



ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Cognitive Development

journal homepage: www.elsevier.com/locate/cogdev

Young children judge defection less negatively when there's a good justification

Leon Li^{*}, Aren Tucker, Michael Tomasello

Duke University, United States

ARTICLE INFO

Keywords:

Reasoning
Promise
Promise-breaking
Commitment
Moral development

ABSTRACT

Morality includes a common ground ranking of values, of which a central theme is that prosocial actions are more justifiable than selfish ones. Learning to distinguish between good versus bad justifications for actions based on a common ground ranking of values is a key aspect of moral development. The current study assessed this type of understanding in young children. In a within-participants design, young children ($N = 64$) saw puppets who promised to show them a cool toy, failed to fulfill their promise, and then gave either a good (prosocial), bad (selfish), or no justification for their defection. Children's judgments about defection following good justifications were less negative than their judgments about defection following bad or no justifications, which did not differ. When asked to justify their judgments, 5-year-olds (but not 3-year-olds) made more normative or promise-referencing statements when reasoning about puppets who gave good justifications as opposed to bad or no justifications. Children's rates of tattling on, liking, and inviting puppets to play did not vary by the type of justification that a puppet gave. Overall, the findings suggest that the capacity to reference a common ground ranking of values, a key component of human cooperation and morality, is present in young children.

Moral reasoning is a cooperative social activity in which individuals reason with one another with reference to common ground values, which serve as the grounds for justification (Li & Tomasello, 2021, 2022; Tomasello, 2019). Moral development involves not only learning the common ground values themselves but also learning the common ground ranking for how different values are prioritized in cases when competing values conflict. One central theme in the common ground ranking of values is that prosocial actions are more justifiable than selfish ones (Gert, 2004). That is, prosocial reasons for action are more acceptable grounds for justifying an action than selfish reasons are. This theme cuts across many kinds of transgressions. For instance, although it is generally wrong to lie or steal, it is at least more justifiable to violate these norms for prosocial reasons (e.g., lying to protect someone from persecution, stealing medicine to save someone's life) than for selfish reasons.

An open question is when children come to understand that prosocial reasons for action are better grounds for justification, according to the common ground ranking of values, than selfish reasons are. Potentially, this understanding may be present by 3 years of age, as 3-year-olds already display an awareness that justifications are rooted in common ground (Köymen & Tomasello, 2020). For instance, 3-year-olds provide more explicit justifications for their ideas when their common ground with their interlocutor is weak as opposed to strong (Köymen, Mammen, & Tomasello, 2016; Köymen, Rosenbaum, & Tomasello, 2014). However, recent research might also motivate the prediction that a full grasp of the common ground ranking of values is not present until age 5. In one study by Kanngiesser, Mammen, and Tomasello (2021), children chose between two kinds of story-endings for stories about characters who had

^{*} Correspondence to: Duke University, 110 Reuben-Cooke Building, Durham, NC 27708, United States.
E-mail address: leon.inbox@gmail.com (L. Li).

<https://doi.org/10.1016/j.cogdev.2022.101268>

Received 11 April 2022; Received in revised form 21 September 2022; Accepted 13 October 2022

Available online 3 November 2022

0885-2014/© 2022 Elsevier Inc. All rights reserved.

made promises: story-endings in which characters fulfilled their promises versus story-endings in which characters broke their promises for either selfish or prosocial reasons. Whereas 5-year-olds preferred the prosocial promise-breaking story-endings more than the selfish promise-breaking story-endings, 3-year-olds actually showed the opposite preference.

In another study by Kanngiesser et al. (2021), children who were 3 or 5 years old promised to help with a clean-up task but were enticed to defect from the task for either a selfish reason (to play with a toy) or a prosocial reason (to help someone fix a toy). An adult then asked children why they were not cleaning. The 3-year-olds provided more relevant justifications in the prosocial condition than in the selfish condition, suggesting a nascent awareness that acting prosocially affords better grounds for justifying one's actions compared to acting selfishly. Yet, only 5-year-olds' justifications referenced social norms more in the prosocial condition than in the selfish condition, whereas 3-year-olds' justifications did not show the same sensitivity, suggesting that 5-year-olds have a more explicit understanding of common ground values and how to appropriately reference them. The interpretation that children's justifications to adults reflect their understanding of common ground values follows from the premise that children view adults as authoritative representatives of their culture; in order to persuade adults of the validity of one's justifications, one must appeal to the common ground values that adults, as representatives of the culture, also view as legitimate.

One limitation of the studies by Kanngiesser et al. (2021) is that they did not directly assess children's judgments about the relative wrongness of prosocial versus selfish instances of promise-breaking. Such judgments would provide a relatively clear indication of children's normative ranking of prosocial versus selfish reasons for promise-breaking. Another limitation is that their research only compared children's perceptions of good (prosocial) versus bad (selfish) reasons for promise-breaking, neglecting a third possibility in which someone breaks a promise but gives no reason for doing so. Thus, it remains unclear whether children would perceive promise-breaking for a bad reason and promise-breaking for no reason to be equally wrong or would instead favor any reason for promise-breaking (even a bad one) over no reason.

Relevant to this, research has shown that children can be unduly influenced by the merely structural elements of reason-giving discourse. One study found that both 3-year-olds and 5-year-olds favored individuals who gave idiosyncratic arguments for receiving favorable treatment as much as they favored individuals who gave legitimate arguments (Schmidt, Svetlova, Johe, & Tomasello, 2016). Another study found that 5-year-olds favored statements backed by circular explanations (even though such explanations were not actually informative) over statements backed by no explanation (Mercier, Bernard, & Clément, 2014). These studies suggest that merely structural elements of reason-giving discourse (e.g., the fact that any reason was given at all) may potentially sway young children as much as the actual contents of reasons.

Building on the prior body of research, the current study compared children's judgments about promise-breaking following good, bad, or no justifications. Young children (3-year-olds and 5-year-olds) encountered puppets who promised to show them a cool toy, left the scene, and then came back without a cool toy to show. The puppets then gave either good (prosocial), bad (selfish), or no justifications for their defection. A within-participants design was used—all children observed and responded to all three kinds of justification for defection. Children gave judgments and justifications about whether it was wrong that the various puppets did not show them a cool toy, as they had promised. The main hypothesis was that children's judgments about promise-breaking would be less negative following prosocial as opposed to selfish or no justifications.

A second prediction was that children's justifications for their judgments would contain more explicitly normative statements (e.g., what one *should*, *has to*, or *is supposed to* do) or references to the fact that a promise was made at 5 years of age than at 3 years of age. This prediction was based on previous evidence suggesting that 5-year-olds have a better grasp of the normative implications of promises compared to 3-year-olds (Kanngiesser et al., 2021), as well as the fact that children's reason-giving skills improve in general from 3 to 5 years of age (Köymen, Jurkat, & Tomasello, 2020; Köymen & Tomasello, 2020; Mammen, Köymen, & Tomasello, 2018). If an understanding that promises create normative obligations is more fully developed in 5-year-olds than in 3-year-olds, then 5-year-olds should be more likely to refer to promises or make explicitly normative statements when reasoning about acts of promise-breaking compared to 3-year-olds.

In addition to the judgments and justifications, three other measures were also assessed: children's rates of tattling on the various puppets, preferentially liking certain puppets over others, and preferentially inviting certain puppets over others to a play date. These measures assessed the social consequences that different kinds of justifications for promise-breaking may have on children's perceptions of or behaviors towards agents. Relevant to this, research has shown that children's trust in agents is impacted by whether agents provide good or bad justifications (Castelain, Bernard, Van der Henst, & Mercier, 2016; Rakoczy, Miosga, & Schultze, 2022) and also by whether agents keep their promises or not (Isella, Kanngiesser, & Tomasello, 2019). Accordingly, the prediction here was that children's responses on these measures of the social consequences of promise-breaking would favor puppets who gave good justifications over those who gave bad or no justifications.

1. Methods

1.1. Participants

The sample included two age groups: 3.5-year-olds ($n = 32$, $M = 42$ months, $SD = 1.66$, range: 39–45) and 5.5-year-olds ($n = 32$, $M = 66$ months, $SD = 1.98$, range: 63–69). Both age groups had 18 girls and 14 boys. The sample size of 64 was based on previous studies examining related topics, which used similar sample sizes (e.g., Engelmann, Herrmann, & Tomasello, 2016; Kanngiesser et al., 2021). Otherwise, since this study used a novel procedure, there was no other basis for estimating an appropriate sample size or expected effect size. Participants were recruited via emails to parents in our university's database of local birth records from the Southeastern United States. Over 75% of children were ethnically white, over 85% of parents held a bachelor's degree or higher, and over 85% of

annual family incomes were \$75,000 or higher. Five additional children participated but were excluded from the final sample due to excessive distraction or unwillingness to participate. This study was approved by the Institutional Review Board of Duke University on June 17, 2021.

1.2. Procedure

The study took place over Zoom. To begin, the experimenter told parents to not influence their child's responses, obtained verbal consent, and helped parents set up Zoom in "Side-by-side" mode so that the experimenter and the experimenter's screen-shared stimuli (e.g., PowerPoint slides) appeared at similar sizes to the children. Children's self-view was hidden to help maintain focus on the experimenter and the stimuli.¹ Then, the experimenter introduced the "Head Puppet" (HP), a squirrel puppet operated by the experimenter. The experimenter then went off-screen. For the rest of the study, HP interacted with the child in place of the experimenter.

In a warm-up phase, HP showed the child three pairs of popular story characters (e.g., Elsa and Olaf from Frozen). For each pair, HP asked the child, "Who would you rather invite to play with you, (Character 1) or (Character 2)?" This was meant to familiarize children with the process of choosing between two puppets. Next, HP suggested, "We should have a playdate together! We can invite some more puppets!" HP showed a slide with the 12 animal puppets that would appear later in the study and introduced each one. After this, HP stated, "Remember, I'm the Head Puppet, and I'm in charge of all the other puppets. I'll make sure everyone is on their best behavior. If you see anyone doing anything wrong, you should tell me. OK?" HP then explained, "All of these puppets have cool toys that they want to show you. They'll take turns showing you their cool toys. After they show you their cool toys, you can decide who you want to invite."

Children responded to six pre-recorded videos. All puppets were voiced by male experimenters. The first video trial was a warm-up trial designed to help children feel comfortable with reporting transgressions to HP. In this video, two puppets (Pig and Cow) both bring sandwich materials, but one puppet (Cow) makes a sandwich incorrectly by placing a box of crayons between two slices of bread. Once the video ended, HP returned and asked the child what happened. If the child did not tattle, HP encouraged the child by first asking, "Did Cow make their sandwich correctly?" and then, if the child still did not tattle, by saying, "It looked like Cow made their sandwich wrong. Don't you think so?" Next, HP reminded the child, "It's really important that you tell me if anyone does anything wrong. OK?" HP then asked the child, "Who would you rather invite to play with you? Pig or Cow?" before proceeding to the next five video trials.

Each of the next five videos featured a different pair of animal puppets (e.g., Lion and Tiger, Cat and Dog, Zebra and Giraffe, Parrot and Bunny, Elephant and Bear). Each trial began with HP introducing the current pair of puppets and then going off-camera. Each puppet then introduced itself and promised to show the child a cool toy. Then, the two puppets exited at the same time and remained off-screen for 15 s before returning at the same time. Children experienced two *motivation trials* in which the puppets showed the child interesting toys when they returned. Children also experienced three *experimental trials* in which the puppets did not show the child anything when they returned but instead gave either a good excuse, a bad excuse, or no excuse to justify their defection (scripts in [Appendix A](#)).

Good excuses referenced an obligation to help others (e.g., "I had to help my friend with his homework instead"). Bad excuses referenced a selfish desire to do something else (e.g., "I wanted to watch TV instead"). Each child experienced one of each of three kinds of experimental trials: (i) one "Good vs. Bad Excuse" trial, in which one puppet gave a good excuse for not bringing anything, whereas the other puppet gave a bad excuse, (ii) one "Good vs. No Excuse" trial, in which one puppet gave a good excuse, whereas the other puppet gave no excuse, and (iii) one "Bad vs. No Excuse" trial, in which one puppet gave a bad excuse, whereas the other puppet gave no excuse.

The motivation trials were always on the second and fourth trials. The experimental trials were always on the third, fifth, and sixth trials. The order of presentation of the three kinds of experimental trials (among the third, fifth, and sixth "slots") was fully counterbalanced across six lists. Moreover, six additional lists were created in which the order of presentation of the two kinds of excuses *within* each video was flipped relative to the original six lists (e.g., now, in the "Good vs. Bad Excuse" trials, the first puppet gave a bad excuse instead of a good excuse, and the second puppet gave a good excuse instead of a bad excuse). Thus, there were 12 lists in total, and each participant experienced the order of presentation described by one of the 12 lists.

At the end of each video, once the puppets had left, HP returned and asked, "Is there anything that you want to tell me?" For the motivation trials, children typically did not have anything negative to report, and HP simply followed up by asking what toys the puppets showed. For the experimental trials, however, children's responses were coded as to whether they tattled on one or both of the puppets (the tattle measure). In the experimental trials only, HP followed up by asking, for each of the two puppets, "Was it wrong that (Puppet) didn't show you anything?" (the judgment measure) and "Why was it wrong/okay?" (the justification measure). In both the motivation and experimental trials, HP then asked children, "Who do you like better, (Puppet 1) or (Puppet 2)?" (the liking measure) and "Who would you rather invite to our playdate? (Puppet 1) or (Puppet 2)?" (the invitation measure). However, only data from the experimental trials was analyzed. After the last video trial, HP thanked the child for participating.

Children's justifications for their judgments were coded into four ranked categories (shown in their order of priority when coding responses that fit more than one category): (i) *promise-referencing statements*: references to the fact that the puppet promised or said he

¹ The procedural decision to hide children's self-view was made after a few children in each age group had already participated.

would do something (e.g., “he promised he would bring it, but he didn’t”), (ii) *normative statements*: references to what the puppet should, had to, needed to, or was supposed to do (e.g., “he should help his friend with his homework”), (iii) *descriptive statements*: describing something about oneself (e.g., “I really wanted to see it”) or what the puppet did or did not have, say, or do (e.g., “he didn’t bring his toy”), and (iv) *uninformative statements*: unclear or irrelevant statements (e.g., “I forgot what Parrot said”). For interrater reliability, a second coder coded 25% ($n = 16$) of the justifications; Cohen’s κ was 0.92.

For data analysis, the four types of justifications were scored numerically based on an ordinal ranking. The ranking scheme reflected the degree to which different types of justifications demonstrated an understanding of promises and their normative implications. Uninformative statements received the lowest value (0) because they do not demonstrate any understanding of promises. Descriptive statements received a higher value (1) because they at least demonstrate a descriptive or implicit awareness of what is occurring in the situation. Normative statements and promise-referencing statements both received the highest value (2) because they both demonstrate one’s ability to justify judgments by appealing to common ground expectations about what promises normatively entail.

2. Results

The data analyses were conducted in R version 4.0.0. The data files and R codes used in the analyses are accessible in the [supplementary materials](#). Note that the analyses of children’s judgments and justifications excluded six instances in which children responded “I don’t know” on the judgment measure.

2.1. Judgments

Children’s judgments were analyzed with generalized linear mixed models (GLMMs) assuming a binomial distribution of responses. Judgments of “Wrong” were coded as 0, and judgments of “Okay” were coded as 1. For interrater reliability, a second coder coded 25% of the judgments ($n = 16$); Cohen’s κ was 0.94. A series of models were created and compared. Model 1, the null model, only included the fixed effect of Trial (Good vs. Bad, Good vs. No, Bad vs. No) and random intercepts for Participant. The Trial level “Good vs. Bad” was the reference level. Model 2, the main effects model, added the fixed effects of Age (3.5, 5.5) and Excuse (Good, Bad, No). The Age level “3.5” and the Excuse level “Good” were the reference levels. Model 3, the full model, added the interaction of Age and Excuse. Model comparisons using likelihood ratio tests showed that Model 2 had a significantly improved fit compared to Model 1, $\chi^2(3) = 14.22$, $p = 0.003$, but Model 3 did not have an improved fit compared to Model 2, $\chi^2(2) = 1.60$, $p = 0.45$. Thus, Model 2 was preferred as the most parsimonious explanation of the data (Table 1).

Likelihood ratio tests based on Model 2 showed that the main effects of Trial ($p = 0.78$) and Age ($p = 0.10$) were not significant, but the main effect of Excuse was significant ($p = 0.003$). To examine how children’s judgments varied by the three levels of Excuse, pairwise comparisons using the Tukey adjustment of p -values were conducted. The comparisons showed that children’s judgments about Good Excuse puppets ($M = 0.43$) were less negative than their judgments about Bad Excuse ($M = 0.31$, $p = 0.03$) and No Excuse puppets ($M = 0.30$, $p = 0.007$), which did not differ, $p = 0.88$ (Fig. 1). Note that children’s judgments about all three types of puppets were mostly negative (i.e., below the midpoint of 0.5), likely because all three types of puppets failed to fulfill their promise to show the child a cool toy.

2.2. Justifications

Children’s justifications were analyzed with multilevel logistic regressions using the *ordinal* package. Uninformative statements were scored as 0, descriptive statements were scored as 1, and normative statements and promise-referencing statements were scored as 2. Model 1, the null model, only included the fixed effect of Trial (Good vs. Bad, Good vs. No, Bad vs. No) and random intercepts for Participant. The Trial level “Good vs. Bad” was the reference level. Model 2, the main effects model, added the fixed effects of Age (3.5, 5.5) and Excuse (Good, Bad, No). The Age level “3.5” and the Excuse level “Good” were the reference levels. Model 3, the full model,

Table 1

Output of the generalized linear mixed model of children’s judgment responses. The reference levels for the categorical factors were 3.5 (Age), Good (Excuse), and Good vs. Bad (Trial). GvN: Good vs. No excuse. BvN: Bad vs. No excuse. * $p < 0.05$; ** $p < 0.01$.

Formula: Judgment ~ 1 + Age + Excuse + Trial + (1 Participant)				
Random effects:				
	Name	Variance	SD	
Participant	(Intercept)	11.44	3.382	
Fixed effects:				
	Estimate	SE	z value	Pr(> z)
(Intercept)	-0.1340	0.7768	-0.173	0.86302
Age[5.5]	-1.6662	1.0438	-1.596	0.11043
Excuse[Bad]	-1.2777	0.4976	-2.568	0.01024 *
Excuse[No]	-1.5169	0.5014	-3.026	0.00248 **
Trial[GvN]	0.2114	0.4718	0.448	0.65418
Trial[BvN]	0.3330	0.4842	0.688	0.49161

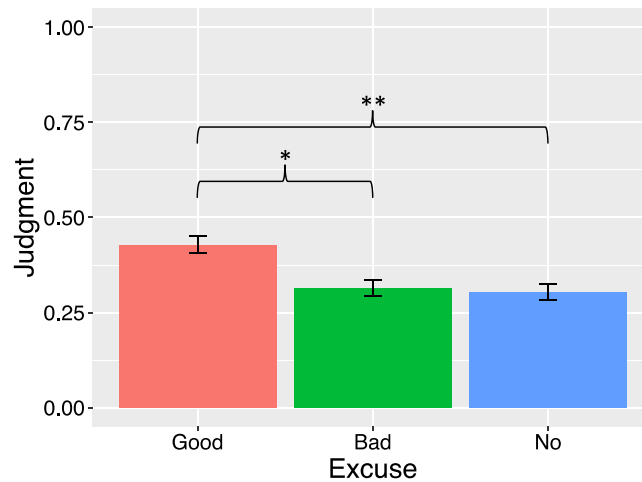


Fig. 1. Children’s judgments about promise-breaking following a good, bad, or no excuse. Judgments of “Wrong” were coded as 0, and judgments of “Okay” were coded as 1. Error bars represent standard errors. * $p < 0.05$; ** $p < 0.01$.

added the interaction of Age and Excuse. Model comparisons using likelihood ratio tests showed that Model 2 had a significantly improved fit compared to Model 1, $\chi^2(3) = 35.76, p < 0.001$, and Model 3 also had a significantly improved fit compared to Model 2, $\chi^2(2) = 15.58, p < 0.001$. Thus, Model 3 was preferred as the best explanation of the data (Table 2).

Likelihood ratio tests based on Model 3 showed that the effect of Trial was not significant, $p = 0.53$, but the interaction between Age and Excuse was significant, $p < 0.001$. To examine the interaction, pairwise comparisons of the three levels of Excuse using the Tukey adjustment of p -values were conducted for each age group separately. The analyses found that 5-year-olds’ justifications about Good Excuse puppets consisted of higher-ranked codes than their justifications about either Bad Excuse puppets ($p < 0.001$) or No Excuse puppets ($p < 0.001$), which did not differ, $p = 0.81$. In contrast, 3-year-olds showed no difference in the rankings of their justifications about Good Excuse, Bad Excuse, and No Excuse puppets, $ps > 0.05$. Fig. 2 shows the counts for children’s justification responses, grouped by Age and Excuse.

To focus specifically on children’s use of normative or promise-referencing statements, the null, main effects, and interaction models for the justification analyses were re-run but with a dichotomized dependent variable. Normative and promise-referencing statements were coded as 1, and uninformative and descriptive statements were coded as 0. Model comparisons again showed that the main effects model had an improved fit compared to the null model, $\chi^2(3) = 33.75, p < 0.001$, and the interaction model had an improved fit compared to the main effects model, $\chi^2(2) = 10.80, p = 0.005$. Likelihood ratio tests based on the interaction model showed that the effect of Trial was not significant, $p = 0.09$, but the interaction between Age and Excuse was significant, $p = 0.005$. To examine the interaction, pairwise comparisons of the three levels of Excuse using the Tukey adjustment of p -values were conducted for each age group separately. The analyses found that 5-year-olds used more normative or promise-referencing statements when reasoning about Good Excuse puppets than when reasoning about Bad Excuse or No Excuse puppets ($ps < 0.001$). In contrast, 3-year-olds’ use of normative or promise-referencing statements did not differ when reasoning about the three types of puppets ($ps > 0.05$).

Table 2

Output of the multilevel logistic regression model of children’s justification responses. The reference levels for the categorical factors were 3.5 (Age), Good (Excuse), and Good vs. Bad (Trial). GvN: Good vs. No excuse. BvN: Bad vs. No excuse. ** $p < 0.01$; *** $p < 0.001$.

Formula: Justification ~ 1 + Age * Excuse + Trial + (1 Participant)					
Random effects:		Name	Variance	SD	
Participant		(Intercept)	4.002	2.000	
Fixed effects:		Estimate	SE	z value	Pr(> z)
Age[5.5]		3.8094	0.6973	5.463	4.68e-08 ***
Excuse[Bad]		-0.1155	0.4418	-0.261	0.793721
Excuse[No]		-0.3814	0.4509	-0.846	0.397657
Trial[GvN]		-0.3693	0.3385	-1.091	0.275308
Trial[BvN]		-0.2502	0.3221	-0.777	0.437423
Age[5.5]:Excuse[Bad]		-2.1145	0.5864	-3.606	0.000311 ***
Age[5.5]:Excuse[No]		-1.8739	0.5983	-3.132	0.001736 **

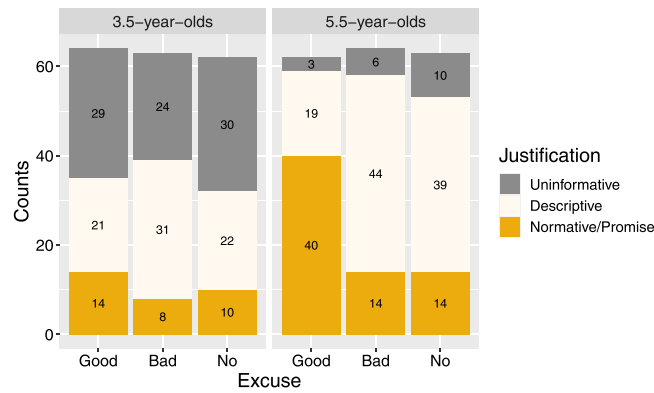


Fig. 2. The counts for children’s justification responses, grouped by Age and Excuse.

2.3. Tatting, liking, and invitation

Children’s tattling, liking, and invitation responses were analyzed using one-sample proportion tests against chance. These tests assessed whether children tattled on, liked, or invited puppets who gave any of the three types of Excuse more than would be expected by chance, defined as one half of the time. For example, to assess whether children invited Good Excuse puppets more than would be expected by chance, the number of times children invited the Good Excuse puppet in the Good vs. Bad Excuse and Good vs. No Excuse trials was divided by the total number of times children invited any puppet in those two trials; that proportion was then compared to the chance level of 0.5. For interrater reliability, a second coder coded 25% of the tattling, liking, and invitation responses (n = 16), assessing whether children chose neither, one, or both of the puppets for each measure. Cohen’s κ values were 0.85 (tattling), 1.00 (liking), and 1.00 (invitation).

No significant effects of Excuse were found for any of the measures. That is, children did not tattle on, like, or invite any of the three types of puppets more than would be expected by chance. When analyzing the two age groups separately, there were still no significant effects of Excuse for any of the three measures. Table 3 shows the counts for children’s tattling, liking, and invitation responses, grouped by Age, Trial, and Excuse.

3. Discussion

Young children encountered puppets who promised to show them a cool toy but then failed to fulfill their promise. Although children’s judgments about promise-breaking were negative in general, their judgments were less negative when the puppets gave prosocial justifications for why they broke their promise (e.g., they had to help someone) as opposed to selfish justifications (e.g., they just wanted to do something fun instead) or no justification. Children were also asked to justify their judgments, and their justification responses revealed an interaction between children’s age (3 or 5 years old) and the type of justification (prosocial, selfish, or none at all) given by the puppet about which they were reasoning. Whereas 5-year-olds used more normative or promise-referencing statements when reasoning about puppets who gave prosocial justifications than when reasoning about puppets who gave selfish or no justifications, 3-year-olds did not differ in the types of statements they used when reasoning about the three types of puppets.

Overall, what the judgment data suggest is that a nascent understanding of the common ground ranking of values underlying moral reasoning is present in young children. Children from 3 to 5 years of age may recognize that prosocial obligations take priority over selfish desires as credible grounds for justifying transgressions, as evidenced by their more lenient judgments of those who broke

Table 3
The counts for children’s tattling, liking, and invitation responses, grouped by Age, Trial, and Excuse.

Age	Trial	Excuse	Tattle	Liking	Invitation
3.5	Good vs. Bad	Good	24	18	20
		Bad	22	15	14
	Good vs. No	Good	23	18	18
		No	24	15	15
	Bad vs. No	Bad	19	17	13
		No	20	12	17
5.5	Good vs. Bad	Good	23	17	13
		Bad	27	14	19
	Good vs. No	Good	25	16	16
		No	26	15	15
	Bad vs. No	Bad	27	15	14
		No	26	16	17

promises for prosocial as opposed to selfish reasons. The current results also suggest that young children are better at evaluating others' justifications than has been implied by previous research. One previous study found that 5-year-olds favored individuals who gave idiosyncratic reasons for receiving favorable treatment as much as they favored individuals who gave legitimate reasons, suggesting that merely structural elements of reason-giving discourse (e.g., the fact that any reason was given at all) swayed them as much as the actual contents of the reasons (Schmidt et al., 2016). Similarly, another study found that 5-year-olds favored statements backed by circular explanations over statements backed by no explanation, even though the circular explanations were not actually informative (Mercier et al., 2014).

In contrast, the current finding that young children's judgments about puppets who gave good justifications were more lenient than their judgments about puppets who gave either bad or no justifications, which did not differ, implies that young children are not always swayed by merely structural elements of reason-giving discourse. In the domain of promise-breaking, at least, young children may weigh the actual contents of justifications above and beyond the fact that any justification was given at all. It is possible that the task demands of the current study were more advantageous than those of previous studies for revealing this type of understanding in young children.

What the justification data suggest is that children increasingly recognize the *normative* status of prosocial obligations as valid reasons for action and grounds for justification from ages 3–5. That is, children from ages 3–5 increasingly view prosocial obligations not just descriptively as things that people do but normatively as things that people *should* do. When the puppets in the prosocial condition explained that they didn't bring a toy because they had to help someone, 5-year-olds recognized that the puppets were not simply giving a *descriptive* account of what they had been doing—unlike in the selfish condition, in which the puppets *were* simply only describing what they had been doing instead—but were indeed articulating a *normative* appeal to a common ground understanding that prosocial obligations may be valid reasons for overriding an existing commitment.

Only 5-year-olds evidenced a recognition of the normative status of prosocial obligations by articulating more normative or promise-referencing statements, not just descriptive statements, when reasoning about puppets who acted prosocially than when reasoning about puppets who acted selfishly. In contrast, the fact that 3-year-olds did not use significantly more normative or promise-referencing statements when reasoning about puppets who acted prosocially than when reasoning about puppets who acted selfishly suggests that 3-year-olds may have viewed both prosocial obligations and selfish desires in a similarly descriptive way. A similar age difference was found by Kanngiesser et al. (2021). In their study, children who were 3 or 5 years old promised to help with a clean-up task but were enticed to defect from the task. The critical manipulation was whether children were enticed to defect for a selfish reason (to play with a toy) or a prosocial reason (to help someone fix a toy). When asked to explain why they had stopped cleaning, only 5-year-olds (not 3-year-olds) referenced social norms more in the prosocial condition than in the selfish condition. Altogether, the results of Kanngiesser et al. (2021) and those of the current study both suggest an age-related improvement from ages 3–5 in children's recognition that prosocial obligations constitute normative reasons for action and grounds for justification.

The conclusions of the current study are importantly qualified by the fact that several null findings were observed. Children's rates of tattling on, preferentially liking, and preferentially inviting different puppets to a play date did not vary by justification type (not even when analyzing each age group separately). The similar rates of tattling on the various puppets may be unsurprising, given that all the puppets failed to fulfill their promise to show the child a cool toy. Yet, children might have been plausibly expected to like or invite puppets who gave prosocial justifications more than puppets who gave selfish or no justifications, especially since they could already distinguish between the puppets in their judgments of wrongness. It is possible that children up to 5 years of age still have an incomplete understanding of the social implications (e.g., attributions of personality or character) of the ways in which agents justify their transgressions.

That is, although young children may recognize that it is less wrong to break a promise for a prosocial as opposed to a selfish reason, they may not yet make the additional inference that someone who breaks a promise for a prosocial reason may be more pleasant or likable than someone who breaks a promise for a selfish reason. Relevant to this, previous studies have shown mixed results in how effective young children are at making trait attributions based on others' behaviors, with some studies suggesting that young children struggle with such inferences (Boseovski & Lee, 2006; Guajardo, Nixon, Royster, & Roccato, 2022; Kalish, 2002; Liu, Gelman, & Wellman, 2007; Rholes & Ruble, 1984). It is possible that making contrasting inferences about different puppets' character traits was especially hard in the current study, in which different puppets all performed the same negatively valenced and salient behavior (defecting from a promise), only differing in their verbal justifications for the behavior.

Another potential challenge to the current conclusions concerns whether children's judgments really reflected an understanding of a *common ground* ranking of values. An alternative interpretation is that children themselves may have viewed promise-breaking for prosocial reasons as less wrong than promise-breaking for selfish reasons, but they did not make the further inference that this ranking was widely endorsed by others in common ground. This view might be mitigated by the body of evidence that young children are aware of the common ground they share with others across many domains, including word labels, cultural artifacts, and generic knowledge (Köymen & Tomasello, 2020; Li, Britvan, & Tomasello, 2021; Tomasello, 2014, 2016). It would not seem unreasonable to suppose that young children, who are aware of sharing common ground with others about a range of domains, would also conceptualize moral values as a type of shared common ground. Another point in favor of the current conclusions is that children in the current study gave their judgments and justifications at the behest of a social interaction partner, a puppet operated by an adult. In order to persuade another person of the validity of one's judgments and justifications, one must appeal to the common ground values that the other person also deems legitimate. Children's responses in this interaction context, as such, plausibly reflected their consideration of what the common ground values were.

3.1. Limitations

One limitation of the current study was the absence of 4-year-olds, whose inclusion would have enabled a better sense of the developmental transition from 3–5 years of age. A second limitation was the fact that puppet interlocutors were used rather than human partners. Although puppets are useful for implementing a tightly controlled experimental design, they may lessen a study's ecological validity (Packer & Moreno-Dulcey, 2022). Some recent articles have argued that puppets are nonetheless valid and valuable tools for probing children's social cognitive abilities (Lillard, 2022; Rakoczy, 2022). It would be interesting for future research to examine whether young children's priority of prosocial over selfish justifications replicates in more naturalistic contexts with human partners.

Relatedly, another limitation of the generalizability of the findings is that only North American children from relatively affluent backgrounds were recruited. Whether children from other societies would prioritize prosocial over selfish justifications is one potential line of future inquiry (we expect that they would, as acting prosocially is universally valued across cultures). Another interesting line of inquiry would be to examine what specifically constitutes good versus bad grounds for justification in different cultures, as different societies may vary in their norms, values, and expectations about what it means to act prosocially. For instance, one cross-cultural study found that American and Chinese children had different views about whether one should tattle on a friend who cheated in a contest (Zhang et al., 2013). American children typically interpreted the concept of friendship as meaning that one should keep secrets for one's friend and not tattle. In contrast, Chinese children were likelier than American children to interpret the concept of friendship as meaning that one should indeed tattle to help the friend become a better person. This interesting example illustrates how the same concept of acting prosocially towards one's friend may be interpreted differently in different cultures.

A fourth limitation was that the types of justifications that the puppets invoked were very constrained. The good justifications referred to simple prosocial obligations (to help someone), and the bad justifications referred to simple desires (to do something fun). In real life, however, the ranking of values that individuals reference when justifying their actions to one another is much more complex. Beyond straightforwardly "good" and "bad" reasons for action, individuals must navigate a heterogeneous and at times tangled hierarchy of values, norms, and obligations that are often in conflict with each other. It would be fascinating for future research to examine how children understand and navigate competing concerns when there is less clarity as to what the "one best solution" is—a challenge even for adults.

3.2. Conclusion

In closing, the current study suggests that children from 3–5 years of age are able to distinguish good versus bad justifications for defection, indicating an early sensitivity to the common ground ranking of values that underlies moral reasoning. The capacity to co-construct and socially reference a common ground ranking of values may have contributed importantly to the success of human cooperation, as such values helpfully inform how individuals regulate their own actions as well as how they coordinate with others' actions socially. Further research on how children come to appreciate and act on the common ground values of their cultures could help shed light on how human cooperation and morality have scaled up so remarkably across evolution and ontogeny.

Appendix A. Scripts of the pre-recorded puppets in the video trials

Head Puppet: "This is (<i>Puppet 1</i>), and this is (<i>Puppet 2</i>). They have really, really cool toys that they want to show you. Do you want to see their cool toys?"			
Head Puppet: "Great! Here they are!"			
<i>*Start video.</i>			
Puppet 1: "Hi I'm (<i>Puppet 1</i>)! I have a really cool toy that I want to show you. I'll go get it and show it to you! I promise."			
Puppet 2: "Hi I'm (<i>Puppet 2</i>)! I also have a really cool toy that I want to show you. I'll go get it and show it to you! I promise."			
<i>*The puppets leave for 15 s and then return with/without the cool toys.</i>			
Good vs. Bad Excuse Trial	Good vs. No Excuse Trial	Bad vs. No Excuse Trial	Motivation Trial
Puppet 1: "Hey, I'm back! I didn't bring my cool toy. The reason I didn't bring my cool toy is because I had to help my mom instead. My mom really needs my help."	Puppet 1: "Hey, I'm back! I didn't bring my cool toy. The reason I didn't bring my cool toy is because I had to help my friend with his homework instead. My friend really needs my help."	Puppet 1: "Hey, I'm back! I didn't bring my cool toy. The reason I didn't bring my cool toy is because I wanted to play a game by myself instead. My game is so much fun!"	Puppet 1: "Hey, I'm back! Here is my [cool toy]!" <i>*Shows the child the cool toy.</i>
"I have to get back to helping my mom. Bye!" <i>*leaves</i>	"I have to get back to helping my friend. Bye!" <i>*leaves</i>	"I want to get back to playing my game. Bye!" <i>*leaves</i>	"Thanks for letting me show you my [cool toy]. Bye!" <i>*leaves</i>
Puppet 2: "I didn't bring my cool toy. The reason I didn't bring my cool toy is because I wanted to watch TV instead. My TV is awesome!"	Puppet 2: "I didn't bring my cool toy. Bye!" <i>*leaves</i>	Puppet 2: "I didn't bring my cool toy. Bye!" <i>*leaves</i>	Puppet 2: "Here is my [cool toy]!" <i>*Shows the child the cool toy.</i>
"I want to get back to watching TV. Bye!" <i>*leaves</i>			"Thanks for letting me show you my [cool toy]. Bye!" <i>*leaves</i>

Appendix B. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.cogdev.2022.101268](https://doi.org/10.1016/j.cogdev.2022.101268).

References

- Boseovski, J. J., & Lee, K. (2006). Children's use of frequency information for trait categorization and behavioral prediction. *Developmental Psychology*, 42(3), 500–513. <https://doi.org/10.1037/0012-1649.42.3.500>
- Castelain, T., Bernard, S., Van der Henst, J.-B., & Mercier, H. (2016). The influence of power and reason on young Maya children's endorsement of testimony. *Developmental Science*, 19(6), 957–966. <https://doi.org/10.1111/desc.12336>
- Engelmann, J. M., Herrmann, E., & Tomasello, M. (2016). Preschoolers affect others' reputations through prosocial gossip. *British Journal of Developmental Psychology*, 34(3), 447–460. <https://doi.org/10.1111/bjdp.12143>
- Gert, B. (2004). *Common morality: Deciding what to do*. Oxford University Press.
- Guajardo, N. R., Nixon, B., Royster, L., & Rocco, R. (2022). The roles of behavioral and affective cues and false belief in children's trait attributions. *Journal of Experimental Child Psychology*, 222, Article 105475. <https://doi.org/10.1016/j.jecp.2022.105475>
- Isella, M., Kanngiesser, P., & Tomasello, M. (2019). Children's selective trust in promises. *Child Development*, 90(6), e868–e887. <https://doi.org/10.1111/cdev.13105>
- Kalish, C. W. (2002). Children's predictions of consistency in people's actions. *Cognition*, 84(3), 237–265. [https://doi.org/10.1016/S0010-0277\(02\)00052-5](https://doi.org/10.1016/S0010-0277(02)00052-5)
- Kanngiesser, P., Mammen, M., & Tomasello, M. (2021). Young children's understanding of justifications for breaking a promise. *Cognitive Development*, 60, Article 101127. <https://doi.org/10.1016/j.cogdev.2021.101127>
- Köymen, B., Jurkat, S., & Tomasello, M. (2020). Preschoolers refer to direct and indirect evidence in their collaborative reasoning. *Journal of Experimental Child Psychology*, 193, Article 104806. <https://doi.org/10.1016/j.jecp.2020.104806>
- Köymen, B., Mammen, M., & Tomasello, M. (2016). Preschoolers use common ground in their justificatory reasoning with peers. *Developmental Psychology*, 52(3), 423–429. <https://doi.org/10.1037/dev0000089>
- Köymen, B., Rosenbaum, L., & Tomasello, M. (2014). Reasoning during joint decision-making by preschool peers. *Cognitive Development*, 32, 74–85. <https://doi.org/10.1016/j.cogdev.2014.09.001>
- Köymen, B., & Tomasello, M. (2020). The early ontogeny of reason giving. *Child Development Perspectives*, 14(4), 215–220. <https://doi.org/10.1111/cdep.12384>
- Li, L., Britvan, B., & Tomasello, M. (2021). Young children conform more to norms than to preferences. *PLOS ONE*, 16(5), Article e0251228. <https://doi.org/10.1371/journal.pone.0251228>
- Li, L., & Tomasello, M. (2021). On the moral functions of language. *Social Cognition*, 39(1), 99–116. <https://doi.org/10.1521/soco.2021.39.1.99>
- Li, L., & Tomasello, M. (2022). Disagreement, justification, and equitable moral judgments: A brief training study. *Journal of Experimental Child Psychology*, 223, Article 105494. <https://doi.org/10.1016/j.jecp.2022.105494>
- Lillard, A. S. (2022). Pretending at hand: How children perceive and process puppets. *Cognitive Development*, 63, Article 101202. <https://doi.org/10.1016/j.cogdev.2022.101202>
- Liu, D., Gelman, S. A., & Wellman, H. M. (2007). Components of young children's trait understanding: Behavior-to-trait inferences and trait-to-behavior predictions. *Child Development*, 78(5), 1543–1558. <https://doi.org/10.1111/j.1467-8624.2007.01082.x>
- Mammen, M., Köymen, B., & Tomasello, M. (2018). The reasons young children give to peers when explaining their judgments of moral and conventional rules. *Developmental Psychology*, 54(2), 254–262. <https://doi.org/10.1037/dev0000424>
- Mercier, H., Bernard, S., & Clément, F. (2014). Early sensitivity to arguments: How preschoolers weight circular arguments. *Journal of Experimental Child Psychology*, 125, 102–109. <https://doi.org/10.1016/j.jecp.2013.11.011>
- Packer, M. J., & Moreno-Dulcey, F. A. (2022). Theory of puppets?: A critique of the use of puppets as stimulus materials in psychological research with young children. *Cognitive Development*, 61, Article 101146. <https://doi.org/10.1016/j.cogdev.2021.101146>
- Rakoczy, H. (2022). Puppet studies present clear and distinct windows into the child's mind. *Cognitive Development*, 61, Article 101147. <https://doi.org/10.1016/j.cogdev.2021.101147>
- Rakoczy, H., Miosga, N., & Schultze, T. (2022). Young children evaluate and follow others' arguments when forming and revising beliefs. *Social Development*, 31(1), 147–164. <https://doi.org/10.1111/sode.12533>
- Rholes, W. S., & Ruble, D. N. (1984). Children's understanding of dispositional characteristics of others. *Child Development*, 55(2), 550–560. <https://doi.org/10.2307/1129966>
- Schmidt, M. F. H., Svetlova, M., Johe, J., & Tomasello, M. (2016). Children's developing understanding of legitimate reasons for allocating resources unequally. *Cognitive Development*, 37, 42–52. <https://doi.org/10.1016/j.cogdev.2015.11.001>
- Tomasello, M. (2014). *A natural history of human thinking*. Harvard University Press.
- Tomasello, M. (2016). The ontogeny of cultural learning. *Current Opinion in Psychology*, 8, 1–4. <https://doi.org/10.1016/j.copsyc.2015.09.008>
- Tomasello, M. (2019). *Becoming human: A theory of ontogeny*. Harvard University Press.
- Zhang, X., Anderson, R. C., Dong, T., Nguyen-Jahiel, K., Li, Y., Lin, T.-Z., & Miller, B. (2013). Children's moral reasoning: Influence of culture and collaborative discussion. *Journal of Cognition and Culture*, 13(5), 503–522. <https://doi.org/10.1163/15685373-12342106>