Cognitive Development 70 (2024) 101450

Contents lists available at ScienceDirect

Cognitive Development

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

Morally-relevant theory of mind is related to viewing gender inequalities as unacceptable

Jacquelyn Glidden, Kathryn M. Yee, Melanie Killen *

University of Maryland, USA

ARTICLE INFO

Keywords:
Theory of mind
Gender inequality
Emotions
Fairness
Resource allocation

ABSTRACT

Previous research has shown that morally-relevant theory of mind enables children to avoid blaming a peer for an accidental transgression. The current study investigated whether this form of theory of mind helps children recognize that gender inequalities are unfair and create negative emotional experiences. Further, the study examined this ability across three perspectives (for themselves, for those who have been advantaged by inequality, and for those who have been disadvantaged by inequality). Participants were 141 children (M_{Age} = 6.67 years, 49% female, 32% ethnic/racial minority) recruited from the mid-Atlantic region of the U.S. Experience with the negative consequences of gender bias and more advanced mental state understanding was associated with more negative evaluations of gender inequalities and more neutral attributions of others' emotions. These findings shed light on the role of different forms of mental state understanding in children's evaluations of inequalities based on gender.

1. Introduction

Over the past decade, researchers have integrated two lines of work, one examining children's developing capacity to understand others' mental states and one examining children's moral judgments (Andrews & Talwar, 2021; Glidden et al., 2021; Gönültaş et al., 2022; Killen et al., 2011; Lagattuta & Kramer, 2022). This work has highlighted the importance of understanding how children's mental state understanding contributes to their capacity to make moral judgments. It is particularly valuable for researchers to be able to measure children's understanding of intentions in social contexts, as this is a particularly relevant context for moral judgments.

Traditionally, children's understanding of intentions has been measured using a prototypic false belief theory of mind scenario. In this measure, children witness a change in the location of an object and must infer that someone who did not witness the change of location will look in an incorrect location (Wellman & Liu, 2004). Killen et al. (2011) investigated the intersection of intention understanding and moral judgments by administering a modified false belief task in which the act involved social and moral considerations. The social and moral considerations inserted into a false belief task involved property ownership and property damage. Specifically, a scenario was described to children in which a child's highly valued object (a cupcake) was placed in a bag and left on a classroom table while the child left the room. During the child's absence, a classroom helper cleaning up transferred the bag from the table to a trash unaware of the contents of the bag. The owner returned to find their bag (and special cupcake) in the trash can. Embedded components of the false belief assessment included where the object's owner would look for the bag when they returned to

* Correspondence to: Department of Human Development and Quantitative Methodology, University of Maryland, College Park, 3942 Campus Drive, Suite 3304, College Park, MD 20742–1131, USA.
E-mail address: mkillen@umd.edu (M. Killen).

https://doi.org/10.1016/j.cogdev.2024.101450
Received 4 May 2023; Received in revised form 6 March 2024; Accepted 28 April 2024
Available online 9 May 2024

the classroom and whether the classroom helper knew what was in the bag. Social and moral assessments included evaluations of the act and attributions of intentions of the classroom helper as well as other assessments about emotion attributions and intentionality. Children's ability to recognize that the classroom helper did not *know* what was in the bag was a more significant predictor of children's view that the accidental transgressor should not be blamed for a wrong-doing than was the prototypic false belief ToM judgment. Thus, children's morally relevant theory of mind (MoToM) skills served as a useful predictor for their moral judgment in a situation involving an ambiguity of intentions and a lack of information in a complex context.

Other researchers have replicated and extended these findings by demonstrating that MoToM measures serve as significant predictors of moral judgments and more so than prototypic ToM measures (Andrews & Talwar, 2021; D'Esterre et al., 2019; Glidden et al., 2021). One such study showed that MoToM skills were significantly related to differentiation of intentional and unintentional falsehoods (D'Esterre et al., 2019), and more so than prototypic false belief assessments. Children with MoToM competence who witnessed an unintentional falsehood when a character claimed to need a resource that they unknowingly already possessed predicted more favorable evaluations of an unintentional transgressor than an intentional transgressor who claimed a need for a resource that they already possessed. Further, children made more accurate attributions of intentions and punishment, above and beyond both age and prototypic ToM.

Using a MoToM measure of children's mental state understanding allows researchers to directly measure how mental state knowledge is being applied to a context with social and moral content. Consequently, the MoToM measure is distinguished from the prototypic ToM tasks in several ways. First, the goal of the MoToM measure is to specifically assess children's understanding of mental states *within* social contexts involving moral cognition and decision making. This contrasts with prototypic ToM measures which specifically measure ToM in acontextual scenarios, with very little information about why actions are being performed or the relationship between the two actors. Children are rarely, if ever, in situations where they need to infer intentions without having social and contextual information about the actor. Previous work examining MoToM skills has mainly focused on understanding of intentions by asking children to infer whether an act was done on purpose (or by accident), but mental state understanding consists of more than understanding others' intentions in a situation in which the actor had good intentions with a bad outcome (or the reverse).

More recently, research has investigated the role of children's mental state knowledge regarding resource allocation and preferences for equity. This work has shown that greater mental state knowledge is often associated with children's rejections of unequal resource allocations, including advantageous allocations that benefit the self (Tsoi & McAuliffe, 2020) and those that are based on stereotypic gender expectations (Rizzo & Killen, 2018). This work has shown that mental state knowledge helps children allocate more fairly. While some research has focused on resource allocation, much less is known about whether MoToM competence enables children to reject unequal resource allocation based on a gender ingroup bias.

Thus, to fill this gap, the current study extended previous MoToM research by shifting away from an accidental transgression scenario used in several prior studies to focusing on a different type of mental state capacity. Specifically, the MoToM measure in the current study required children to set aside their own knowledge that a character is gender biased and predict what a new character will think and how they will feel. This MoToM measure required morally-embedded (within the context of fairness) access to knowledge (a new character does not know that another character is biased) to predict another's beliefs (how the new character thinks prizes will be distributed). This application of MoToM skills has the potential to expand our understanding of children's contextually embedded ToM skills.

1.1. Theoretical perspective

The theoretical model that motivated this study was the social reasoning developmental (SRD) model which draws on theories from developmental psychology (social domain theory) and social psychology (social identity theory) to investigate children's evaluations of intergroup decisions and the extent to which their evaluations are grounded in reasoning about social norms, morality, group identity, and mental state knowledge (Elenbaas et al., 2020; McGuire et al., 2018; Rutland & Killen, 2015). With respect to mental state knowledge, SRD theory has asserted that MoToM serves as a significant underlying mechanism accounting for change. Similar to Piaget, (1932, 1977) who asserted that cognitive development accounts for change more than chronological age, SRD posits that recognizing others' mental states in the form of attributions of intentions of others along with social experiences related to intergroup contact can serve as a developmental mechanism.

Previous evidence to support this proposition includes a study by D'Esterre et al. (2019) which showed that MoToM predicted children's responses for all assessments of intentional and unintentional falsehood above and beyond age as well as prototypic ToM. Further, a recent study found that MoToM skills served as a mediator between ingroup bias regarding team membership and children's social exclusion judgments (Glidden et al., 2021). Children were inducted into a team as part of a pumpkin growing contest. It was theorized that MoToM skills would serve as a specific and ecologically valid measure of mental state understanding within the competitive intergroup contexts. The results revealed that children with better MoToM skills were less likely to have ingroup bias contribute to their decisions to exclude a peer in contrast to children with low MoToM skills, who were more likely to do so (Glidden et al., 2021). In the Glidden et al. (2021) study, the authors stated that chronological age was intentionally not included in the analyses because the focus was on the underlying mechanism of change, which was hypothesized to be MoToM competence. What has not been investigated is whether MoToM skills are significantly related to children's evaluations of gender inequalities; nor has this line of research examined connections between MoToM and emotion attributions of those experiencing gender inequalities. Thus, a central goal for the present study was to examine a new focus for MoToM skills in relation to evaluations and emotion attributions regarding gender inequalities.

1.2. Children's understanding of gender biases and inequalities

The decision to focus on gender biases and inequalities was based on research that has demonstrated that, from very early in development, gender is a salient social category for children (Bigler & Liben, 2007; Horn & Sinno, 2013). Research has simultaneously shown that children often reject explicit exclusion and discrimination based on gender (Theimer et al., 2001), but display gender-ingroup biases when allocating and sharing resources (Renno & Shutts, 2015). Moreover, an emerging and robust body of research has focused on children's conceptions and awareness of social inequalities (Heck et al., 2022). Early in life children become aware of inequalities and demonstrate a desire to rectify unequal distributions of resources (Elenbaas et al., 2016; Li et al., 2014; Paulus, 2015). When distributing resources themselves, children often prefer equal distributions and will sometimes discard resources rather than create inequalities of resources, known as inequity aversion (Blake & McAuliffe, 2011; McAuliffe et al., 2013; Shaw & Olson, 2012). Children also judge others' uneven distributions negatively (Cooley & Killen, 2015). Throughout childhood, children prefer fair distributions of resources and will even sacrifice resources in order to achieve an equal outcome (McAuliffe et al., 2013). It remains an important task for developmental scientists to investigate the various contextual factors that lead children to perpetuate or reject gender biases, especially in the context of resource allocations. In this study, then, we focused on two outcome measures regarding gender inequalities, children's evaluations as well as emotional attributions of gender inequalities for those from different perspectives, advantaged and disadvantaged.

1.3. Emotion attributions following inequalities

Attributions of emotions in moral contexts are influenced by a variety of factors, such as situational factors, group norms, and social hierarchies (Killen & Malti, 2015; Smetana & Ball, 2018). Shared group membership can moderate children's likelihood of attributing positive emotions to others who are excluded. For example, in one study, adolescents who were members of a disadvantaged social group were more likely to attribute positive emotions to an excluder than were adolescents who were members of the advantaged social group (Malti et al., 2012). Research has not yet examined the emotions that children attribute to peers who are socially and contextually advantaged or disadvantaged by a specific action, such as a gender-biased resource allocation, or the role of morally-relevant theory of mind skills in these emotion attributions.

When evaluating gender ingroup inclusion preferences in the absence of an advantaged or disadvantaged outcome, children attribute positive emotions to both themselves and to the gender ingroup peer they include (Peplak et al., 2017). Specifically, 4- and 8-year-old children overwhelmingly chose to include a gender ingroup peer (as opposed to a gender outgroup peer), and when asked how they felt about including the gender ingroup peer and how the gender ingroup peer felt about being included, the majority of children expected happy feelings for both themselves and the target. This provides some evidence that children are likely to attribute positive emotions to included ingroup members.

In the context of transgressions, however, young children often fail to attribute negative emotions to transgressors who achieve a gain (e.g., push someone off a swing to get a turn) instead attributing happy emotions, termed the "happy victimizer effect" (Arsenio, 2014). By the ages of 6 and 7, children begin to realize that transgressions often result in negative emotions, such as sadness and guilt, for both the victims and the transgressors. Arsenio (2014) has attributed this age-related change to one that involves an increasing coordination of judgments, such as being able to recognize that a transgressor may feel happy for their gain but ambivalent about the victim's emotion experience. Similarly, someone being advantaged by an unfair allocation of resources may feel happy to gain resources but ambivalent about someone who has been disadvantaged.

Thus, while very young children consistently fail to attribute negative emotions following moral transgressions, older children acknowledge that transgressions often result in negative emotions (Chilver-Stainer et al., 2014; Gasser et al., 2013). Yet, inconsistent age findings across studies may suggest that other cognitive skills, like MoToM skills, or context play a role in how children attribute emotions. No research that we know of has examined whether MoToM skills bear on children's attribution of emotions in gender inequality contexts. Investigating emotional attributions of those experiencing inequalities requires differentiating whether individuals are experiencing advantaged or disadvantaged status within the unequal context.

1.4. Advantaged and disadvantaged perspectives within inequality contexts

Individuals experience different positions within inequalities, with some occupying higher-status positions, referred to as *advantaged* perspectives, and others occupying lower-status positions, referred to as *disadvantaged* perspectives. In the context of resource inequalities, researchers have used the term advantaged to refer to individuals who have more resources than others and the term disadvantaged to refer to individuals who have fewer resources than others (Elenbaas et al., 2020; McGuire et al., 2018; Rizzo & Killen, 2020). Research has shown that children are more likely to rectify inequalities when they are personally disadvantaged by them compared to when they are personally advantaged by them (Blake & McAuliffe, 2011; Rizzo et al., 2023). In this study, we examined whether MoToM competence was related to multiple perspectives of being advantaged or disadvantaged.

1.5. The present study

The goal of the present study, then, was to investigate whether and how children's MoToM mental state understanding was related to children's evaluations of gender inequalities when a character displayed an intentional gender bias. This study also posed a new research question regarding children's attributions of emotions by focusing on both characters who benefitted and characters who

were disadvantaged by gender bias. Emotion attributions were systematically analyzed for these multiple perspectives. This novel examination of children's MoToM skills helps to inform our understanding of the cognitive mechanisms at play when children encounter bias.

Specifically, the study determined whether and how children's cognitive (MoToM) skills and experiential factors (experiencing gender advantage or disadvantage) were related to their evaluations of gender-biased resource allocations and their attributions of others' emotions following inequalities, across three perspectives. Children witnessed an intentional resource inequality: a prize distributor declared that they hold an ingroup gender preference and then distributed resources unfairly based on gender. After witnessing this intentional inequality, children assessed the expectations and mental states of others who were unaware of the inequality, evaluated the inequality, and then attributed emotions for themselves and others who were advantaged and disadvantaged by the inequality.

1.6. Hypotheses

There were two sets of hypotheses. The first set pertained to evaluations of gender inequalities for the three perspectives and how MoToM competence was related to these evaluations; the second set focused on the emotion attributions regarding gender inequalities for the three perspectives and the role of MoToM competence.

Hypotheses about judgments about gender inequalities. The first set of hypotheses tested the prediction that children would judge gender-biased resource inequalities differently depending on their perspective (own, advantaged characters', and disadvantaged characters'). Specifically, based on previous work (Rizzo & Killen, 2020), we predicted that participants would expect characters who were advantaged by their gender to judge the inequality as more acceptable and characters who were disadvantaged by their gender would judge the inequality as less acceptable (H1a). Further, participants' own judgments of the inequality would depend on their status as being advantaged or disadvantaged by the inequality (condition). Children who were advantaged would judge the inequality as more acceptable than children who were disadvantaged (H1b). Based on work showing that MoToM skills impact children's moral judgments (Glidden et al., 2021), we hypothesized that children's MoToM skills would be related to their judgments. Specifically, we hypothesized that children with more advanced MoToM skills would judge gender-biased resource inequality as less acceptable than children with less advanced MoToM skills (H1c).

Hypotheses about emotion attributions about gender inequalities. The second set of hypotheses tested the prediction that, following gender-biased resource inequalities, children would attribute different emotions to the self and others depending on their advantaged or disadvantaged status. Based on Malti and colleagues (2012) work showing that adolescents attribute positive emotions to characters who are advantaged because of their nationality, we predicted that participants would attribute more positive emotions to characters advantaged by their gender than those who were disadvantaged by their gender (H2a). Further, we expected that children's experience being advantaged or disadvantaged by gender inequalities would be related to their emotion attributions across perspectives such that children who were advantaged would feel more positive than children who were disadvantaged (H2b). Based on findings that 6–7 year old children are more likely to attribute negative or neutral emotions following a transgression, while younger children attribute positive emotions (Arsenio, 2014), we hypothesized that children with more advanced MoToM competencies would expect that characters who were advantaged by an unfair resource allocation would feel emotionally ambivalent (e.g., neutral or negative even though they experienced something positive: receiving more prizes) while children with less advanced MoToM skills would attribute more positive emotions (H2c).

2. Method

2.1. Participants

Participants were 141 children between the ages of 4 and 10 years ($M_{Age} = 6.67$ years, $SD_{Age} = 1.88$; 49% female) recruited in person from preschools and summer camps prior to March 2020 ($n = 109$) and online using Zoom during 2020 ($n = 32$). All participants were interviewed one-on-one, face-to-face, with identical formats. Recruitment for online interviews was conducted in the same geographic region as the in-person interviews. Sample size was determined using a priori power analyses using G*Power (Faul et al., 2009), which revealed that in order to detect small to medium effects, a minimum of approximately 100 participants would be necessary to test the hypotheses; additional participants were tested to adequately fill the cells for each condition. Participants were ethnically diverse (68% European American, 11% African American, 6% Asian American, 6% Hispanic, 6% Multiracial, 3% chose not to respond) and came from middle to upper-middle income families in the Mid-Atlantic region of the United States. Data collection occurred between September 2019 and August 2020.

2.2. Design

We investigated the relationship between morally-relevant Theory of Mind (MoToM) skills, experience being advantaged or disadvantaged due to gender (condition, between subjects), and three perspectives: child's own, advantaged characters,' disadvantaged characters' (within-subjects). The focus of this paper was on MoToM as the developmental variable given our theory that age is a proxy for MoToM. We ran the models for age as well and have included them in the Supplemental Materials (Table S1). Three types of judgments: an evaluation (is it okay or not?), a prediction about how the characters would feel after being advantaged or disadvantaged by the gender bias, and a justification for their decision (why?) were assessed. (Justification data were not included in this

paper as it would considerably lengthen the manuscript and added complexity that was beyond the scope of the focus for this paper.).

2.3. Procedure

This project was approved by the Institutional Review Board at University of Maryland, #1482371. All participants received written or digital parental consent to participate and gave verbal assent prior to study administration. Trained research assistants individually administered the task to all participants. In-person interviews were conducted in a quiet space in participants' schools and online individually administered (face-to-face with camera on) interviews were conducted over Zoom. All interviews lasted approximately 20–30 min. The research assistants read the children stories from a script which was presented using a brightly illustrated PowerPoint presentation on a laptop computer (via screenshare in Zoom for online participants). Researchers used a printed protocol to record children's Likert-type response and all sessions were audiotaped.

Participants were first introduced to a 6-point Likert-type scale and were trained on its use. Once children demonstrated their comprehension of the scale and were able to reliably and comfortably use the midpoints and both endpoints of the scale, the researcher began the interview.

Group Assignment. All participants were told they were invited to join an online puzzle club where they could solve puzzles to win prizes. Children were introduced to their gender-matched teammate and an opposing team of a different gender. The two groups were explicitly labeled as the "Girls group" and the "Boys group", which is consistent with previous research examining children's understanding of gendered groups (Mulvey et al., 2016; Rizzo & Killen, 2020).

Puzzles and Conditions. Participants were told the puzzle competition would be a Spot-the-Difference puzzle. A practice puzzle was presented, where participants practiced identifying differences between two black and white images (see Fig. 1). All participants identified at least three differences in the practice trial. Then, participants were told it was time for the competition and started the first test puzzle. Research assistants recorded the number of differences that participants could find in 90 s. If a participant could not find at least three differences in the 90 s period, the research assistant helped the child until they found three differences. At this point the research assistant announced, "Great job!" and the child was told that both teams had tied and found the same number of differences. Children were then introduced to the prize distributor Alex. The prize distributor was always gender biased, giving more prizes to the team that was gender-matched to Alex. In the Advantaged Condition, participants were gender matched to Alex and would receive more prizes. In the Disadvantaged Condition participants were not gender matched to Alex and would receive fewer prizes. For example, a girl in the Disadvantaged Condition would see Alex as a boy and Alex would distribute two prizes to the girls' group and four prizes to the boys' group.

Participants completed two puzzle competitions, each with the same uneven prize distribution favoring Alex's gendered team, and then answered two memory check questions: 1) Who did a better job on the puzzles? and 2) Who has more prizes? If participants answered incorrectly, they were shown the prize distribution slides again and asked the question up to two additional times. All participants were able to correctly answer the memory checks within three repetitions.

After completing the puzzle competition and memory checks, participants answered questions about their own judgments of the resource inequality and attributed emotions for themselves. Participants then moved into the second phase of the study where they witnessed a third-party puzzle competition. First, participants were informed that Alex is always gender biased towards his or her ingroup. When Alex is depicted as a girl then participants hear the following:

> "After the puzzle competition Alex says: I always give more prizes to the girls, because I'm a girl. I *never* give the prizes based on who won. I *always* give more prizes to the girls."

When Alex was depicted as a boy then the desire for his allocation of the prizes was to give more to the boys. Then participants met the two new teams and were told "Both teams are brand new to the puzzle competition. They have never met Alex or watched Alex give out the prizes." Following this, participants watched the two new teams participate and tie in the puzzle competition and then receive an unequal resource distribution. Next, participants answered morally-relevant Theory of Mind (MoToM) questions about the

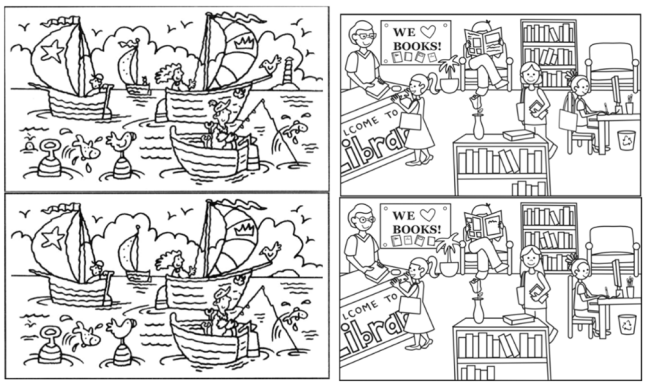

Fig. 1. *Spot-the-Difference Puzzle Stimuli.*

members of the new teams, followed by the judgments and emotion attribution questions for the advantaged characters and disadvantaged characters. The order of the conditions was fixed with participants viewing the self-condition prior to the advantaged or disadvantaged condition. Thus, there were four PowerPoint documents reflecting each of four protocols: Female/Advantaged, Female Disadvantaged, Male/Advantaged, and Male Disadvantaged.

2.4. Assessments

Judgment of the Inequality. Children judged the acceptability of the resource inequality three times throughout the task to measure three perspectives: 1) for themselves, 2) for advantaged characters, and 3) for disadvantaged characters. The first judgment of inequality question was worded as: "How OK or not OK ***do you think*** it is that some kids got more prizes than others?" as shown in Fig. 2. The second question for the advantaged characters was worded as: "How OK or not OK ***does Morgan think*** it is that some kids got more prizes than other kids?" The third question for the disadvantaged character was worded as: "How OK or not OK ***does Jordan think*** it is that some kids got more prizes than other kids?" Participants responded on a six-point Likert-type scale from 1 (Really not OK) to 6 (Really OK).

Emotion Attributions. Children attributed emotions across three perspectives: 1) for themselves, 2) for advantaged characters, and 3) for disadvantaged characters. Specifically, children were asked "How good or bad do you/[Advantaged Character]/[Disadvantaged Character] feel about getting X prizes?". Participants responded on a six-point Likert-type scale from 1 (Really Bad) to 6 (Really Good).

Morally-relevant Theory of Mind (MoToM). Between each participants' own puzzle competition and their evaluation of a third-party puzzle competition, children answered two MoToM questions. After being introduced to the two new competition teams and informed that neither team knew Alex nor knew how Alex would give out the prizes, the participant was informed that both teams had tied in the puzzle competition. Then children were asked, "What does Morgan think? How does Morgan think Alex will give out the prizes?", where Morgan and Jesse were members of the same team. The MoTom questions were presented as a forced choice option such that participants could choose: 1) one prize to the girls and 3 prizes to the boys; 2) two prizes to each group; or 3) three prizes to the girls and 1 prize to the boys. Pictures accompanying these questions display the option for prizes (see Fig. 3). Participants were asked this question twice, once for the girls' team and once for the boys' team. Participants could respond correctly by saying that the characters would expect equal prize distributions following a tie or could respond incorrectly by saying that the characters expected a biased prize distribution.

These questions served as a morally-relevant False Belief question, where the participants had to set aside their own knowledge that Alex would distribute prizes in a gender-biased way and predict that the character would expect equal prizes to be given out after a tie. The questions were scored as fail (0) or pass (1). Participants who did not pass either MoToM question were given a score of 0 for "Lowest"; a score of 1 was "Middle" and a score of 2 was "highest."

Debrief. At the end of the interview session, the interviewer provided a debrief for the girl and boy protocols by stating: "Look! Here's Alex again! Alex realized that s/he was wrong to give more prizes to the girls/boys. S/he wants to apologize and promises to be fair from now on!".

3. Results

Though we had no hypotheses or predictions of differences for children who participated in the study online versus in-person, we conducted preliminary data analyses to compare the response patterns of both groups. To examine the potential, confound of testing online vs in-person, we ran all analyses separately for both groups of participants. The analyses revealed no significant differences for the main effects and interactions whether children were interviewed in-person or on a face-to-face Zoom video call. The mean differences were in the same direction. Therefore, these two groups of participants were combined for all analyses.

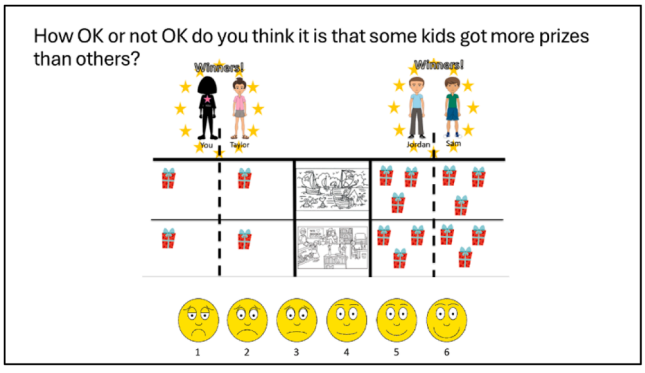

Fig. 2. *Picture Accompanying the Judgment of Inequality for the Self Perspective.*

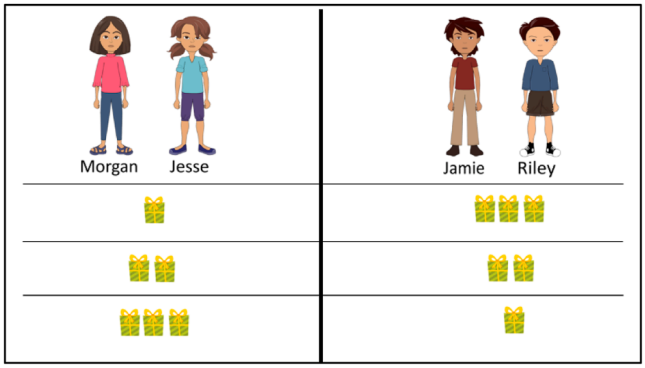

Fig. 3. *Picture Accompanying the Morally Relevant Theory of Mind (MoToM) Assessment. Note*: This picture was part of the PowerPoint slide for the question: *How Does Morgan Think Alex will Give out the Prizes?*.

3.1. Judgments of the inequality across perspectives

We tested whether participants would judge the gender resource inequality differently for the three perspectives by conducting an ANOVA with perspective as a repeated measure (own perspective, advantaged perspective, and disadvantaged perspective) and with the inequality condition (disadvantaged, advantaged) and morally-relevant Theory of Mind (MoToM) included as covariates. This analysis revealed significant effects of perspective, $F(2, 128) = 49.49$, $p < .001$, $\eta_p^2 = .28$; a perspective by condition interaction, $F(2, 128) = 3.31$, $p < .05$, $\eta_p^2 = .03$, and a significant effect for MoToM, $F(2, 128) = 9.65$, $p < .001$, $\eta_p^2 = .13$. Follow up analyses with Bonferroni corrections are reported next.

Main Effect of Perspective. Confirming H1a, participants more negatively evaluated the inequality ($M = 2.21$, $SE = 0.14$) compared to how they expected the advantaged characters would evaluate the inequality ($M = 3.67$, $SE = 0.15$, $p < .001$). There were no differences between participants own judgments and how they predicted disadvantaged characters would judge the inequality ($M = 2.21$, $SE = 0.14$, $p > .05$). Participants' evaluations of an inequality were aligned with how they expected disadvantaged characters would feel, which was more negative than how they expected advantaged characters would feel.

Condition by Perspective Interaction. Children's experience being advantaged or disadvantaged by gender inequalities was

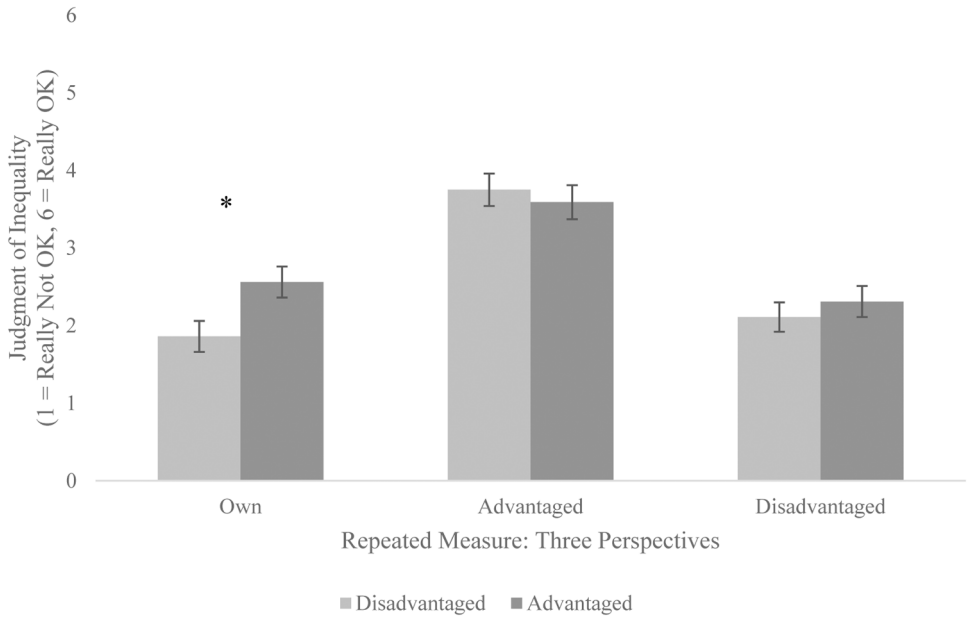

Fig. 4. *Children's Judgment of the Gender Inequality as a Function of Being Advantaged or Disadvantaged. Note.* Three perspectives with gender inequality are presented: 1) Own = Participant as advantaged or disadvantaged evaluating their own inequality; 2) Advantaged = Participant as advantaged or disadvantaged evaluating the inequality for the third party Advantaged character; and 3) Disadvantaged = Participants as advantaged or disadvantaged evaluating the inequality for the third party Disadvantaged character. Children's experience with gender inequality affected their judgments only when considering their "own" perspective, but not when they experienced being advantaged or disadvantaged for other characters. *Significant at $p < .05$.

related to their judgments of the inequality across perspectives (see Fig. 4). For the perspective by condition interaction, there were only significant effects within children's own perspectives, not when participants were predicting judgments for other advantaged and disadvantaged characters. Partially confirming H1b, within the own perspective, participants in the disadvantaged condition judged the inequality as worse ($M = 1.86$, $SE = 0.20$) than participants in the advantaged condition ($M = 2.56$, $SE = 0.20$, p=.01). Interestingly, experience with advantage/disadvantage due to gender did not affect children's judgment predictions for advantaged (p=.59) and disadvantaged characters (p=.47). Children's own experiences (condition) was influential on their judgments, but participants own experience did not bear on their predictions for others' judgments.

Morally-Relevant Theory of Mind. Confirming H1c, participants with the lowest MoToM skills rated the inequality as more acceptable ($M = 3.26$, $SE = 0.17$) than participants with mid-level MoToM skills ($M = 2.55$, $SE = 0.21$, p=.03) and participants with the highest MoToM skills ($M = 2.28$, $SE = 0.16$, $p < .001$). There was no significant difference in judgment of the inequality for participants with mid-level and high MoToM skills (p=.89). Having less developed MoToM skills was associated with children judging a gender-biased inequality more positively.

Summary. These findings supported our hypotheses that children's MoToM skills and experience being advantaged or disadvantaged due to their gender would affect their judgments of inequalities, and that these judgments would differ across perspectives (child's own, advantaged characters', disadvantaged characters'). MoToM skills were associated with children's judgments of the inequality such that advanced MoToM skills were positively related to children judging the inequality more negatively. A perspective by condition interaction showed that children's own experience with gender inequality was related to children thinking about the inequality differently: though both participants with experience being unfairly.

advantaged and disadvantaged due to gender viewed the inequality as unacceptable, children who experienced disadvantage viewed the inequality as more unacceptable. Thus, both experience with inequality and MoToM skills were influential for children's judgments of a gender-biased resource inequality.

3.2. Emotion attributions for inequalities across perspectives

We tested whether participants would attribute emotions differently for the three perspectives by conducting an ANOVA with perspective as a repeated measure (own perspective, advantaged perspective, and disadvantaged perspective) and with the inequality condition (disadvantaged, advantaged) and morally-relevant Theory of Mind (MoToM) included as covariates. As expected, this analysis revealed significant effects of perspective, $F(2, 128) = 214.54$, $p < .001$, $\eta_p^2 = .62$; a perspective by condition interaction, $F(2, 128) = 50.25$, $p < .001$, $\eta_p^2 = .28$; MoToM, $F(2, 131) = 3.34$, $p = .04$, $\eta_p^2 = .05$; and a perspective by MoToM interaction, $F(4, 128) = 2.82$, $p = .03$, $\eta_p^2 = .04$.

Main Effect of Perspective. Confirming H2a, participants felt not as happy ($M = 3.75$, $SE = 0.13$) as they expected advantaged characters felt ($M = 5.15$, $SE = 0.10$, $p < .001$), but happier than they predicted disadvantaged characters felt ($M = 2.13$, $SE = 0.10$, $p < .001$). Further, the predictions for advantaged and disadvantaged characters were also significantly different from each other ($p < .001$). All three emotion attribution values significantly differed from each other, with participants believing that advantaged

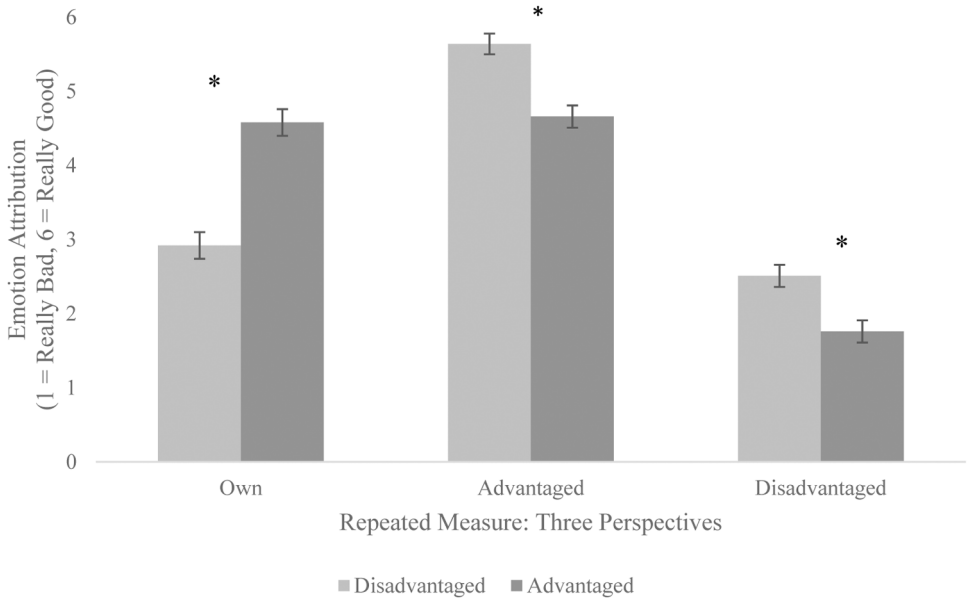

Fig. 5. *Children's Emotion Attributions about Experiencing Gender Inequality for Three Perspectives: Own, Advantaged, Disadvantaged. Note***.** Children's experience being advantaged or disadvantaged due to a gender inequality affected their emotion attributions in all three perspectives (own, advantaged characters', and disadvantaged characters'). *Significant at $p < .05$.

characters felt happiest, participants themselves feeling slightly happy (regardless of condition), and disadvantaged characters feeling the worst.

Perspective by Condition interaction. Children's experience being advantaged or disadvantaged by gender inequalities was related to their emotion attributions across perspectives, confirming H2b (see Fig. 5). Regarding participants own emotion attributions, children who had been disadvantaged due to gender bias felt worse ($M = 2.92$, $SE = 0.18$) than children who had been advantaged ($M = 4.58$, $SE = 0.18$, $p < .001$). Regarding participants emotion attributions of advantaged characters, disadvantaged participants predicted that advantaged characters would feel better ($M = 5.64$, $SE = 0.14$) than advantaged participants predicted ($M = 4.66$, $SE = 0.15$, $p < .001$). Lastly, disadvantaged participants predicted that disadvantaged characters would feel better ($M = 2.51$, $SE = 0.15$) than advantaged participants predicted ($M = 1.75$, $SE = 0.15$, $p < .001$) as shown in Fig. 5. Overall, participants who had experienced disadvantage due to gender were more optimistic about the other characters, believing both advantaged and disadvantaged characters would feel slightly more positive compared to the predictions of participants who experienced advantage.

Main Effect of MoToM. Confirming H2c, participants with the lowest MoToM skills attributed more positive emotions ($M = 3.93$, $SE = 0.12$) than participants with the highest MoToM skills ($M = 3.55$, $SE = 0.11$, p=.05). There were no differences for participants with mid-level MoToM skills compared to lowest or highest MoToM skills. Having advanced MoToM skills was associated with more neutral emotion attributions, while having the lowest MoToM skills was associated with more positive attributions.

Perspective by MoToM interaction. A central finding for the role of MoToM was revealed by the findings that MoToM skills and emotion attributions differed depending on the perspective (see Fig. 6). Confirming H2c, participants with the lowest MoToM skills felt happier ($M = 4.24$, $SE = 0.20$) than participants with mid-level MoToM skills ($M = 3.44$, $SE = 0.26$, p=.045) and participants with the highest MoToM skills ($M = 3.56$, $SE = 0.19$, p=.43).

for their own perspective. Participants with mid-level and high MoToM did not differ from each other in regard to their own emotion attributions. When attributing emotions to advantaged characters, participants with the lowest MoToM skills predicted that advantaged characters would feel happier ($M = 5.52$, $SE = 0.16$) compared to predictions made by participants with the highest MoToM skills ($M = 4.89$, $SE = 0.16$, p=.02), confirming H2c, as shown in Fig. 6. Further, when attributing emotions to disadvantaged characters, there were no differences in emotion attribution based on MoToM skills. All participants believed the disadvantaged.

characters would feel sad ($M = 2.13$, $SE = 0.18$, $ps = 1.0$, across the MoToM levels). Children's own emotions and their ability to attribute emotions to others was affected by the MoToM skills, but only when the emotions were relatively positive.

Summary. These findings supported our hypotheses that children's MoToM skills and experience being advantaged or disadvantaged due to their gender would be related to their emotion attributions following an inequality and that these emotion attributions would differ across perspectives (child's own, advantaged characters', disadvantaged characters'). Generally, participants believed that advantaged characters felt happiest, participants themselves felt.

slightly happy (regardless of condition), and disadvantaged characters felt the worst. A perspective by condition interaction showed that children's own experience with gender-inequality was associated with them attributing emotions differently: participants

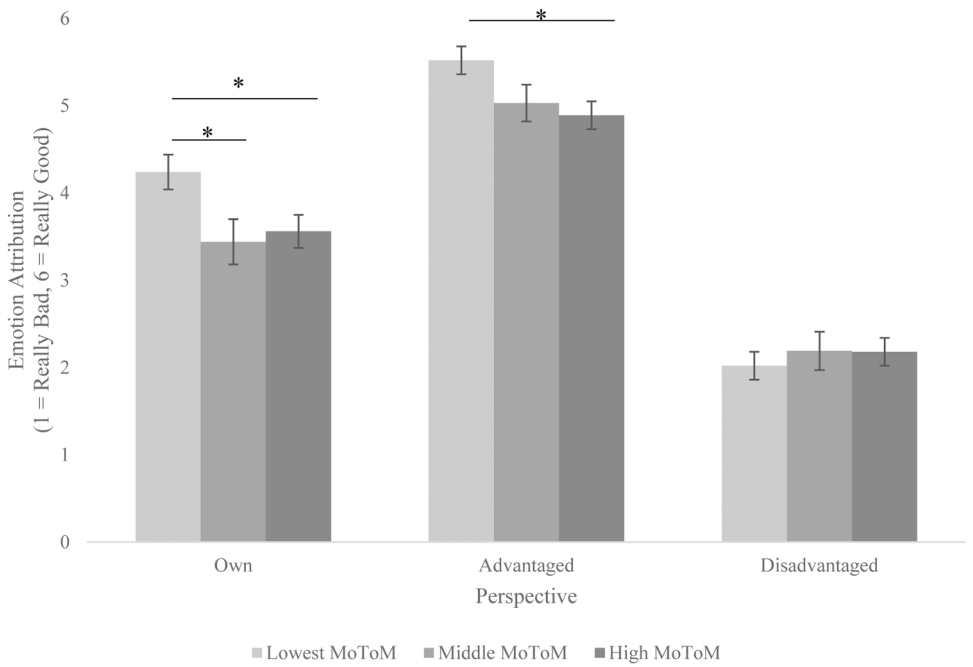

Fig. 6. Morally-Relevant Theory of Mind and Children's Emotion Attributions. *Note.* Children's morally-relevant Theory of Mind (MoToM) skills affected their attributions of emotions following an inequality only when considering their own perspective and the perspective of advantaged characters, not when thinking of other characters who had been disadvantaged. *Significant at $p < .05$.

who experienced disadvantage due to gender were more optimistic about the other characters, believing both advantaged and disadvantaged characters would feel slightly more positive compared to the predictions of participants who experienced advantage. Importantly for the central goal of the study, advanced MoToM skills were related to children attributing more neutral emotions, while the participants with the lowest MoToM skills were more likely to attribute highly positive emotions. A perspective by MoToM interaction revealed that when considering their own emotions and the emotions of advantaged characters, children with the lowest MoToM skills attributed the most positive emotions compared to children with the highest MoToM skills, but there was no effect of MoToM when children were considering the emotions of disadvantaged characters given that the majority of participants expected the disadvantaged characters to feel negative emotions.

4. Discussion

Inequalities, such as being advantaged or disadvantaged based on one's gender, are present throughout the lifespan and there is growing evidence that young children perceive and internalize information about inequalities (Bigler & Liben, 2007; Horn & Sinno, 2013; Li et al., 2014; Rizzo & Killen, 2020). Though we know children witness and experience inequalities in their everyday lives, only recently has research examined how children's cognitive abilities, such as their understanding of other's mental states, contribute to their ability to judge the inequality and consider other's emotions following an inequality. Motivated by the social reasoning developmental (SRD) model (Rutland & Killen, 2015) which proposes that children bring moral, societal, and psychological issues to bear when evaluating intergroup contexts, we examined how children's cognitive skills (their morally-relevant Theory of Mind, MoToM) and their own experience with a gender-biased inequality impacted their ability to judge that same inequality in the future and attribute emotions to others experiencing the inequality. The type of gender inequality we examined was unequal and unfair resource allocation based on a gender ingroup bias explicitly expressed by a resource allocator. As theorized by the SRD framework, the findings revealed that evaluating gender inequalities involves weighing moral (fairness), societal (group dynamics), and psychological considerations (mental state knowledge). Further, SRD asserts that the acquisition of mental state knowledge reflects a developmental mechanism that promotes change which was demonstrated in this study for the first time regarding evaluation and emotion attributions of gender inequalities. These findings contribute to the SRD model by the application of the theory to a new assessment, specifically emotion attributions.

4.1. Morally-relevant theory of mind

For the first time, this study examined the role of children's morally-relevant theory of mind skills in their judgments and evaluations in the context of gender inequality. Specifically, we examined the role of MoToM skills and experience with intentional inequality due to gender ingroup bias on children's inequality judgments and emotion attributions. We found that children's MoToM competence was valuable for their judgment of the gender-biased resource allocation, as well as for their ability to attribute emotions differentially for themselves, others who were advantaged, and others who were disadvantaged.

Specifically, children with less advanced MoToM skills judged the resource inequality as more acceptable compared to children with more advanced MoToM skills. This is in line with previous work showing that mental state understanding helps children to reject unequal resource distributions (Tsoi & McAuliffe, 2020). Our findings extend previous work showing that prototypic ToM capacity helps children reject inequity by focusing on the intersection of mental state understanding and moral evaluations. In the present study, participants who were able to infer mental states more accurately in the context of gender bias were more likely to rate that gender bias as less acceptable than their peers with less accurate mental state understanding.

Children's MoToM skills helped them differentiate emotion attributions. Generally, children with less developed MoToM skills attributed more positive emotions across contexts, but those with more developed MoToM skills attributed less positive and more neutral emotions. MoToM skills were particularly advantageous when children considered their own emotions or emotions of others who were advantaged, but not when thinking about disadvantaged characters. This finding is in line with previous work showing that MoToM skills were not necessary for interpreting straightforward transgressions (Glidden et al., 2021). Specifically, previous work has shown that MoToM skills do not serve as a mediator between group biases and moral decision making when children were presented with straightforward moral transgressions (e.g., cheating on purpose). In the current study, MoToM skills did not assist children in the most straightforward emotion attributions: attributing emotions to children who were disadvantaged by gender bias. In this case, all children predicted the characters would feel poorly after experiencing gender bias. This highlights an interesting pattern in the literature where mental state understanding is more helpful in complex contexts compared to straightforward situations. Yet, complex contexts are situations where it may be more difficult to utilize mental state understanding skills, especially for children just developing these skills.

Previous work had not yet demonstrated the connection between children's MoToM skills and their differential attribution of emotions. While belief-emotion is a component of the initial ToM task (Wellman & Liu, 2004), there has not yet been work showing how children's understanding of mental states in morally-salient contexts is related to their emotion attributions for others. This study showed that children with the least developed MoToM skills attributed more positive emotions to characters that were disadvantaged by gender inequality. Children with advanced MoToM skills attributed neutral emotions, for themselves, others who were advantaged, and others who were disadvantaged. This suggests that more developed MoToM skills allow children to mentally consider that others might not feel happy about gender-biased resource distributions, even when they benefit from it. However, it is not the case that children with more developed MoToM skills attributed negative emotions to those who were advantaged by gender inequality. Adults may feel negatively after benefiting from resource inequalities, especially when those inequalities come at the expense of someone else.

When do children begin to understand or expect this? Future research should investigate when children begin to feel negatively or attribute negative emotions following being advantaged by inequalities.

Recently, researchers have investigated the relationship between children's developing concerns for fairness and their social emotions. This work has found that children younger than 5.5 years old display positive emotion in response to unfair advantages, but by 8.5 years old demonstrate negative emotion in response to unfair advantage (Gerdemann et al., 2022). Findings from the present study suggest that ToM skills may underlie these differences, with older children typically having more advanced ToM skills than younger children. However, future research should continue to explore the variables impacting children's simultaneously developing concerns for fairness and their emotional responses to unfairness.

Experiences with Gender Bias Contextual factors such as inequality in status inform children's moral decision making (Elenbaas et al., 2020; Sims, Yee, et al., 2023). Children consider multiple factors when making moral judgments, especially in intergroup contexts. The current study investigated the role of children's own experiences with gender bias as a contextual factor that might impact their moral judgments and emotion attributions across groups. This is an especially important area of research as there is growing evidence that young children are both aware of and experience bias based on gender (Elenbaas et al., 2020; Perszyk et al., 2019).

In the current study, children's own experience with gender inequality was related to them thinking about the inequality differently: though participants with experience being unfairly advantaged and disadvantaged due to gender viewed the inequality as unacceptable, children who experienced disadvantage viewed the inequality as more unacceptable. This nuanced finding suggests that children's experiences with inequality may inform their future moral judgments and evaluations of that same inequality. In this case, children who experienced the negative consequences of gender bias were more likely to say that same bias was unacceptable later, compared to their peers who benefited from gender bias. Importantly, future research might investigate possible ways of helping children who have been advantaged due to inequality to later reject that inequality.

In past research, children sometimes struggle to reject resource inequalities when the recipients are from different social groups, such as gendered or racial groups (Elenbaas & Killen, 2016; Li et al., 2014; Renno & Shutts, 2015). There is growing evidence that children complexly balance concerns for fairness with their understanding of social relationships and status hierarchies. Recent work looking specifically at gender-biased resource inequalities has shown that children struggle to rectify inequalities when a historically disadvantaged group receives fewer resources (Sims, Yee, et al., 2023). Evidence from the current study suggests that one component of this complex process is children's experience with inequalities. Future research should continue to investigate how children's experience with bias and inequality impact their understanding of inequalities and decisions to rectify inequalities.

Children's experience with gender-biased resource inequality also impacted their emotion attributions. A perspective by condition interaction showed that children's own experience with gender-inequality was associated with them attributing emotions differently: participants who experienced disadvantage due to gender were more optimistic about the other characters, believing both advantaged and disadvantaged characters would feel slightly more positive compared to the predictions of participants who experienced gender advantage. This finding aligns with previous research showing that children who were disadvantaged by gender inequality were more likely to pass prototypic ToM measures afterwards, while children who were advantaged were less likely to pass prototypic ToM (Rizzo & Killen, 2018). These authors suggested that experience being disadvantaged motivated children to think about mental states of others more, perhaps reasoning and thinking about why someone would give them fewer resources based on gender. However, children who were advantaged were not motivated in this way.

In the current study, children who were previously advantaged may be more likely to think that disadvantaged children feel bad about not getting as many prizes while simultaneously down-playing their own experience to compensate. Children who were advantaged by an inequality consistently reported that they did not feel as positive as others' who had been advantaged. Children were making the interesting distinction that some people may feel good about receiving more resources due to a gender inequality, but they personally do not. This also suggests that children are complexly analyzing the situation, and future research should explore how other types of emotions (e.g., anger, guilt, frustration) might bear on evaluations of advantaged and disadvantaged status in a resource inequality context.

Future directions for this line of research may also be informed by current research on psychological distancing pertaining to children's reasoning where children reason better about another person's future than their own (Lee & Atance, 2016) and children's performance on executive functioning tasks increased with self-distancing which might be related to theory of mind (White & Carlson, 2016). In addition, research on evaluative information has demonstrated that moral reasoning is also related to trait judgments which can include emotion attributions. Including a MoToM measure for understanding connections between moral reasoning and trait judgments could shed light on social development (Marble & Boseovski, 2020). Unpacking the relationships between social-cognitive developmental mechanisms such as MoToM and chronological age would shed light on how developmental processes unfold. We recognize that the chronological age of the child, and other characters and experiences contribute to changes in cognitive development. Investigating distinctions between different mechanisms such as chronological age, mental state ability, and cognitive processing skills such as executive functioning is complex. Our goal here was to demonstrate the MoToM was related to viewing gender inequalities as unacceptable. Future research should more thoroughly examine multiple mechanisms of change to partial out the various contributions to children's cognitive development. Finally, including MoToM skills as a factor for enabling children to give priority to moral reasoning in situations with group dynamics would contribute to the literature (Chalik & Rhodes, 2020; Sims, Yee, et al., 2023).

5. Conclusions

This study reported novel findings regarding our understanding of how children's morally-relevant theory of mind (MoToM) skills and experience with gender bias contribute to intergroup moral judgments. This is the first time that research has investigated the role of children's MoToM skills in their judgments and evaluations in a context of bias and inequality. Further, this study utilized a novel MoToM measure, providing a new tool for analyzing children's access to knowledge and belief understanding in a moral context. Interestingly, children's MoToM competency was related to both their judgments of gender-biased inequalities and their emotion attributions following inequalities. Advanced MoToM skills allowed children to both negatively evaluate the inequality and to recognize that being advantaged by a gender inequality does not necessarily result in positive emotions but may leave the individual feeling more neutral. Further, children's own experiences with gender-biased inequality informed their judgments and emotion attributions. Children with experience being disadvantaged by gender inequality appeared more optimistic in their emotion attributions for others who were advantaged and disadvantaged, compared to children who were previously advantaged. It is possible that having experienced the negative outcomes of the inequality gave children a different perspective on the consequences (e.g., receiving fewer prizes did not leave them feeling very emotionally upset) compared to children who were advantaged and did not share the experience (e.g., did not know how bad they would feel if they got fewer prizes).

There are many active lines of research investigating how children think about inequality and reject inequity (Elenbaas et al., 2020; Heck et al., 2022; Killen et al., 2022; McAuliffe et al., 2017). This study adds to this literature by showing that children's moral cognition is an important factor for their understanding of inequality. There are multiple implications for future research, specifically at the intersection of identity, mental state understanding, and bias. The intersection of mental state understanding and experience with bias provides an interesting starting point for intervention work aimed at reducing bias and inequality among children.

CRediT authorship contribution statement

Kathryn M. Yee: Writing – review & editing. **Melanie Killen:** Writing – review & editing. **Jacquelyn Glidden:** Writing – review & editing.

Data availability

Data will be made available on request.

Acknowledgments

We thank the members of the Social and Moral Development Lab at the University of Maryland for their invaluable contributions to this project. We gratefully recognize the research assistants who helped with data collection for this project and would especially like to thank Dylan Cooper, Jacob Glassman, Jessica Monaghan, and Arvid Samuelson. The third author was supported by grants from the National Science Foundation, BCS1728918, and the National Institutes of Health, NICHD R01HD093698 while working on this paper. This study was not preregistered. We are very appreciative to the schools, students, and parents who participated in this study.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cogdev.2024.101450.

References

Andrews, K., & Talwar, V. (2021). Considering morally relevant theory of mind through a social domain theory perspective. *Journal of Moral Education*, 1–15. https://doi.org/10.1080/03057240.2021.1938521

Arsenio, W. (2014). Moral emotion attributions and aggression. In M. Killen, & J. G. Smetana (Eds.), *Handbook of moral development* (2nd ed..,, pp. 235–255). Psychology Press.

Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science, 16*(3), 162–166. https://doi.org/10.1111/j.1467-8721.2007.00496.x

Blake, P. R., & McAuliffe, K. (2011). I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition, 120*(2), 215–224. https://doi.org/10.1016/j.cognition.2011.04.006

Chalik, L., & Rhodes, M. (2020). Groups as moral boundaries: A developmental perspective. In J. B. Benson (Ed.), *Advances in child development and behavior* (pp. 63–93). Elsevier Academic Press. https://doi.org/10.1016/bs.acdb.2020.01.003.

Chilver-Stainer, J., Gasser, L., & Perrig-Chiello, P. (2014). Children's and adolescents' moral emotion attributions and judgements about exclusion of peers with hearing impairments. *Journal of Moral Education, 43*(3), 235–249. https://doi.org/10.1080/03057240.2014.913515

Cooley, S., & Killen, M. (2015). Children's evaluations of resource allocation in the context of group norms. *Developmental Psychology, 51*(4), 554–563. https://doi.org/10.1037/a0038796

D'Esterre, A. P., Rizzo, M. T., & Killen, M. (2019). Unintentional and intentional falsehoods: The role of morally relevant theory of mind. *Journal of Experimental Child Psychology, 177*, 53–69. https://doi.org/10.1016/j.jecp.2018.07.013

Elenbaas, L., & Killen, M. (2016). Children rectify inequalities for disadvantaged groups. *Developmental Psychology, 52*(8), 1318–1329. https://doi.org/10.1037/dev0000154

Elenbaas, L., Rizzo, M. T., & Killen, M. (2020). A developmental-science perspective on social inequality. *Current Directions in Psychological Science, 29*(6), 610–616. https://doi.org/10.1177/0963721420964147

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Gasser, L., Malti, T., & Buholzer, A. (2013). Children's moral judgments and moral emotions following exclusion of children with disabilities: Relations with inclusive education, age, and contact intensity. *Research in Developmental Disabilities, 34*(3), 948–958. https://doi.org/10.1016/j.ridd.2012.11.017

Gerdemann, S. C., McAuliffe, K., Blake, P. R., Haun, D. B. M., & Hepach, R. (2022). The ontogeny of children's social emotions in response to (un)fairness. *Royal Society Open Science, 9*(8), Article 191456. https://doi.org/10.1098/rsos.191456

Glidden, J., D'Esterre, A., & Killen, M. (2021). Morally-relevant theory of mind mediates the relationship between group membership and moral judgments. *Cognitive Development, 57*, Article 100976. https://doi.org/10.1016/j.cogdev.2020.100976

Gönültaş, S., Richardson, C., & Mulvey, K. L. (2022). But they weren't being careful! Role of theory of mind in moral judgments about victim and transgressor negligence. *Journal of Experimental Psychology, 212*, Article 105234. https://doi.org/10.1016/j.jecp.2021.105234

Heck, I. A., Shutts, K., & Kinzler, K. D. (2022). Children's thinking about group-based social hierarchies. *Trends in Cognitive Sciences, 26*(7), 593–606. https://doi.org/10.1016/j.tics.2022.04.004

Horn, S. S., & Sinno, S. (2013). Gender, sexual orientation, and discrimination based on gender and sexual orientation. In M. Killen, & J. G. Smetana (Eds.), *Handbook of Moral Development* (2nd ed., pp. 343–365). Psychology Press. https://doi.org/10.4324/9780203581957-28.

Killen, M., Elenbaas, L., & Ruck, M. D. (2022). Developmental perspectives on social inequalities and human rights. *Human Development, 66*(4-5), 329–342. https://doi.org/10.1159/000526276

Killen, M., & Malti, T. (2015). Moral judgments and emotions in contexts of peer exclusion and victimization. In *Advances in Child Development and Behavior* (Vol. 48, pp. 249–276). Elsevier,. https://doi.org/10.1016/bs.acdb.2014.11.007

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

Lagattuta, K. H., & Kramer, H. J. (2022). Theory of mind and moral cognition: Developmental changes in integrating mental states and moral judgments. In M. Killen, & J. G. Smetana (Eds.), *Handbook of Moral Development* (3rd ed.). Routledge Press.

Lee, W. S. C., & Atance, C. M. (2016). The effect of psychological distancing on children's reasoning about future preferences. *Plos One, 11*(10), Article e0164382. https://doi.org/10.1371/journal.pone.0164382

Li, V., Spitzer, B., & Olson, K. R. (2014). Preschoolers reduce inequality while favoring individuals with more. *Child Development, 85*(3), 1123–1133. https://doi.org/10.1111/cdev.12198

Malti, T., Killen, M., & Gasser, L. (2012). Social judgments and emotion attributions about exclusion in Switzerland. *Child Development, 83*(2), 697–711. https://doi.org/10.1111/j.1467-8624.2011.01705.x

Marble, K. E., & Boseovski, J. J. (2020). Content counts: A trait and moral reasoning framework for children's selective social learning. In J. B. Benson (Ed.), *Advances in child development and behavior* (pp. 95–136). Elsevier Academic Press. https://doi.org/10.1016/bs.acdb.2020.01.004.

McAuliffe, K., Blake, P. R., Kim, G., Wrangham, R. W., & Warneken, F. (2013). Social influences on inequity aversion in children. *PLoS ONE, 8*(12), Article e80966. https://doi.org/10.1371/journal.pone.0080966

McAuliffe, K., Blake, P. R., Steinbeis, N., & Warneken, F. (2017). The developmental foundations of human fairness. *Nature Human Behaviour, 1*(2), Article 0042. https://doi.org/10.1038/s41562-016-0042

McGuire, L., Rizzo, M. T., Killen, M., & Rutland, A. (2018). The development of intergroup resource allocation: The role of cooperative and competitive in-group norms. *Developmental Psychology, 54*, 1499–1506. https://doi.org/10.1037/dev0000535

Mulvey, K. L., Buchheister, K., & McGrath, K. (2016). Evaluations of intergroup resource allocations: The role of theory of mind. *Journal of Experimental Child Psychology, 142*, 203–211. https://doi.org/10.1016/j.jecp.2015.10.002

Paulus, M. (2015). Children's inequity aversion depends on culture: A cross-cultural comparison. *Journal of Experimental Child Psychology, 132*, 240–246. https://doi.org/10.1016/j.jecp.2014.12.007

Peplak, J., Song, J.-H., Colasante, T., & Malti, T. (2017). Only you can play with me!" Children's inclusive decision making, reasoning, and emotions based on peers' gender and behavior problems. *Journal of Experimental Child Psychology, 162*, 134–148. https://doi.org/10.1016/j.jecp.2017.04.019

Perszyk, D. R., Lei, R. F., Bodenhausen, G. V., Richeson, J. A., & Waxman, S. R. (2019). Bias at the intersection of race and gender: Evidence from preschool-aged children. *Developmental Science, 22*(3). https://doi.org/10.1111/desc.12788

Piaget, J. (1932). *The moral judgment of the child*. Free Press.

Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. Viking Press.

Renno, M. P., & Shutts, K. (2015). Children's social category-based giving and its correlates: Expectations and preferences. *Developmental Psychology, 51*(4), 533–543. https://doi.org/10.1037/a0038819

Rizzo, M. T., & Killen, M. (2018). Theory of mind is related to children's resource allocations in gender stereotypic contexts. *Developmental Psychology, 54*(3), 510–520. https://doi.org/10.1037/dev0000439

Rizzo, M. T., & Killen, M. (2020). Children's evaluations of individually and structurally based inequalities: The role of status. *Developmental Psychology, 56*(12), 2223–2235. https://doi.org/10.1037/dev0001118

Rutland, A., & Killen, M. (2015). A developmental science approach to reducing prejudice and social exclusion: Intergroup processes, social-cognitive development, and moral reasoning. *Social Issues and Policy Review, 9*(1), 121–154. https://doi.org/10.1111/sipr.12012

Shaw, A., & Olson, K. R. (2012). Children discard a resource to avoid inequity. *Journal of Experimental Psychology: General, 141*(2), 382–395. https://doi.org/10.1037/a0025907

Sims, R., Yee, K. M., & Killen, M. (2023). Morality and conceptions of social status, inequalities, and group norms. In M. Killen, & J. G. Smetana (Eds.), *Handbook of Moral Development* (3rd ed., pp. 71–87). Routledge.

Smetana, J. G., & Ball, C. L. (2018). Young children's moral judgments, justifications, and emotion attributions in peer relationship contexts. *Child Development, 89*(6), 2245–2263. https://doi.org/10.1111/cdev.12846

Theimer, C. E., Killen, M., & Stangor, C. (2001). Young children's evaluations of exclusion in gender-stereotypic peer contexts. *Developmental Psychology, 37*(1), 18–27. https://doi.org/10.1037//0012-1649.37.1.18

Tsoi, L., & McAuliffe, K. (2020). Individual differences in theory of mind predict inequity aversion in children. *Personality and Social Psychology Bulletin, 46*(4), 559–571. https://doi.org/10.1177/0146167219867957

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*(2), 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

White, R. E., & Carlson, S. M. (2016). What would Batman do? Self-distancing improves executive functioning in young children. *Developmental Science, 19*(3). https://doi.org/10.1111/desc.12314