), pp. 41–65.
29. M. P. Burns, J. A. Sommerville, "I pick you": The impact of fairness and race on infants' selection of social partners. *Front. Psychol.* **5**, 93 (2014).
30. A. Geraci, L. Surian, The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Dev. Sci.* **14**, 1012–1020 (2011).
31. K. Lucca, J. Pospisil, J. A. Sommerville, Fairness informs social decision making in infancy. *PLoS One* **13**, e0192848 (2018).
32. T. D. DesChamps, A. E. Eason, J. A. Sommerville, Infants associate praise and admonishment with fair and unfair individuals. *Infancy* **21**, 478–504 (2016).
33. M. Meristo, L. Surian, Do infants detect indirect reciprocity? *Cognition* **129**, 102–113 (2013).
34. L. Surian, L. Franchin, Toddlers selectively help fair agents. *Front. Psychol.* **8**, 944 (2017).
35. J. J. Boseovski, K. Lee, Children's use of frequency information for trait categorization and behavioral prediction. *Dev. Psychol.* **42**, 500–513 (2006).
36. K. M. Cain, G. D. Heyman, M. E. Walker, Preschoolers' ability to make dispositional predictions within and across domains. *Soc. Dev.* **6**, 53–75 (1997).
37. E. E. Chen, K. H. Corriveau, P. L. Harris, Person perception in young children across two cultures. *J. Cogn. Dev.* **17**, 447–467 (2016).
38. K. A. Heller, T. J. Berndt, Developmental changes in the formation and organization of personality attributions. *Child Dev.* **52**, 683–691 (1981).
39. L. Heiphetz, N. Strohminger, S. A. Gelman, L. L. Young, Who am I? The role of moral beliefs in children's and adults' understanding of identity. *J. Exp. Soc. Psychol.* **78**, 210–219 (2018).
40. E. L. Uhlmann, D. A. Pizarro, D. Diermeier, A person-centered approach to moral judgment. *Perspect. Psychol. Sci.* **10**, 72–81 (2015).
41. J. De Freitas et al., Consistent belief in a good true self in misanthropes and three interdependent cultures. *Cogn. Sci. (Hauppauge)* **42** (suppl. 1), 134–160 (2018).
42. L. Heiphetz, Moral essentialism and generosity among children and adults. *J. Exp. Psychol. Gen.* **148**, 2077–2090 (2019).
43. G. E. Newman, P. Bloom, J. Knobe, Value judgments and the true self. *Pers. Soc. Psychol. Bull.* **40**, 203–216 (2014).
44. J. K. Hamlin, The case for social evaluation in preverbal infants: Gazing toward one's goal drives infants' preferences for Helpers over Hinderers in the hill paradigm. *Front. Psychol.* **5**, 1563 (2015).
45. Y. Kanakogi et al., Preverbal infants affirm third-party interventions that protect victims from aggressors. *Nat. Hum. Behav.* **1**, 0037 (2017).
46. Y. E. Lee, J. E. Yun, E. Y. Kim, H. J. Song, The development of infants' sensitivity to behavioral intentions when inferring others' social preferences. *PLoS One* **10**, e0135588 (2015).
47. D. Premack, A. J. Premack, Infants attribute value +/- to the goal-directed actions of self-propelled objects. *J. Cogn. Neurosci.* **9**, 848–856 (1997).
48. L. Surian, M. Ueno, S. Itakura, M. Meristo, Do infants attribute moral traits? Fourteen-month-olds' expectations of fairness are affected by agents' antisocial actions. *Front. Psychol.* **9**, 1649 (2018).
49. J. K. Hamlin, K. Wynn, P. Bloom, Social evaluation by preverbal infants. *Nature* **450**, 557–559 (2007).
50. J. K. Hamlin, K. Wynn, P. Bloom, Three-month-olds show a negativity bias in their social evaluations. *Dev. Sci.* **13**, 923–929 (2010).
51. F. Margoni, L. Surian, Infants' evaluation of prosocial and antisocial agents: A meta-analysis. *Dev. Psychol.* **54**, 1445–1455 (2018).
52. C. Fawcett, U. Liszkowski, Infants anticipate others' social preferences. *Infant Child Dev.* **21**, 239–249 (2012).
53. J. K. Hamlin, K. Wynn, P. Bloom, N. Mahajan, How infants and toddlers react to antisocial others. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 19931–19936 (2011).
54. L. J. Walker, K. H. Hennig, Differing conceptions of moral exemplarity: Just, brave, and caring. *J. Pers. Soc. Psychol.* **86**, 629–647 (2004).
55. L. J. Walker, R. C. Pitts, Naturalistic conceptions of moral maturity. *Dev. Psychol.* **34**, 403–419 (1998).
56. Y. J. Choi, Y. Luo, 13-month-olds' understanding of social interactions. *Psychol. Sci.* **26**, 274–283 (2015).
57. M. E. McCullough, R. Kurzban, B. A. Tabak, Cognitive systems for revenge and forgiveness. *Behav. Brain Sci.* **36**, 1–15 (2013).
58. R. Baillargeon, R. M. Scott, L. Bian, Psychological reasoning in infancy. *Annu. Rev. Psychol.* **67**, 159–186 (2016).
59. G. Gergely, G. Csibra, Teleological reasoning in infancy: The naive theory of rational action. *Trends Cogn. Sci.* **7**, 287–292 (2003).
60. F. Faul, E. Erdfelder, A. G. Lang, A. Buchner, G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191 (2007).