CHILD DEVELOPMENT

Child Development, May/June 2021, Volume 92, Number 3, Pages 959–975

The Homogeneity and Heterogeneity of Moral Functioning in Preschool

Enda Tan, Amori Y. Mikami, Anastasiya Luzhanska, and J. Kiley Hamlin
University of British Columbia

The current study examined relations between distinct aspects of moral functioning, and their cognitive and emotional correlates, in preschool age children. Participants were 171 typically developing 3- to 6-year-olds. Each child completed several tasks, including (a) moral tasks assessing both performance of various moral actions and evaluations of moral scenarios presented both verbally and nonverbally; and (b) non-moral tasks assessing general cognitive skill, executive functioning, theory-of-mind, and emotion recognition. Shyness and empathic concern were assessed from video acquired during participation. Results demonstrated positive associations among distinct moral actions, as well as among distinct moral evaluation tasks, but few associations between tasks assessing moral actions and moral evaluation. Empathic concern and inhibitory control each emerged as important predictors of preschoolers' moral functioning.

From an early age, children demonstrate a wide range of skills which allow them to navigate the complex social and moral world. As third parties, children evaluate others based on their moral actions, viewing intentional moral transgressions as wrong (Killen, Mulvey, Richardson, Jampol, & Woodward, 2011), judging prosocial individuals to be nice, and allocating punishment to antisocial individuals (Van de Vondervoort & Hamlin, 2017). As first parties, children engage in their own prosocial behaviors, helping others achieve instrumental goals (Svetlova, Nichols, & Brownell, 2010), comforting distressed others (Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992), and sharing resources with others (Warneken, Lohse, Melis, & Tomasello, 2011). The diversity of these morally relevant skills (including moral evaluations and moral behaviors) begs two important questions: Should we view children's moral functioning as a unified construct? And if so, what are the mechanisms that unify these seemingly distinct mental and behavioral processes? The current study aimed to address these questions by assessing the consistency between different aspects of moral functioning, and their cognitive and emotional correlates, in preschool-age children.

This work was supported in part by Social Sciences and Humanities Research Council of Canada (SSHRC) Grant and UBC Hampton Fund Research Grant to K. Hamlin. We thank Shannon Bridson, Hannah de Vries, Vivian Wong, Rachelle Graham, Brianna Vissers, Thaddeus Grams, Adam Shillington, Kristi Atkinson, and Giulia Santaca for their assistance. The data reported in the paper are deposited in the Department of Psychology at UBC. The authors have no conflict of interest to declare.

Correspondence concerning this article should be addressed to Enda Tan, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, British Columbia, Canada V6T 1Z4. Electronic mail may be sent to enda.tan@psych.ubc.ca.

Relations Between Different Forms of Moral Behaviors

Relations Between Types of Prosocial Behaviors

Are there consistencies in children's production of different types of moral behaviors? Some scholars argue that different forms of prosocial behaviors are unrelated, primarily because they are based on distinct mechanisms. For instance, Dunfield (2014) argues that helping, comforting, and sharing are elicited by different negative states (i.e., instrumental need, emotional distress, and unmet material desire, respectively), and that these negative states are represented by different neural mechanisms. In a similar vein, Paulus (2014) proposes that different prosocial behaviors are driven by distinct motivations: While instrumental helping is motivated by a desire to fulfill others' goals or to engage in social interactions, comforting is driven by emotional contagion or empathic concerns, and sharing is driven by compliance to an explicit request, conformity to social norms, and a desire to interact. These theories suggest that, given the diverse mechanisms underlying different forms of prosocial behaviors, prosocial behaviors should not be viewed as a homogeneous construct (Dunfield, 2014; Paulus, 2014).

Empirically, the heterogeneity of prosocial behaviors is supported by studies finding no

 0009-3920/2021/9203-0013
DOI: 10.1111/cdev.13458

significant correlations between helping, comforting, and sharing in both concurrent (Dunfield & Kuhlmeier, 2013; Dunfield, Kuhlmeier, O'Connell, & Kelley, 2011) and longitudinal assessments (Paulus et al., 2015). Furthermore, research shows that helping, comforting, and sharing recruit distinct neural mechanisms (Paulus, Kühn-Popp, Licata, Sodian, & Meinhardt, 2013; Steinbeis, 2018) and have different genetic etiologies (Knafo-Noam, Vertsberger, & Israel, 2018). Third, developmental research demonstrates that helping, comforting, and sharing each emerge at different ages and show divergent developmental trajectories (Dunfield & Kuhlmeier, 2013). Finally, there is evidence that helping, comforting, and sharing are differentially impacted by neurocognitive conditions (e.g., autism) as well as environmental forces (for review, see Dunfield, 2014). Taken together, these studies support the idea that distinct prosocial behaviors should not be considered a unified construct.

Despite the evidence for the multifaceted nature of prosocial behaviors, there are reasons to believe that different forms of prosocial behaviors may share at least some similarities. First, prosocial behaviors are all likely shaped by socialization processes, including explicit teaching, observational learning, and imitation (Dahl & Brownell, 2019). Second, it is proposed that rudimentary aspects of prosocial behaviors are rooted in the evolutionary history of the human species. Specifically, prosocial behaviors are hypothesized to derive from a universal moral instinct grounded in our biology (Hauser, 2006); these prosocial behaviors promote positive interactions between group members and ultimately give rise to large-scale societies (Tomasello & Vaish, 2013). This evolutionarily shaped moral instinct, if exists, will promote consistency between different forms of prosocial behaviors.

Supporting this homogeneity view, two recent studies have found consistencies between helping, sharing, and empathic responses in toddlers (Gross et al., 2015; Newton, Thompson, & Goodman, 2016). For preschoolers, although studies have generally found no relations between helping and sharing (Cassidy, Werner, Rourke, Zubernis, & Balaraman, 2003; Eisenberg, Pasternack, Cameron, & Tryon, 1984; Yarrow et al., 1976), significant correlations have been found between helping and comforting (Eisenberg & Lennon, 1980) and between comforting and sharing (Yarrow et al., 1976). These findings demonstrate the nuances in relations between different varieties of prosocial behaviors, and suggest that there is at least some homogeneity across distinct prosocial behaviors.

Relations Between Prosocial Behaviors and Third-Party Distribution

Aside from helping, sharing, and comforting, children may also demonstrate moral behavior through how they distribute resources between third parties. At the population level, cultural differences have been found: When distributing resources between two similar recipients (who had contributed equally to a task), American children would rather discard extra resources in order to attain equal distributions (Shaw & Olson, 2012), whereas Ugandan children would rather distribute resources unequally than discard extra resources (Paulus, 2015). These findings suggest group-level differences in how children distribute resources among third parties.

At the individual level, however, to our knowledge, no studies have investigated whether children's third-party distributions are related to their first-party prosocial behaviors. We reasoned that there are similarities and dissimilarities between these two types of moral actions: On the one hand, third-party resource distribution and prosocial behaviors are both actions that have moral implications. On the other hand, prosocial behaviors typically require more personal involvement than do third-party distributions. Specifically, prosocial behaviors often require physical activities (e.g., lifting a box), emotional engagement (e.g., hugging the victim), and direct social interactions. In contrast, third-party distributions can be performed in a relatively detached manner. The distribution can be made "behind the scene," and the distributor does not necessarily need to directly (i.e., physically, emotionally, and socially) interact with the recipient. Thus, third-party distribution may or may not correlate with prosocial behaviors. The current study sought to assess the relations between first-party prosocial behaviors and third-party resource distributions, in order to shed further light on the relations between distinct forms of moral behavior.

Relations Between Different Forms of Moral Evaluation

Traditionally, children's moral evaluations have been examined using verbally presented moral scenarios (for review, see Killen & Smetana, 2006). In more recent years, a growing number of studies have used nonverbally presented moral scenarios to explore infant and children's understanding of moral principles (Van de Vondervoort & Hamlin, 2017, 2018). It remains largely unclear whether and how these two forms of moral evaluations are

related to each other. To our knowledge, only one study has examined the consistencies in children's responses to verbally and nonverbally presented socio-moral scenarios. Smetana, Ball, Jambon, and Yoo (2018) showed children (aged 2–5 years) nonverbal (puppet show) and verbal scenarios depicting moral and conventional transgressions, and compared their responses across paradigms. Children who chose the moral (vs. conventional) transgressors as more wrong in the puppet show task were somewhat more likely to view moral (vs. conventional) transgressions as more unacceptable (p = .10) and more rule- and authority-independent (p = .06) in the verbal scenario task. These findings, although marginal, provide preliminary evidence for consistencies in children's moral-conventional distinctions across verbally and nonverbally presented socio-moral scenarios. However, as this study examined children's evaluations of moral versus conventional transgressions, it remains unclear whether children's evaluations within the moral domain itself are consistent across different paradigms. In the current study, we assessed children's evaluations of others' moral behaviors using three different paradigms, including both verbally and nonverbally presented scenarios and conflicting mental states and outcomes. By examining the correlations between these measures, we hoped to gain insights into the relations between various measures of moral evaluations.

Relations Between Moral Behaviors and Moral Evaluations

Previous studies have also explored links between moral behaviors and moral reasoning. These studies typically assess moral reasoning by asking children to make decisions about moral dilemmas, and then categorizing children into moral developmental stages based on how they justify their decisions. Within these paradigms, justifications based on self-serving considerations or punishment avoidance are viewed as less mature, whereas justifications based on empathic concerns or internalized moral principles are viewed as more mature (e.g., Colby & Kohlberg, 1987; Eisenberg & Hand, 1979). In a review of 75 studies examining the relation between moral dilemma reasoning and moral behavior, higher levels of moral reasoning predicted more altruistic behaviors, higher honesty, and lower delinquent behaviors in preschoolers, school-age children, adolescents, and adults (Blasi, 1980; see also Eisenberg & Hand, 1979; Miller, Eisenberg, Fabes, & Shell, 1996). In a more recent meta-analysis, researchers examined 151 studies and found a consistent effect of moral reasoning on moral actions across age (from preschool through adulthood), sex, and methodology (Villegas de Posada & Vargas-Trujillo, 2015). Together, these studies suggest that children's ability to reason about moral dilemmas and to justify their moral decisions is associated with their own prosocial behaviors.

Although there is evidence that how children justify their choices in response to moral dilemmas predicts prosocial behaviors, it remains unclear whether children's basic attitudes toward others' moral actions have similar links with prosocial behaviors. Reasoning about moral dilemmas requires children to explicitly state the logic behind their decisions, and hence requires effortful, deliberative mental processes. By contrast, positive and negative attitudes toward others' moral actions may derive from more intuitive processes. Crucially, a recent study found that children's prosocial behaviors were associated with later controlled, but not early automatic, neural processes (Cowell & Decety, 2015), suggesting that intuitive forms of moral evaluations may be less predictive of prosocial behavior. Consistent with this idea, no associations were found between 5- and 9-year-old children's attitudes toward third-party moral transgressions and their prosocial behaviors (Malti, Gasser, & Gutzwiller-Helfenfinger, 2010). The current study assessed children's evaluations of verbally and nonverbally presented moral scenarios. We predicted no correlations between these evaluations and children's own moral behaviors.

Nature of the Relations Between Different Aspects of Moral Functioning

What explains the relations between different aspects of moral functioning (if they do exist)? One possibility is that these associations reflect overlap in underlying cognitive and emotional mechanisms. To test this hypothesis, it is important to first clarify how basic cognitive and emotional mechanisms are related to different aspects of moral functioning. In the recent literature, individual aspects of moral functioning have been linked to a host of cognitive and emotional factors, including general intelligence (Derryberry, Wilson, Snyder, Norman, & Barger, 2005), executive functioning (Paulus et al., 2015), theory-of-mind (Imuta, Henry, Slaughter, Selcuk, & Ruffman, 2016; Killen et al., 2011), emotion recognition (Grossmann, 2018), empathy (Ball, Smetana, & Sturge-Apple, 2017), and social inhibition (Beier, Terrizzi, Woodward, & Larson, 2017; Karasewich, Kuhlmeier, Beier, & Dunfield, 2018; Smetana et al.,

2012). Although these processes have been examined in relation to local aspects of moral functioning, no studies have yet explored how cognitive and emotional processes influence a wide range of moral behaviors and evaluations within the same children. By assessing the relations between various cognitive and emotional mechanisms and diverse measures of moral development, we can determine which mechanisms play the greatest role(s) in preschool moral functioning. Furthermore, examining whether or not different aspects of moral functioning are associated with the same cognitive or emotional mechanisms would allow us to further examine to what extent they should be seen as related constructs.

The Current Study

The primary goal of the current study was to elucidate the relations between different forms of moral behaviors and moral evaluations in 171 preschool-aged children (aged 3–6). We assessed children's moral functioning in an experimental setting using a wide range of behavioral tasks ($N = 7$). The scope of our measures made our study well suited for exploring the nuanced relations between different aspects of moral functioning in preschool. The comprehensiveness of our measures also allowed us to focus on several relatively underexplored relations, including between third-party resource distribution and other moral behaviors, and between evaluations of verbally and nonverbally presented moral scenarios. It also allowed for further examination of the relation between moral evaluation and moral behavior.

For moral behaviors, we assessed children's instrumental helping (Bryan, Master, & Walton, 2014), verbal and physical comforting to distressed others (Dunfield & Kuhlmeier, 2013), first-party sharing, and third-party resource distribution (Gummerum, Hanoch, Keller, Parsons, & Hummel, 2010). We included the third-party distribution task because this test has the same structure (e.g., similar number of stickers, similar prompts) as the sharing task, but the distribution is made between two third-party recipients. Including this task would allow us to directly compare and explore the relations between first-party sharing and third-party resource distribution.

For moral evaluations, we examined evaluations of moral scenarios presented nonverbally (e.g., via puppet shows) depicting failed attempts to help and harm (Van de Vondervoort & Hamlin, 2018). Furthermore, we assessed children's evaluations of moral scenarios presented verbally, depicting both prototypical moral transgressions and accidental transgressions (Killen et al., 2011). We included the accidental transgression scenario because it has the same structure (e.g., verbally presented, same rating scale) as the prototypic moral transgression scenario, and has a theory-of-mind component that is similar to the nonverbally presented scenario. Including this test could help us better assess the consistencies in children's responses to verbally and nonverbally presented socio-moral scenarios. The diverse moral evaluation tasks included in the current study enabled us to probe children's attitudes toward third-party moral actions presented in different forms (verbal vs. nonverbal), representing different mental states (intentional vs. accidental), and different outcomes (successful vs. failed, positive vs. negative).

An additional goal of the study was to explore whether and how various cognitive and emotional mechanisms relate to different aspects of moral functioning. For this purpose, we assessed children's general cognitive skills, executive functioning, theory-of-mind, and emotional recognition abilities. Our original study design did not include formal measures of social inhibition and empathy. However, these two factors have been shown to be important predictors of preschoolers' moral functioning (Ball et al., 2017; Beier et al., 2017; Karasewich et al., 2018). In order to explore the roles of these processes in explaining preschool functioning, we coded children's signs of shyness and empathic concern from videos acquired during participation.

The overarching goal of the study is relatively broad, and this breadth may have drawbacks; for example, limited ability to probe specific relations in depth. However, a holistic approach may provide unique insights into the overall structure of moral functioning as well as the mechanisms supporting this structure; these insights may be less attainable when focusing solely on a few aspects of moral functioning. To reduce the cost of this broad approach, our data analysis plan was designed to maintain specificity (e.g., using bivariate correlations to explore local relations of theoretical interest), while at the same time increasing power (e.g., using exploratory factor analysis to reduce the number of variables).

Method

Participants

Participants were 171 typically developing children (85 females; $M_{age} = 63.10$ months, range = 47.30–81.10 months) from middle-class families

living in a metropolitan area in the Pacific Northwest of North America. These children were predominantly Caucasian and East Asian. One additional child was excluded due to color blindness and inability to distinguish color-labeled stimuli. This sample size provides a power ($1 - \beta$) of .98 to detect moderate correlations ($\rho s \geq .3$) between different aspects of moral functioning in preschool.

Procedure

Participants, randomly assigned to one of six test orders, completed both moral and nonmoral tasks (described below) during a single visit. These tasks involved observations of children's responses to the examiner, sharing of resources, responses to verbally presented hypothetical scenarios, and responses to a puppet show. Children were administered the tasks in private, individual interviews with an examiner. The procedure as well as a general analysis plan for the study were preregistered on Open Science Framework (https://osf.io/fhzsq/). One deviation from our preregistered analysis plan is justified below; a second was made prior to data analysis and is specified in Supporting Information.

Measures

Instrumental Helping (Moral Construct)

Using a paradigm adapted from Bryan et al. (2014), children were invited to join the experimenter in playing with Legos. The experimenter ensured that the child was fully engaged before providing four opportunities to help. Each opportunity lasted 20 s, and help was verbally prompted but not explicitly requested. First, the experimenter noticed some wooden blocks on the floor, which she picked up and put in a bucket. Because the bucket was too small, some blocks kept falling out. The child could help by picking up at least one block from the floor and putting it into the bucket. After picking up all the blocks, the experimenter carried the overflowing bucket to a storage bin and, with both hands occupied, pretended to have difficulty opening the lid of the bin. The child could help by opening the lid for the experimenter. Next, the experimenter returned to the table and invited the child to join her in coloring. Once the child was fully engaged, the experimenter started putting away the Lego pieces that were still on the table from the previous Lego activity. The child could help by putting away at least one piece of Lego. Finally, the experimenter "accidentally" knocked over a cup of markers and started picking up markers from the floor. In this situation, the child could help by picking up at least one marker and putting it back into the cup. Children received 1 point for providing help in each situation, resulting in helping scores from 0 to 4. Reaction times were also coded; analyses can be found in Supporting Information.

Comforting Behavior (Moral Construct)

Children's comforting behaviors were measured using the "banged knee scenario" (Dunfield et al., 2011). While sitting across the table from the child, the experimenter stood up and said, "Oh! I just need to grab something." As she stood up, she "accidentally" banged her knee against the edge of the table and sat down hard on the floor with a pained expression. She rubbed her knee and said, "Oh! My knee, I banged my knee!", looking at her knee for 5 s and alternating her gaze between the child and her knee for the next 5 s. Ten seconds later, the experimenter looked at the floor and asked, "What should we do?" pausing for another 5 s. The coding scheme was adapted from Dunfield and Kuhlmeier (2013). Children received two points for providing verbal comforting (e.g., "Are you okay?") and two points for providing physical comforting (e.g., hugging the experimenter). Showing concerned expressions was not coded as comforting behaviors. Scores on this task fall into three categories (0, 2, 4); higher scores indicate more comforting behaviors. Reaction times were also coded; analyses can be found in Supporting Information.

Sharing Behavior (Moral Construct)

Children's first-party sharing behaviors and third-party distribution were measured using the Dictator Game (Gummerum et al., 2010). In the sharing task, children distributed six stickers between themselves and an anonymous child who would come in later. The experimenter placed two envelopes and six stickers in front of the child. The child's name was written on one of the envelopes. The experimenter told children to put the stickers they wanted to keep for themselves in the envelope with their name on it, and put the stickers they wanted to give to the other child in the other envelope. Children's understanding of the rules was verbally confirmed. The dependent variable of this task was the number of stickers children distributed to the anonymous child. The range of possible

scores was 0–6; higher scores indicate more generous sharing.

The third-party distribution task was similar to the sharing task except that the stickers were divided between two identical stuffed animal dolls (third-party recipients). The experimenter placed two dolls, two envelopes, and six stickers in front of the child. The dolls were positioned 1.5 feet apart facing the child, and each doll was given an envelope. Children were asked to divide the stickers between the two dolls by placing the stickers in the envelopes. Our original preregistration specified that we would code the number of stickers children gave to the less favored animal doll (range = 0–3); however, because the distribution was highly skewed (76% children distributed stickers evenly), we chose to dichotomize this variable. Hence, the dependent variable of this task was whether children distributed the stickers evenly or not. Unequal distributions received a score of 0, and equal distributions received a score of 1.

Evaluations of Moral Scenarios (Moral Construct)

Children's evaluations of verbally presented moral scenarios were measured using the prototypic moral transgression test and the morally relevant theory-of-mind test (MoToM), both adapted from Killen et al. (2011). Children's evaluations of nonverbal moral scenarios were measured using a puppet show task adapted from Van de Vondervoort and Hamlin (2018). Although for completeness, we administered the entirety of each task, because our goal is to explore the relations between moral behaviors and moral evaluations we only focus on questions pertaining to moral evaluations in the main text. Analyses for other questions are provided in Supporting Information.

For verbally presented moral scenarios, picture cards were used to facilitate comprehension, and characters' genders were matched to the gender of the child. Children were first familiarized with a 4-point Likert scale represented by four emotional faces. The experimenter explained the meanings of the faces (from unhappy to happy: 4 = *really bad*; 3 = *a little bad*; 2 = *a little good*; 1 = *really good*), and confirmed children's understanding of the scale using two questions ("Can you show me how much you like pizza?" "Can you show me how much you like Brussels sprouts?").

In the prototypic moral transgression test, the experimenter told children the following story: "This is David/Diane. This is Martin/Mary. David/Diane is playing on the swings outside. Martin/Mary comes over and pushes him/her off the swing so that he/she can get on it. David/Diane falls down on the ground and hurts his/her knee." Children then responded to four questions about the transgressor: (a) *intentions of the transgressor* (e.g., "When Martin pushed David, did Martin think he was doing something that was all right or not all right?"; Likert scale 1–4); (b) *justifications for the* intentions *of the transgressor* ("Why?" leading to a verbal response); (c) *moral evaluation* (e.g., "When Martin pushed, do you think he was doing something that was all right or not all right?"; Likert scale 1–4); and (d) *justification for moral evaluation* ("Why?" leading to a verbal response). The next two questions were about the victim: (e) *attributions of the emotional state of the victim* (e.g., "How will David feel about getting pushed?" leading to a verbal response); and (f) *attributions of the victim emotion toward the transgressor* (e.g., "How will David feel about Martin?"; Likert scale 1–4). For the purposes of the current study, we focused on the *moral evaluation* question. Higher scores on this measure represent harsher evaluations of the prototypic transgressor.

In the MoToM test, children were presented with the following scenario:

> This is Tommy/Tammy and this is Josh/Jane. Tommy/Tammy has brought in a cupcake from home and is keeping it in a paper bag. Tommy/Tammy puts the paper bag on the table then goes outside to play. Josh/Jane is helping the teacher clean up the classroom and sees the paper bag. Josh/Jane throws the paper bag in the trash.

Children first responded to five questions about the accidental transgressor: (a) *false belief of the accidental transgressor* (e.g., "What did Josh, the boy who threw out the paper bag, think was in the bag?" leading to a verbal response); (b) *intentions of the accidental transgressor* (e.g., "When Josh threw out the bag, did he think he was doing something that was all right or not all right?"; Likert scale 1–4); (c) *justification for intentions of the accidental transgressor* ("Why?" leading to a verbal response); (d) *moral evaluation* (e.g., "When Josh threw out the bag, do you think he was doing something that was all right or not all right?"; Likert scale 1–4); and (e) *justifications for moral evaluation* ("Why?" leading to a verbal response). Three additional questions were about the victim: (f) *false belief of the victim* (e.g., "Now Tommy wants to eat the cupcake that he brought in from home . . . Where will

Tommy look for his cupcake?" verbal response—table or trash); (g) *attributions of the emotional state of the victim* (e.g., "How will Tommy feel about losing his cupcake?" verbal response); and (h) *attributions of the victim emotion toward the accidental transgressor* (e.g., "How will Tommy feel about Josh?"; Likert scale 1–4). For the purposes of the current study, we focused on the *moral evaluation* question. Higher scores on this measure represent harsher evaluations of the accidental transgressor.

For nonverbal moral scenarios, children viewed puppet shows depicting unsuccessful attempts to help or hinder another's goal. At the beginning of the show, a protagonist puppet originally hidden underneath the table emerged and said "hi" to the participant. The protagonist then moved toward a clear box (with an attractive toy inside) and looked inside the box twice (twisting its body and saying, "Look, a toy!"). He then jumped onto the front corner of the box lid and tried unsuccessfully to lift the lid ("Too heavy!"). At this point, one of the two additional characters (racoon puppets resting on the opposite sides of the table) intervened. In failed helper events, the character (failed helper) ran forward, gasped the box lid, and tried to help the protagonist lift the lid ("Open!") but failed ("No toy!"). In the failed hinderer events, the character (failed hinderer) ran forward and attempted to hinder the protagonist by jumping on the box and slamming the box shut ("Close!"), but later the protagonist successfully opened the box ("Toy!").

After watching the puppet shows, children responded to five questions about their evaluations of the failed helper and the failed hinderer: (a) *social preference* ("Which one of these guys do you like the most?"); (b) *moral evaluation* ("Which one of these guys was nicer?"); (c) *degree of moral evaluation* ("Was he a lot nice, or a little bit nice?"); (d) *punishment assignment* ("I think one of these guys should get in trouble. Who should get in trouble?"); (e) *justifications for punishment* ("Why should he get in trouble?"). Four additional questions were asked to assess children's understanding of the puppet shows: (f) *failed helper intention* (the experimenter pointed at the failed helper and asked: "Did he try to open the box or close the box?"); (g) *failed helper outcome* ("Did the penguin get the toy?"); (h) *failed hinderer intention* (the experimenter pointed at the failed hinderer and asked: "Did he try to open the box or close the box?"); (i) *failed hinderer outcome* ("Did the penguin get the toy?"). The order of the questions was counterbalanced across children. We reasoned that if children engaged in intention-based moral judgments, they should view the failed helper as nicer (*moral evaluation*) and assign punishment to the failed hinderer (*punishment assignment*). Thus, for the purpose of the current study, we focused on these two questions. Specifically, we computed a composite score for these two questions (*moral evaluation* and *punishment assignment*) and labeled the measure as *evaluation of nonverbal moral scenarios*. Scores on this measure could range from 0 to 2; higher scores indicate more intention-based evaluations.

Cognitive Skill (Nonmoral Construct)

Children's global cognitive functioning was estimated using a two-subtest short form of the Wechsler Preschool and Primary Scale of Intelligence, 4th ed. (WPPSI–IV; Wechsler, 2012a). The WPPSI–IV is a widely used intelligence test for children aged 2.5–7.5 years of age, with strong psychometric properties. Children were administered the Matrix Reasoning subtest and the Information subtest, and the average of the standard scores on these two subtests was calculated. These two subtests have high reliability coefficients (Information = .89, Matrix Reasoning = .90), and are the two subtests of the WPPSI with the highest individual correlations with Full Scale IQ (Information = .64, matrix reasoning = .60; Wechsler, 2012b).

Executive Functioning (Nonmoral Construct)

Children's executive functioning was assessed using the delay of gratification task (Mischel, Ebbesen, & Raskoff Zeiss, 1972; Rodriguez, Mischel, & Shoda, 1989) and the Head Toes Knees Shoulders task (HTKS; Ponitz et al., 2008). The delay of gratification task measures children's ability to inhibit material desires. At the onset of the test, the experimenter confirmed with children that they preferred having two (vs. one) chocolate-covered marshmallows. For children who had allergies or disliked marshmallows, other snack options (e.g., yogurt drops, graham crackers) were offered. The experimenter then put one marshmallow in front of children and said,

> Okay! Well, I need to do something in the other room. If you wait until I come back by myself, without eating this marshmallow, then you can have two marshmallows to eat instead. But if you don't want to wait, you can knock on the door and make me come back anytime you want to. But if you knock on the door and make me come back, you can't have two marshmallows.

Children's understanding of the instructions was confirmed via questioning (e.g., "What do you need to do if you want to have two marshmallows?", "What will happen if you knock on the door?"). Children were then left in the room with one marshmallow. If children ate the marshmallow or knocked on the door, the experimenter returned and allowed them to finish the marshmallow. If children waited 15 min without eating the marshmallow or knocking on the door, the experimenter returned and offered them two marshmallows. The dependent variable of this task was children's waiting times (range = 0–15 min). Longer waiting times indicate better inhibitory control (Duckworth, Tsukayama, & Kirby, 2013).

The HTKS task (Ponitz et al., 2008) measures children's inhibitory control, working memory, and attentional control. Children were invited to play a game in which they must act in the opposite way of the experimenter's commands. When the experimenter said, "Touch your head (or toes)!" children were supposed to touch their toes (or head). If children passed the head or toes phase of the game (by correctly responding to five or more commands of the 10 commands or if children were in kindergarten or beyond), they proceeded to the second phase where two additional body parts (knees and shoulders) were added. In the second phase, children were supposed to touch their toes (head) when the experimenter said, "Touch your head (toes)!" and touch their knees (shoulders) when the experimenter said, "Touch your or shoulders (knees)!" For each command, children received a score of 2 if they produced the correct response right away, a score of 1 if they self-corrected without prompting, and a score of 0 if they failed to touch the correct body part. In total, there were 26 commands to respond to; possible scores on this task could range from 0 to 52.

Theory-of-Mind (Nonmoral Construct)

Using the full version of the Wellman battery (Wellman & Liu, 2004), children viewed seven scenarios enacted by toy figurines and picture props and answered questions about characters' mental states. The questions tap children's understanding of diverse desires (knowing that people can have different desires), diverse beliefs (knowing that people can have different beliefs about the same situation), knowledge access (knowing that other people may not know what children themselves know), contents false belief (knowing that other people can have false beliefs about what is in a container), explicit false belief (predicting how a person will behave given his/her false belief), belief emotion (predicting how a person will feel given his/her false belief), and real-apparent emotion (knowing that people can display emotions that are different from how they actually feel). Details for each scenario are described in the original article (Wellman & Liu, 2004). Possible scores on this task range from 0 (passing none) to 7 (passing all); higher scores indicate better theory-of-mind skills.

Emotion Recognition (Nonmoral Construct)

We assessed children's ability to identify dynamic faces depicting five basic emotions (happiness, sadness, anger, fear, pain). The stimuli were taken from the STOIC database (Roy et al., 2007), and the procedure was adapted from the emotional recognition paradigm described by Kimonis et al. (2016). Before the test, children were asked to recall times in which they experienced the five basic emotions. The experimenter then invited children to play a game on the laptop. The task consisted of five practice trials (5 expressions × 1 actor) and 40 test trials (5 expressions × 8 actors). Feedback was provided for the practice trials but not for the test trials. The experimenter intervened when children provided task-irrelevant answers or emotions outside of the five options; the experimenter did not tell children whether their answers were right or wrong. Each trial began with a blank slide, followed by a fixation cross (500 ms) and a brief grayscale movie clip (1,000 ms) depicting an emotional expression. Children were asked to identify the emotion out loud. The black slides were manually advanced by the experimenter to make sure that children were looking at the screen during the stimulus presentation. The order of the movies was counterbalanced across children. Children received one point for each correct response. Possible scores on this task could range from 0 to 40; higher scores indicate better emotional recognition abilities.

Shyness (Nonmoral Construct)

A coder watched study videos and rated children's signs of shyness during the first 2 min of the whole study. Based on Mish (1995), shyness was defined as being easily frightened, disposed to avoid persons or things, hesitant to commit themselves, sensitively diffident or retiring, secluded, or hidden. The coder scored shyness along four dimensions: verbal responses, facial expressions, body language, and reactivity to experimenter's

prompts. For each dimension, shyness was rated on a 5-point scale (1 = *not shy at all*; 3 = *average shyness*; 5 = *very shy*). A shyness score was then calculated for each child by averaging scores across these four dimensions. A second independent coder scored a subset (n = 51, 30%) of the whole sample. The intraclass correlation coefficient between the two raters was .86 (two-way random, consistency, average measure; Shrout & Fleiss, 1979).

Empathic Concern (Nonmoral Construct)

We coded children's empathic concern for distressed others during the comforting task ("banged knee scenario"). An independent coder watched study videos and rated children's facial and behavioral cues of empathic concern toward the experimenter. The coding scheme was based on Ball et al. (2017) and focused on children's facial expressions (e.g., eyebrows pulled down) and body movements (e.g., forward head orientation). Scores ranged from 0 (no facial or nonverbal indicators of concern) to 4 (exceptionally strong, sustained display of concern in face and body); higher scores indicate higher empathic concern.

Data Analytic Plan

We first tested whether different forms of moral behaviors and moral evaluations formed a unitary construct, by exploring the relations between our moral task measures using bivariate correlations and factor analysis. If factor analysis supported data reduction to a smaller number of measures, we retained these major dimensions in subsequent analyses, in order to reduce the number of total analyses conducted.

Next, we explored how various cognitive and emotional mechanisms relate to different aspects of moral functioning, by correlating nonmoral measures with major dimensions obtained in the factor analysis. Finally, to explore the strongest predictors of the major dimensions of moral functioning, we performed regression analyses predicting these dimensions from nonmoral variables while controlling for age and gender. Besides these main analyses, we also performed supplementary analyses for each moral task to provide further insights into moral functioning in preschool (see Supporting Information).

Results

Consistency Between Aspects of Moral Functioning

Bivariate correlations between different aspects of moral functioning are shown in Table 1. For moral behaviors, we found positive correlations between comforting, instrumental helping, and sharing: Children who showed more comforting behaviors were more helpful to the experimenter and more generous in sharing resources with others. Unlike with first-party distributions (sharing resources between oneself and an anonymous child), no correlations were observed between third-party distribution (dividing resources among two animal dolls) and prosocial behaviors of instrumental helping, comforting, or sharing. These results provide evidence for associations between different forms of prosocial behaviors and dissociation between prosocial behaviors and third-party distributions.

We also found significant correlations between different types of moral evaluations: Children who evaluated prototypic moral transgression more negatively also evaluated accidental transgression more negatively (the direction of this effect will be discussed below), and showed more intention-based judgments in the nonverbal moral evaluation task. These results provide evidence for links between

Table 1
Descriptive Statistics and Bivariate Correlations Between Subtypes of Moral Functioning

	1	2	3	4	5	6	7	*M* (*SD*)
1. Instrumental helping	—	.21**	.11	.03	.05	.15*	−.04	2.91 (1.46)
2. Comforting		—	.15*	−.12	.08	.11	−.10	1.11 (1.15)
3. Sharing			—	.04	.08	.09	.02	1.45 (1.52)
4. Third-party distribution				—	.13	−.08	.11	0.76 (0.43)
5. Evaluation of prototypic moral transgression					—	.20*	.17*	2.61 (0.84)
6. Evaluation of accidental transgression						—	−.02	2.40 (1.03)
7. Evaluation of nonverbal moral scenarios							—	1.57 (0.75)

*p < .05. **p < .01 (two-tailed).

different forms of moral evaluations. Finally, there were no significant correlations between moral behaviors and moral evaluations, except that children who engaged in more instrumental helping tended to evaluate accidental transgression more negatively. Taken together, these findings provide preliminary evidence for homogeneity within prosocial behaviors (instrumental helping, comforting, sharing), homogeneity within moral evaluations, and heterogeneity between prosocial behaviors and moral evaluations.

We next conducted a principal-axis factor analysis on the moral measures. This analysis served two purposes: First, it provided further information about the relations between different forms of moral behaviors and moral evaluations. Second, it allowed us to extract major dimensions of moral functioning, in order to reduce the number of variables entered into subsequent analyses (where we explored the relations between moral measures and nonmoral measures). Our data contained sufficient shared variance for factor analysis, Kaiser-Meyer-Olkin value = .55, Bartlett's test of sphericity, $\chi^2(21) = 42.38$, $p = .004$. Because varimax (orthogonal) and oblimin (oblique) rotations yielded similar solutions, we only report the varimax solution. Based on Kaiser's stopping rule (Ford, MacCallum, & Tait, 1986; Kaiser, 1974; Zwick & Velicer, 1986), we retained two factors with eigenvalues > 1 (third factor eigenvalue = 0.99 and accounted for 14.15% of the variance). Model comparisons and inspection of the scree plot, factor matrix, and reproduced correlations confirmed that this two-factor solution offered the best summary of the data.

Results are shown in Table 2 (to facilitate interpretation, loadings below 0.250 were omitted). Variables loading on the first factor included prosocial behaviors (comforting, instrumental helping, and sharing) and moral evaluation of the accidental transgression. The constituent variables were all positively related to the latent factor, and this latent factor accounted for 21.44% of total variance. Variables loading on the second factor included children's moral evaluations of prototypic moral transgression and nonverbal moral scenarios, as well as third-party distributions. The constituent variables were all positively related to the latent factor, and this latent factor accounted for 18.92% of total variance. Overall, this two-factor solution accounted for 40.36% of the variance in children's performance on moral behavior and evaluation tasks.

The two-factor solution presented above should be interpreted in light of the overall low factor loadings. This might be attributable to the fact that different aspects of moral functioning were assessed using very different paradigms and measurement scales. Despite this drawback, the general pattern in the factor matrix provides further insights into the structure of moral functioning. We reasoned that most of the variables on the first factor (instrumental helping, comforting, sharing, but not evaluation of accidental transgression; see discussion below) represent first-party prosocial behaviors. These behaviors require higher personal involvement (e.g., more time and energy investment, more social interaction with the recipients) and directly promote recipients' welfare. The variables on the second factor represent third-party moral functioning, whereby children act as independent judges or arbitrators, evaluating others' behaviors and allocating resources between others. These behaviors require less personal involvement (e.g., children do not need to interact with the recipients) and do not directly promote recipients' welfare. Together, these results look similar to those observed in our correlation matrix, and provide evidence for homogeneity within prosocial behaviors, homogeneity within third-party moral functioning, and heterogeneity between prosocial behaviors and third-party moral functioning.

Table 2
Principal-Axis Factor Analysis of Moral Behaviors and Moral Evaluations

Variable	Loadings 1	Loadings 2	Eigenvalue	Variance explained (%)
Factor 1			1.50	21.44
Comforting	.492			
Instrumental helping	.366			
Evaluation of accidental transgression	.357			
Sharing	.286			18.92
Factor 2			1.32	
Evaluation of prototypic moral transgression	.335	.532		
Evaluation of nonverbal moral scenarios		.397		
Third-party distribution		.279		

Cognitive and Emotional Factors Related to Moral Functioning

We next explored how cognitive and emotional factors relate to moral functioning. To reduce the number of moral functioning variables, based on our factor analysis, we calculated composite scores for prosocial behaviors (comforting, helping, and

sharing) and third-party moral functioning (evaluation of prototypic moral transgression, evaluation of nonverbal moral scenarios, and third-party distribution). Composite scores were calculated by adding unweighted standardized scores of constituent variables (DiStefano, Zhu, & Mindrila, 2009; Gorsuch, 1983). Higher scores on prosocial behaviors and third-party moral functioning represent more prosocial behaviors and higher levels of third-party moral functioning, respectively. The obtained composite scores were then correlated with cognitive and emotional factors.

Although it emerged as part of the first factor in our analysis, we chose to exclude evaluation of the accidental transgression from the first factor (prosocial behaviors) for three reasons: First, theoretically, evaluation of accidental transgressions is not a form of prosocial behavior. Second, correlational analyses showed that this variable was correlated with both instrumental helping and evaluation of the prototypic moral transgression. Thus, it is unclear which conceptual group this variable should actually be categorized into. Third, despite the usefulness of this paradigm in assessing theory-of-mind in moral contexts, there is no clear evidence for the reliability and validity of this *moral evaluation* question with preschool populations: In past studies, children at this age have failed to resist negative evaluations of accidental transgressions despite knowing the accidental nature of the act (Killen et al., 2011). Indeed, in the current study, children who evaluated the prototypic moral transgression more negatively also evaluated the accidental transgression more negatively, despite their distinct moral implications. Thus, it is unclear whether this measure taps meaningful individual differences in preschoolers. Confirmatory factor analyses revealed that, after removing moral evaluation of the accidental transgression, the two-factor model showed good model fit and significant factor loadings (see Supporting Information).

As shown in Table 3, prosocial behaviors and third-party moral functioning were both associated with age: Older children performed more prosocial behaviors and showed higher levels of third-party moral functioning (but age did not moderate the relations between different aspects of moral functioning; see Supporting Information). Girls also performed more prosocial behaviors than boys. Crucially, prosocial behavior was not correlated with third-party moral functioning (see Supporting Information for further evidence that the two factors can be viewed as orthogonal), and these two variables were predicted by different cognitive and emotional factors. Specifically, prosocial behaviors were associated with both shyness, mental state understanding, and empathic concern: Children who performed more prosocial behaviors tended to be less shy, better at understanding others' mental states (scoring higher on both the Wellman battery and the emotional recognition task), and more empathic to others' distress. Third-party moral functioning was correlated with delay of gratification: Children who scored higher on third-party moral functioning were better at inhibiting material desires in the Marshmallow Task. Together, these results provide further evidence for the dissociation between first-party prosocial behaviors and third-

Table 3
Bivariate Correlations Between Moral Functioning, Cognitive and Emotional Factors, Age, and Gender

	1	2	3	4	5	6	7	8	9	10	11
1. Prosocial behaviors	—										
2. Third-party moral functioning	.01	—									
3. WPPSI mean	.09	−.03	—								
4. Shyness score	−.19*	.04	−.08	—							
5. Delay of gratification (waiting time)	.03	.20*	.17*	.14	—						
6. Self-regulation and executive function (HTKS total)	.12	.13	.43**	.10	.33**	—					
7. Theory-of-mind (Wellman battery)	.23**	.08	.30**	.03	.31**	.54**	—				
8. Emotional recognition	.26**	.07	.12	−.09	.15	.43**	.29**	—			
9. Empathic concern	.35**	−.07	−.01	−.20*	−.05	−.06	−.10	−.01	—		
10. Age	.25**	.22**	.01	.02	.29**	.52**	.44**	.51**	−.03	—	
11. Gender	.15*	.03	.11	.10	.09	.11	.16*	.06	.14	.06	—

Note. Gender: 1 = female and 0 = male. WPPSI = Wechsler Preschool and Primary Scale of Intelligence; HTKS = Head Toes Knees Shoulders.
*$p < .05$. **$p < .01$ (two-tailed).

party moral functioning, and suggest that these two aspects of moral functioning are served by distinct cognitive and emotional mechanisms.

Comparing Predictors of Moral Functioning

Finally, to explore the strongest predictors of moral functioning above and beyond age and gender, we performed hierarchical regressions with the composite scores of prosocial behaviors and third-party moral functioning as dependent variables. Age and gender were entered in the first step as statistical controls. Cognitive and emotional factors were entered in the second step. Preliminary analyses showed no multicollinearity issues (VIFs < 2.20, tolerance values > 0.45) or other violations of statistical assumptions. The results are shown in Table 4. In the first step, age was a significant predictor of both prosocial behaviors and third-party moral functioning; older children showed more of these constructs than did younger children. In the second step, after controlling for age and gender, the strongest predictor of prosocial behaviors was empathic concern, and the strongest predictor of third-party moral functioning was delay of gratification. These results demonstrate that the relations between empathic concern and prosocial behaviors and the relations between inhibitory control and third-party moral functioning were not reducible to age- and gender-related differences. Taken together, these findings support the idea that first-party prosocial behaviors and third-party moral functioning are associated with different mechanisms, and suggest that empathy and executive functioning (inhibitory control) are important predictors of moral development during the preschool years.

Table 4
Hierarchical Regressions for Prosocial Behaviors and Third-party Moral Functioning

	Prosocial behaviors		Third-party moral functioning	
Predictor	β (Step 1)	β (Step 2)	β (Step 1)	β (Step 2)
Age	.21*	.10	.26**	.23
Gender	.15	.11	−.08	−.05
IQ		.05		−.16
Shyness		−.15		−.04
Delay of gratification		.03		.24*
Executive functioning (HTKS)		−.07		.07
Theory-of-mind		.17		−.05
Emotional recognition		.10		−.10
Empathic concern		.35**		−.06

Note. Gender: 1 = female and 0 = male. β = standardized regression coefficient. HTKS = Head Toes Knees Shoulders.
*$p < .05$. **$p < .01$.

Discussion

The current study systematically examined the relations between different aspects of moral functioning, and their cognitive and emotional correlates, in preschool-aged children. We found that different forms of prosocial behaviors were related to one another, providing evidence for homogeneity within prosocial behaviors in preschool. Specifically, preschoolers who showed more comforting behaviors were more likely to provide instrumental helping and shared more resources with others. We note that the correlations were stronger between helping and comforting and between comforting and sharing, but weaker (and statistically insignificant) between helping and sharing. This pattern is consistent with past studies with preschoolers showing correlations between helping and comforting (Eisenberg & Lennon, 1980) and between comforting and sharing (Yarrow et al., 1976), but not between helping and sharing (Cassidy et al., 2003; Eisenberg et al., 1984; Yarrow et al., 1976). We note that the current study assessed children's prosocial behaviors in a specific situation. Future studies should incorporate trait measures (e.g., parental reports) to explore whether associations between prosocial behaviors are stronger with more reliable measures.

Interestingly, past studies have generally found dissociations between distinct prosocial behaviors in infants and younger children (e.g., Dunfield & Kuhlmeier, 2013; Dunfield et al., 2011; Paulus et al., 2013; Pettygrove, Hammond, Karahuta, Waugh, & Brownell, 2013; but see Gross et al., 2015; Newton et al., 2016). The heterogeneity found in infants and younger children and the homogeneity found in the current and other preschool studies (Eisenberg & Lennon, 1980; Yarrow et al., 1976) suggest that there might be an increase in cohesion between different forms of prosocial behaviors from infancy to preschool (see also Knafo-Noam, Uzefovsky, Israel, Davidov, & Zahn-Waxler, 2015 for evidence of consistent prosocial behaviors in 7-year-old children). These early behavioral tendencies may later consolidate into a consistent prosocial personality (Eisenberg et al., 1999).

What explains the correlations between different types of prosocial behaviors? One possibility is that

these correlations reflect similarities in underlying cognitive and emotional mechanisms. Nonetheless, we reason that this reductionist account is implausible: Scholars have argued that helping, sharing, and comforting are based on different cognitive and motivational mechanisms (Dunfield, 2014; Paulus, 2014, 2018), and recruit different neurocognitive processes (Paulus et al., 2013; Steinbeis, 2018). In the current study, although instrumental helping and sharing were associated with mental state understanding, and comforting and sharing were associated with empathic concern, there was not a single mechanism that correlated with all three types of prosocial behaviors (see Table S1). Thus, the associations between different forms of prosocial behaviors should be attributed to processes that were not measured in the current study (see below for a discussion of other candidate processes).

Besides the associations between prosocial behaviors, we also found associations between moral evaluations of verbally presented prototypic moral transgression and nonverbally presented moral scenarios. We note that children's evaluations of verbally presented prototypic moral transgression and nonverbally presented moral scenarios were measured using very different stimuli (verbal vignettes vs. nonverbal puppet shows), scenarios (prototypic transgression vs. failed attempts to help or harm), and measuring scales (four-point Likert scales vs. dichotomous questions). The fact that we found associations despite these dissimilarities speaks to the strength of the cohesion between different forms of third-party moral evaluations. Interestingly, although evaluations of verbally presented prototypic moral transgression and nonverbal moral scenarios were correlated, they were not predicted by the same cognitive and emotional factors: Whereas evaluations of verbally presented prototypic moral transgression were predicted by emotional recognition, evaluations of nonverbal moral scenarios were predicted by delay of gratification (see Table S1). Hence, the data again suggest that the associations observed between distinct moral evaluations may not be attributable to shared cognitive or emotional mechanisms.

The idea that the cohesion within first-party prosocial behaviors and within third-party moral evaluations seems not to be attributable to specific bottom-up mechanisms is consistent with recent neuroscience evidence that there is no unique neural signature that accounts for all aspects of moral cognition. Indeed, our findings are in line with others suggesting that moral evaluation is supported by a wide range of domain-general mechanisms (Young & Dungan, 2012). However, other mechanisms underlying the cohesion within both first-party prosocial behaviors and third-party moral evaluations might exist. For example, there may be situational characteristics central to moral functioning that we did not measure with our paradigms (e.g., how comfortable the child was with the experimenter, or whether a reciprocal norm was established in the situation; see Barragan & Dweck, 2014). Alternatively or in addition, socialization processes, which presumably could serve to promote different forms of prosocial behaviors relatively equally, could generate coherent individual differences. Given that the impact of socialization on prosocial behaviors likely takes time to unfold, one could examine this hypothesis by determining if the cohesion between different forms of prosocial behaviors increases from infancy to preschool.

In contrast to the homogeneity found within first-party prosocial behaviors and third-party moral evaluations, we found heterogeneity between first-party prosocial behaviors and third-party moral functioning. Specifically, prosocial behaviors were associated with neither third-party resource distribution nor moral evaluations (except for *evaluation of accidental transgression*; see prior discussion). With respect to moral behaviors, the dissociation between prosocial behaviors and third-party distribution is consistent with the idea that prosocial behaviors and third-party distribution rely on distinct psychological mechanisms. Whereas prosocial behaviors require personal involvement and direct social interactions, third-party distribution only requires children to make decisions "at a distance." Indeed, our data showed that mental state understanding was associated with helping and sharing, and empathic concern was associated with comforting and sharing, but neither mental state understanding nor empathic concern was associated with third-party distribution (see Table S1). These findings provide support for the idea that third-party distribution may rely less on socio-cognitive processes than do prosocial behaviors. We also note that the recipients of the third-party distribution task were two identical animal dolls. Hence, children's decision to distribute equally may reflect a tendency to create equal divisions, without need-based considerations.

The dissociation between moral evaluations and prosocial behaviors stands in contrast with previous studies showing correlations between children's reasoning about moral dilemmas (justifications for moral choices) and their prosocial behaviors (for review, see Villegas de Posada & Vargas-Trujillo,

2015). Considering that our moral evaluation tasks all involved an antisocial scenario, one explanation for the dissociation is that prosocial behaviors might be only linked to evaluations of prosocial (but not antisocial) scenarios. Indeed, recent work shows that children's prosocial behaviors are specifically predicted by their prosocial values (Abramson, Daniel, & Knafo-Noam, 2018; Benish-Weisman, Daniel, Sneddon, & Lee, 2019). Alternatively, the fact that prosocial behaviors were predicted by reasoning about moral dilemmas but not by evaluations of others' moral behaviors suggests that prosocial behaviors may be more related to the deliberative (as opposed to intuitive) aspects of moral cognition. Consistent with this idea, a recent study found that 3- to 5-year-old children's sharing behaviors were predicted by later controlled (as opposed to early automatic) neural responses to third-party moral interactions (Cowell & Decety, 2015), suggesting that the controlled (vs. automatic) aspects of moral cognition may have more bearing on children's own prosocial behaviors. Unfortunately, the current study did not include a moral dilemma task (e.g., Eisenberg & Hand, 1979), which prevents us from making a direct comparison between these two aspects of moral cognition. Future research should include both third-party moral evaluations of prosocial events and moral dilemmas to explore whether prosocial behaviors are specifically associated with deliberative reasoning about moral choices or with evaluations of prosocial events.

What are the strongest predictors of prosocial behaviors? Correlational analyses revealed that prosocial behavior, as a construct, was correlated with age, gender, shyness, theory-of-mind, emotional recognition, and empathic concern. Specifically, children who performed more prosocial behaviors were less shy, better at understanding others' mental states, better at recognizing facial expressions (see also Table S2 for correlation between prosocial behaviors and real-apparent emotion understanding), and more empathic. These findings are consistent with the idea that prosocial behaviors require the abilities to understand others' mental states (Imuta et al., 2016), recognize overt expressions of negative states (Grossmann, 2018; Rajhans, Altvater-Mackensen, Vaish, & Grossmann, 2016), overcome social inhibition (Beier et al., 2017; Karasewich et al., 2018), and show sympathy to others (Holmgren, Eisenberg, & Fabes, 1998). After controlling for age and gender, empathic concern was the strongest predictor of prosocial behaviors above and beyond other cognitive and emotional factors. This result is in line with past research suggesting that children's prosocial behaviors are driven by a genuine concern for others' plight (Hepach, 2017; Vaish, Carpenter, & Tomasello, 2009).

What are the strongest predictors of third-party moral functioning? Correlational analyses showed that third-party moral functioning, as a construct, was predicted by age and delay of gratification. The effect of delay of gratification still held after controlling for age, gender, and other cognitive and emotional factors. This finding is compatible with theories viewing inhibitory control as an important building block for higher level socio-cognitive processes (Barkley, 1997; Birch & Bloom, 2004). In the context of third-party moral functioning, inhibitory control allows children to resist the distraction of irrelevant stimuli, to suspend personal views, and to consider others' mental states which may be different from one's own. These processes are crucial for generating moral judgments and making fair distributions. Indeed, a recent study found that inhibitory control (but not working memory) was associated with moral reasoning in children with autism spectrum disorder (Kretschmer, Lampmann, & Altgassen, 2014). This finding highlights the unique role of inhibitory control in the development of moral cognition.

Taken together, results from the current study provide evidence for homogeneity within first-party prosocial behaviors and within third-party moral functioning, as well as heterogeneity between these two aspects of moral functioning in preschool. These findings highlight the nuanced relations between first-party and third-party moral behaviors, between moral behaviors and moral cognition, and between different forms of moral evaluations. Our findings also underline the importance of empathic concern and inhibitory control in the development of moral behaviors and cognition. Future research should explore the relations between different aspects of moral functioning and their correlates in other age groups. Studies along this line will help us depict a more nuanced and complete picture of moral development.

References

Abramson, L., Daniel, E., & Knafo-Noam, A. (2018). The role of personal values in children's costly sharing and non-costly giving. *Journal of Experimental Child Psychology*, *165*, 117–134. https://doi.org/10.1016/j.jecp.2017.03.007

Ball, C. L., Smetana, J. G., & Sturge-Apple, M. L. (2017). Following my head and my heart: Integrating preschoolers' empathy, theory of mind, and moral

judgments. *Child Development*, *88*, 597–611. https://doi.org/10.1111/cdev.12605

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*, 65–94. https://doi.org/10.1037/0033-2909.121.1.65

Barragan, R. C., & Dweck, C. S. (2014). Rethinking natural altruism: Simple reciprocal interactions trigger children's benevolence. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 17071–17074. https://doi.org/10.1073/pnas.1419408111

Beier, J. S., Terrizzi, B. F., Woodward, A. M., & Larson, E. G. (2017). Shyness and social conflict reduce young children's social helpfulness. *Child Development*, *88*, 1922–1929. https://doi.org/10.1111/cdev.12681

Benish-Weisman, M., Daniel, E., Sneddon, J., & Lee, J. (2019). The relations between values and prosocial behavior among children: The moderating role of age. *Personality and Individual Differences*, *141*, 241–247. https://doi.org/10.1016/j.paid.2019.01.019

Birch, S. A., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Sciences*, *8*, 255–260. https://doi.org/10.1016/j.tics.2004.04.011

Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin*, *88*, 1–45. https://doi.org/10.1037/0033-2909.88.1.1

Bryan, C. J., Master, A., & Walton, G. M. (2014). "Helping" versus "being a helper": Invoking the self to increase helping in young children. *Child Development*, *85*, 1836–1842. https://doi.org/10.1111/cdev.12244

Cassidy, K. W., Werner, R. S., Rourke, M., Zubernis, L. S., & Balaraman, G. (2003). The relationship between psychological understanding and positive social behaviors. *Social Development*, *12*, 198–221. https://doi.org/10.1111/1467-9507.00229

Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgment: Theoretical foundations and research validation*. Cambridge, MA: Cambridge University Press.

Cowell, J. M., & Decety, J. (2015). The neuroscience of implicit moral evaluation and its relation to generosity in early childhood. *Current Biology*, *25*, 93–97. https://doi.org/10.1016/j.cub.2014.11.002

Dahl, A., & Brownell, C. A. (2019). The social origins of human prosociality. *Current Directions in Psychological Science*, *28*, 274–279. https://doi.org/10.1177/0963721419830386

Derryberry, W. P., Wilson, T., Snyder, H., Norman, T., & Barger, B. (2005). Moral judgment developmental differences between gifted youth and college students. *Journal of Secondary Gifted Education*, *17*, 6–19. https://doi.org/10.4219/jsge-2005-392

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation*, *14*, 20. https://doi.org/10.7275/da8t-4g52

Duckworth, A. L., Tsukayama, E., & Kirby, T. A. (2013). Is it really self-control? Examining the predictive power of the delay of gratification task. *Personality and Social Psychology Bulletin*, *39*, 843–855. https://doi.org/10.1177/0146167213482589

Dunfield, K. A. (2014). A construct divided: Prosocial behavior as helping, sharing, and comforting subtypes. *Frontiers in Psychology*, *5*, 958. https://doi.org/10.3389/fpsyg.2014.00958

Dunfield, K. A., & Kuhlmeier, V. A. (2013). Classifying prosocial behavior: Children's responses to instrumental need, emotional distress, and material desire. *Child Development*, *84*, 1766–1776. https://doi.org/10.1111/cdev.12075

Dunfield, K. A., Kuhlmeier, V. A., O'Connell, L., & Kelley, E. (2011). Examining the diversity of prosocial behavior: Helping, sharing, and comforting in infancy. *Infancy*, *16*, 227–247. https://doi.org/10.1111/j.1532-7078.2010.00041.x

Eisenberg, N., Guthrie, I. K., Murphy, B. C., Shepard, S. A., Cumberland, A., & Carlo, G. (1999). Consistency and development of prosocial dispositions: A longitudinal study. *Child Development*, *70*, 1360–1372. https://doi.org/10.1111/1467-8624.00100

Eisenberg, N., & Hand, M. (1979). The relationship of preschoolers' reasoning about prosocial moral conflicts to prosocial behavior. *Child Development*, *50*, 356–363. https://doi.org/10.2307/1129410

Eisenberg, N., & Lennon, R. (1980). Altruism and the assessment of empathy in the preschool years. *Child Development*, *51*, 552–557. https://doi.org/10.2307/1129290

Eisenberg, N., Pasternack, J. F., Cameron, E., & Tryon, K. (1984). The relation of quantity and mode of prosocial behavior to moral cognitions and social style. *Child Development*, *55*, 1479–1485. https://doi.org/10.2307/1130017

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, *39*, 291–314. https://doi.org/10.1111/j.1744-6570.1986.tb00583.x

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Gross, R. L., Drummond, J., Satlof-Bedrick, E., Waugh, W. E., Svetlova, M., & Brownell, C. A. (2015). Individual differences in toddlers' social understanding and prosocial behavior: Disposition or socialization? *Frontiers in Psychology*, *6*, 600. https://doi.org/10.3389/fpsyg.2015.00600

Grossmann, T. (2018). How to build a helpful baby: a look at the roots of prosociality in infancy. *Current Opinion in Psychology*, *20*, 21–24. https://doi.org/10.1016/j.copsyc.2017.08.007

Gummerum, M., Hanoch, Y., Keller, M., Parsons, K., & Hummel, A. (2010). Preschoolers' allocations in the dictator game: The role of moral emotions. *Journal of*

Economic Psychology, *31*, 25–34. https://doi.org/10.1016/j.joep.2009.09.002

Hauser, M. (2006). *Moral minds: How nature designed our universal sense of right and wrong*. New York, NY: Ecco/HarperCollins.

Hepach, R. (2017). Prosocial arousal in children. *Child Development Perspectives*, *11*, 50–55. https://doi.org/10.1111/cdep.12209

Holmgren, R. A., Eisenberg, N., & Fabes, R. A. (1998). The relations of children's situational empathy related emotions to dispositional prosocial behaviour. *International Journal of Behavioral Development*, *22*, 169–193. https://doi.org/10.1080/016502598384568

Imuta, K., Henry, J. D., Slaughter, V., Selcuk, B., & Ruffman, T. (2016). Theory of mind and prosocial behavior in childhood: A meta-analytic review. *Developmental Psychology*, *52*, 1192–1205. https://doi.org/10.1037/dev0000140

Kaiser, H. F. (1974). Analysis of factorial simplicity. *Psychometrika*, *39*, 31–36. https://doi.org/10.1007/BF02291575

Karasewich, T., Kuhlmeier, V., Beier, J. S., & Dunfield, K. (2018). Getting help for others: an examination of indirect helping in young children. *Developmental Psychology*, *55*, 606–611. https://doi.org/10.1037/dev0000654

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

Killen, M., & Smetana, J. G. (2006). *Handbook of moral development*. Mahwah, NJ: Erlbaum.

Kimonis, E. R., Fanti, K. A., Anastassiou-Hadjicharalambous, X., Mertan, B., Goulter, N., & Katsimicha, E. (2016). Can callous-unemotional traits be reliably measured in preschoolers? *Journal of Abnormal Child Psychology*, *44*, 625–638. https://doi.org/10.1007/s10802-015-0075-y

Knafo-Noam, A., Uzefovsky, F., Israel, S., Davidov, M., & Zahn-Waxler, C. (2015). The prosocial personality and its facets: Genetic and environmental architecture of mother-reported behavior of 7-year-old twins. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00112

Knafo-Noam, A., Vertsberger, D., & Israel, S. (2018). Genetic and environmental contributions to children's prosocial behavior: Brief review and new evidence from a reanalysis of experimental twin data. *Current Opinion in Psychology*, *20*, 60–65. https://doi.org/10.1016/j.copsyc.2017.08.013

Kretschmer, A., Lampmann, S.-A., & Altgassen, M. (2014). Relations between moral reasoning, theory of mind and executive functions in children with autism spectrum disorders. *International Journal of Developmental Disabilities*, *60*, 174–183. https://doi.org/10.1179/2047387714Y.0000000045

Malti, T., Gasser, L., & Gutzwiller-Helfenfinger, E. (2010). Children's interpretive understanding, moral judgments, and emotion attributions: Relations to social behaviour. *British Journal of Developmental Psychology*, *28*, 275–292. https://doi.org/10.1348/026151009X403838

Miller, P. A., Eisenberg, N., Fabes, R. A., & Shell, R. (1996). Relations of moral reasoning and vicarious emotion to young children's prosocial behavior toward peers and adults. *Developmental Psychology*, *32*, 210–219. https://doi.org/10.1037/0012-1649.32.2.210

Mischel, W., Ebbesen, E. B., & Raskoff Zeiss, A. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology*, *21*, 204–218. https://doi.org/10.1037/h0032198

Mish, F. X. (1995). *Merriam-Webster's collegiate dictionary* (10th ed.). Springfield, MA: Merriam-Webster.

Newton, E. K., Thompson, R. A., & Goodman, M. (2016). Individual differences in toddlers' prosociality: Experiences in early relationships explain variability in prosocial behavior. *Child Development*, *87*, 1715–1726. https://doi.org/10.1111/cdev.12631

Paulus, M. (2014). The emergence of prosocial behavior: Why do infants and toddlers help, comfort, and share? *Child Development Perspectives*, *8*, 77–81. https://doi.org/10.1111/cdep.12066

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

Paulus, M. (2018). The multidimensional nature of early prosocial behavior: A motivational perspective. *Current Opinion in Psychology*, *20*, 111–116. https://doi.org/10.1016/j.copsyc.2017.09.003

Paulus, M., Kühn-Popp, N., Licata, M., Sodian, B., & Meinhardt, J. (2013). Neural correlates of prosocial behavior in infancy: Different neurophysiological mechanisms support the emergence of helping and comforting. *NeuroImage*, *66*, 522–530. https://doi.org/10.1016/j.neuroimage.2012.10.041

Paulus, M., Licata, M., Kristen, S., Thoermer, C., Woodward, A., & Sodian, B. (2015). Social understanding and self-regulation predict pre-schoolers' sharing with friends and disliked peers: A longitudinal study. *International Journal of Behavioral Development*, *39*, 53–64. https://doi.org/10.1177/0165025414537923

Pettygrove, D. M., Hammond, S. I., Karahuta, E. L., Waugh, W. E., & Brownell, C. A. (2013). From cleaning up to helping out: Parental socialization and children's early prosocial behavior. *Infant Behavior and Development*, *36*, 843–846. https://doi.org/10.1016/j.infbeh.2013.09.005

Ponitz, C. E. C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly*, *23*, 141–158. https://doi.org/10.1016/j.ecresq.2007.01.004

Rajhans, P., Altvater-Mackensen, N., Vaish, A., & Grossmann, T. (2016). Children's altruistic behavior in context: The role of emotional responsiveness and culture.

Scientific Reports, *6*, 24089. https://doi.org/10.1038/srep24089

Rodriguez, M. L., Mischel, W., & Shoda, Y. (1989). Cognitive person variables in the delay of gratification of older children at risk. *Journal of Personality and Social Psychology*, *57*, 358–367. https://doi.org/10.1037/0022-3514.57.2.358

Roy, S., Roy, C., Éthier-Majcher, C., Fortin, I., Belin, P., & Gosselin, F. (2007). *STOIC: A database of dynamic and static faces expressing highly recognizable emotions*. Retrieved from http://mapageweb.umontreal.ca/gosselif/cv.html

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

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Smetana, J. G., Ball, C. L., Jambon, M., & Yoo, H. N. (2018). Are young children's preferences and evaluations of moral and conventional transgressors associated with domain distinctions in judgments? *Journal of Experimental Child Psychology*, *173*, 284–303. https://doi.org/10.1016/j.jecp.2018.04.008

Smetana, J. G., Rote, W. M., Jambon, M., Tasopoulos-Chan, M., Villalobos, M., & Comer, J. (2012). Developmental changes and individual differences in young children's moral judgments. *Child Development*, *83*, 683–696. https://doi.org/10.1111/j.1467-8624.2011.01714.x

Steinbeis, N. (2018). Neurocognitive mechanisms of prosociality in childhood. *Current Opinion in Psychology*, *20*, 30–34. https://doi.org/10.1016/j.copsyc.2017.08.012

Svetlova, M., Nichols, S. R., & Brownell, C. A. (2010). Toddlers' prosocial behavior: from instrumental to empathic to altruistic helping. *Child Development*, *81*, 1814–1827. https://doi.org/10.1111/j.1467-8624.2010.01512.x

Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology*, *64*, 231–255. https://doi.org/10.1146/annurev-psych-113011-143812

Vaish, A., Carpenter, M., & Tomasello, M. (2009). Sympathy through affective perspective taking and its relation to prosocial behavior in toddlers. *Developmental Psychology*, *45*, 534–543. https://doi.org/10.1037/a0014322

Van de Vondervoort, J. W., & Hamlin, J. K. (2017). Preschoolers' social and moral judgments of third-party helpers and hinderers align with infants' social evaluations. *Journal of Experimental Child Psychology*, *164*, 136–151. https://doi.org/10.1016/j.jecp.2017.07.004

Van de Vondervoort, J. W., & Hamlin, J. K. (2018). Preschoolers focus on others' intentions when forming sociomoral judgments. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01851

Villegas de Posada, C., & Vargas-Trujillo, E. (2015). Moral reasoning and personal behavior: A meta-analytical review. *Review of General Psychology*, *19*, 408–424. https://doi.org/10.1037/gpr0000053

Warneken, F., Lohse, K., Melis, A. P., & Tomasello, M. (2011). Young children share the spoils after collaboration. *Psychological Science*, *22*, 267–273. https://doi.org/10.1177/0956797610395392

Wechsler, D. (2012a). *WPPSI–IV: Wechsler Preschool and Primary Scale of Intelligence—fourth edition*. Bloomington, MN: Pearson, Psychological Corporation.

Wechsler, D. (2012b). *WPPSI–IV technical and interpretive manual*. Bloomington, MN: Pearson, Psychological Corporation.

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

Yarrow, M. R., Waxler, C. Z., Barrett, D., Darby, J., King, R., Pickett, M., & Smith, J. (1976). Dimensions and correlates of prosocial behavior in young children. *Child Development*, *47*, 118–125. https://doi.org/10.2307/1128290

Young, L., & Dungan, J. (2012). Where in the brain is morality? Everywhere and maybe nowhere. *Social Neuroscience*, *7*, 1–10. https://doi.org/10.1080/17470919.2011.569146

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, *28*, 126–136. https://doi.org/10.1037/0012-1649.28.1.126

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, *99*, 432–442. https://doi.org/10.1037/0033-2909.99.3.432

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Appendix S1. Supplementary Materials and Analyses.