# Reciprocal Associations Between Young Children's Developing Moral Judgments and Theory of Mind

Judith G. Smetana and Marc Jambon
University of Rochester

Clare Conry-Murray
Pennsylvania State University, Beaver

Melissa L. Sturge-Apple
University of Rochester

Associations between young children's developing theory of mind (ToM) and judgments of prototypical moral transgressions were examined 3 times across 1 year in 70 American middle class 2.5- to 4-year-olds. Separate path models controlling for cross-time stability in judgments, within-time associations, and children's age at Wave 1 indicated that across both 6-month intervals, children who evaluated moral acts as more wrong independent of authority had more mature ToM 6 months later; in addition, judgments of moral transgressions as less permissible at Wave 2 also led to more advanced ToM at Wave 3. Children with more advanced ToM judged that moral rules are more alterable, however, and rated moral transgressions as less deserving of punishment. Finally, more advanced ToM initially led to evaluations of moral transgressions as less independent of rules and then to judgments of moral transgressions as more independent of rules. During the preschool years, early moral judgments and theory of mind appear to develop as reciprocal, bidirectional processes.

*Keywords:* moral judgments, theory of mind, preschool children

The development of both preschool children's moral judgments and young children's theory of mind (ToM), or their understanding of others' mental states (Harris, 2006; Wellman, 1990; Wellman, Fang, & Peterson, 2011), have both been topics of intense interest over the past 20 years. Although these two lines of research have proceeded independently, there has been increased recognition that they may be interrelated (Baird & Sokol, 2004; Chandler, Sokol, & Wainryb, 2000; Dunn, 2006; Knobe, 2005: Leslie, Knobe, & Cohen, 2006; Wainryb & Brehl, 2006; Wellman & Miller, 2008). Most of the available studies have at least implicitly assumed that a more advanced understanding of others' mental states is necessary for more mature moral judgments, but little research has tested this hypothesis directly or considered whether development in these two conceptual domains is mutually influential through

bidirectional, transactional processes. To address this gap, the present study examined reciprocal associations between preschool children's moral judgments and ToM longitudinally over 1 year.

## Development of Moral Judgments and ToM

According to social domain theory (Smetana, 2006; Turiel, 1983, 2006), children's understanding of morality—or their prescriptive understanding of right and wrong regarding others' welfare, justice, and rights—is distinct from their understanding of social conventions, or the arbitrary social norms that structure social life. Numerous studies have shown that children as young as 3 years of age (and more reliably by age 4) treat familiar, hypothetical, prototypical moral and social-conventional rules and transgressions as distinct in their judgments and as developing from qualitatively different social interactions (Turiel, 1983). In these studies, judgments typically have been assessed along several dimensions or criteria that define the domains (Smetana, 2006). That is, in contrast to conventions, young children view moral rules as unalterable; they also see moral transgressions as generalizably wrong (i.e., wrong in different contexts), not permissable, wrong independent of rules and authority dictates, and more deserving of punishment than conventional violations.

It is typically assumed that evaluations of moral transgressions depend on an awareness of others' intentions and motivations. Therefore, the widely replicated findings on young children's moral judgments are surprising when considered in the context of research on the development of preschool children's understanding of others. Young children are able to identify some emotions at a fairly young age (Denham, 1998), which may affect their moral judgments. But beyond this, research on ToM has shown that

mental state understanding develops in a consistent sequence in early childhood (Wellman & Liu, 2004; Wellman et al., 2011). Young children first develop an understanding of diverse desires, followed by diverse beliefs, and at around ages 4 or 5 years, an understanding of false beliefs, or the awareness that others might have beliefs that differ from reality and their own. Later, children become aware of how others will feel in the face of a mistaken belief (Wellman & Liu, 2004). Although some recent research has suggested that these abilities develop in toddlerhood or even infancy, these studies have employed implicit behavioral measures (e.g., Vaish, Carpenter, & Tomasello, 2009) or preferential looking and thus provide only indirect or inferential evidence for this hypothesis. Thus, it is unclear whether the development of children's moral judgments and children's understanding of others, including their mental states, proceeds independently or transactionally over time.

## Theory of Mind as Influencing Moral Judgments

Although young children distinguish moral and conventional acts, evidence suggests that young children's moral evaluations are limited in ways that may reflect their incomplete grasp of ToM. For instance, research has shown that 3-year-olds can predict future behavior in normal causal situations (e.g., hitting causes pain) but not in noncanonical, situations, where others are described as responding in atypical ways (e.g., hitting causes pleasure; Zelazo, Helwig, & Lau, 1996). Although 3-year-olds viewed acts resulting in harm as unacceptable, Zelazo et al. (1996) demonstrated that with age, children became better able to use intention information and consider both outcomes and intentions in judging the severity and punishment deserved for moral transgressions. More generally, research indicates that young children have difficulty in coordinating intentions, actions, and outcomes and instead describe their own intentions as matching the actual outcomes of their actions (Phillips, Baron-Cohen, & Rutter, 1998; Schult, 2002).

These studies suggest that ToM may be important in making moral judgments. Although most research has examined diverse and false beliefs using factual beliefs or physical knowledge (e.g., involving a change of location task), children's understanding of diverse moral beliefs also has been studied. For instance, Wainryb and Ford (1998) had children evaluate situations where hypothetical actors engaged in harmful (immoral) behaviors based on moral beliefs (e.g., that it is acceptable to be nicer to girls than boys) or informational beliefs (e.g., that it is acceptable to give a bigger snack to girls than boys because girls need more food) different than their own. They found that 3-year-olds uniformly gave negative evaluations of the situations because they could not grasp that others might have different informational or moral beliefs, even when they were explicitly told that this was so. However, 5- and 7-year-olds understood that others might legitimately have different informational beliefs, which led to more positive evaluations of unfair social practices than when the same practices were described as based on different moral beliefs. According to Wainryb and Brehl (2006), an understanding of others' mental states may lead to a greater appreciation of the complexities inherent in social situations. Therefore, they concluded that, "Prior to the age of 5 children's primitive psychological understandings significantly constrain their construals and judgments of moral situations" (p. 147).

Wainryb and Ford (1998) did not assess false beliefs using standard ToM tasks, but other studies have more directly compared moral and factual false beliefs. For instance, Flavell, Mumme, Green, and Flavell (1992) investigated children's judgments of stories where story characters were depicted as having beliefs different from their own or from other story characters. They found that 3-, 4-, and 5-year-olds' understanding of factual beliefs (as assessed in terms of standard false beliefs tasks), morality, social conventions, values, and ownership of property mostly proceeded synchronously within individuals and with age across the different task domains. They concluded that advances in all of these domains represented children's developing representational conception of the mind.

Using a somewhat different design, Killen, Mulvey, Richardson, Jampol, and Woodward (2011) administered traditional false beliefs tasks as well as a morally relevant false beliefs task (describing a transgressor who accidentally throws away another child's desired object) to 3-, 5-, and 7-year-olds. Children who did not pass the standard false belief tasks were unable to correctly attribute the transgressor's intentions in the morally relevant false belief task and therefore viewed the accidental transgression as more wrong and deserving of punishment. These results were seen as consistent with Wainryb and Brehl's (2006) claim that mental state understanding is important for complex moral judgments.

Finally, Lane, Wellman, Olson, LaBounty, and Kerr (2010) followed a sample of 3.5-year-olds longitudinally over 2 years to examine the influence of cognitive and affective perspective-taking (assessed in terms of emotion, false belief, and appearance-reality emotion understanding) on young children's prosocial moral reasoning. They found that better emotion and false belief understanding predicted more other-oriented (Level 2) prosocial moral reasoning, as defined within Eisenberg, Fabes, and Spinrad's (2006) stage theory, and that appearance-reality emotion understanding was concurrently associated with more advanced (Level 3) prosocial moral reasoning. The researchers concluded that a mature understanding of both emotional and mental states may help children progress beyond Level 2, other-oriented reasoning to societally focused (Level 3) prosocial moral reasoning. These conclusions are limited, however, as prosocial moral reasoning was not measured at the first wave.

Thus, these studies (Flavell et al., 1992; Killen et al., 2011; Lane et al., 2010; Wainryb & Ford, 1998) all subscribe in varying degrees to the view that the development of ToM underlies moral judgments. This position is based on a longstanding assumption that cognitive development is fundamental for social development. For instance, Kohlberg (1971) assumed that cognitive attainments were prerequisites for progress in moral judgment development; he mapped Piaget's logical stages onto his moral judgment stages and viewed the former as "necessary but not sufficient" for development of the latter. Likewise, researchers have assumed that ToM tasks assess basic abilities that are then applied to the social world. As Killen et al. (2011) noted, although the standard ToM tasks measure children's understanding of mental states, they do not include information about actors' social relationships. Therefore, these researchers asserted that children may find it more difficult to apply mental state understanding to moral situations than to standard ToM tasks, because the social world is complex and

ambiguous, and others' motivations, intentions, emotions, and desires may be difficult to infer.

## Moral Judgments as Influencing Theory of Mind

Others have offered the alternate view, however, that moral judgments may influence ToM. For instance, in a series of experimental studies with adults, Knobe (2005, 2010) found that the presence or absence of salient moral concerns (harmful outcomes) influenced whether individuals viewed acts as intentional. He concluded that moral judgments are used in ToM judgments and not the reverse. Although studies with adults do not address developmental questions, Leslie et al. (2006) extended this research to a study of 3-, 4- and 5-year-old children. They examined judgments regarding the positive or negative side effects (emotional reactions) resulting from positively or negatively intended behaviors. Four- and 5-year-olds (but not 3-year-olds) judged that a foreseen side effect was intentional, even when the story character denied that this was the case, when the consequences of the act were harmful but not when the consequences were positive. Thus, along with Knobe, Leslie et al. (2006) claimed that the moral nature of the act influences ToM judgments.

Children are immersed from birth in social relationships, and their attempts to understand social interactions and behaviors, such as why others commit wrongdoings or why such behaviors are done to them, may facilitate more advanced mental state understanding. For instance, children can use objective information (how much harm did the violation cause?) to evaluate the permissibility or severity of wrongdoings, especially for moral transgressions. But children also may be motivated to try to understand why others behave as they do, leading to more advanced ToM. Adult responses (e.g., "Why did you hurt him? You didn't have to hit," "Look what you did—you made him cry") and victims' responses to moral transgressions potentially can stimulate children's understanding of ToM. Indeed, Ruffman, Perner, and Parkin (1999) found that parents who disciplined their children by asking them to focus on the victim's feelings had offspring who were more advanced in their false belief understanding. These findings suggest that individual differences in parental responses to morally relevant situations may influence ToM. Moral judgments were not directly assessed, however, and the cross-sectional design of the study limited the conclusions that could be drawn about the directions of effects.

## The Present Study

Much recent theorizing and research have described development as occurring through reciprocal, transactional processes (Sameroff, 2009), but researchers have not considered whether children's developing moral judgments and theory of mind have bidirectional effects. That is, children's moral experiences and judgments may lead them to better understand others' beliefs, desires, and intentions, but at the same time, children's developing grasp of others' mental states also may produce advances in moral thinking.

Thus far, assumptions about associations between developing theory of mind and moral judgments have been examined primarily in cross-sectional studies. The contemporaneous assessments used in previous studies cannot fully test the central temporal premise of developmental change over time, however (Little, Card, Preacher, & McConnell, 2009), or more specifically, whether ToM predicts subsequent change in moral judgments after controlling for prior levels of moral judgments and vice versa. Thus, the present study employed a longitudinal design, providing a particularly powerful test of the dynamic associations between the development of ToM and moral judgments across 1 year. Children were between the ages of 2.5 and 4 years when they were first assessed. This age range was chosen because it represents the developmental period of significant growth in both moral judgments (which first emerge between 3 and 4 years of age) and ToM understanding (between 2 and 5 years, as assessed on verbal tasks). Constructs were assessed three times across a year to capture bidirectional influences that occur during this period of rapid developmental growth.

Children's understanding of theory of mind was assessed employing five of Wellman & Liu's (2004) tasks. Consistent with research on preschoolers' moral understanding (reviewed in Smetana, 2006), we examined moral judgments using prototypical moral transgressions that did not provide any explicit information about transgressors' desires, beliefs, or intentions. We assessed different moral criterion judgments, including nonpermissibility, authority and rule independence (whether the act would be wrong if the teacher did not see it or if there were no rule, respectively), rule nonalterability (whether moral rules can be changed), generalizability (whether the act is permissible in a different setting), and ratings of deserved punishment. Longitudinal paths from moral judgments to ToM understanding and vice versa were examined. Previous research has yielded no consistent evidence of gender differences in either moral judgments or ToM (Harris, 2006; Smetana, 2006; Wellman, 1990), but we also examined whether there were gender differences in the associations between ToM and moral judgments over time, although no specific hypotheses were tested.

There is some evidence that children's understanding of different moral criteria develop at different rates during the preschool years (Smetana, 1981; Smetana & Braeges, 1990; Smetana et al., in press). Therefore, we examined different moral criteria in separate analyses, as we expected that associations between children's ToM and moral evaluations would vary for different moral judgments. We did not expect to find significant effects of ToM on children's evaluations of the nonpermissibility, deserved punishment, and generalizability of acts, as these judgments can be made straightforwardly by observing social interactions, but we hypothesized that these moral judgments may have effects on ToM. We also hypothesized that there would be longitudinal, bidirectional links between ToM and moral judgments of rule and authority independence and rule alterability.

To test our models, path analysis within structural equation modeling (SEM) was used to examine the timing and direction of potential links between young children's ToM understanding and their moral judgments. Using SEM path models has several advantages over the analytic methods used in previous research (Kline, 2006). Path analysis provides for the testing of relationships among multiple manifest variables within longitudinal process models, allows for simultaneous assessment of multiple outcome variables, and produces evidence of model fit and misspecification. Thus, our use of SEM models allowed us to examine the unique, bidirectional associations between ToM and

moral judgments over time while controlling for the rank-order stability of judgments over time and the correlations between manifest variables assessed at the same time.

## Method

### Sample

The sample for this study included 70 children, 33 boys and 37 girls, who ranged in age from 2.44 to 4.27 years ($M = 3.39$, $SD = 0.48$) at the first interview. They were drawn from four daycare centers serving middle class families in a suburb of a city in the Northeastern United States; all parents of children between the ages of 2.5 and 4 years at each of the four daycare centers were invited to participate. The sample was 83% European American, 5% Asian, 1% African American, and 9% other (primarily biracial). Nearly all (91%) of the children lived in two-parent homes, with the remainder living in single-parent homes or other family configurations. All of the parents had at least some college education.

At the second wave of the study, which occurred 6 months after the initial interviews, we reinterviewed all 70 participants ($M = 3.85$ years, $SD = 0.48$). With two exceptions, interviews took place at participants' day care centers. Two children had changed centers and thus were interviewed in their homes. At the third wave, which occurred 1 year after Wave 1 interviews, we reinterviewed 65 children ($M = 4.34$ years, $SD = 0.48$). Three children had moved out of town, and two others could not be located. Thus, the overall retention rate was 93%. Attrition analyses comparing children who dropped or were retained showed that there were no differences between the groups in family background (mothers' or fathers' education or family marital status), children's age, ToM, or moral judgments at Wave 1.

### Measures

**Social rules interview.** The stimulus items pertained to events that were common and familiar to children at these ages and were drawn from items used in previous research with preschool children (Nucci & Turiel, 1978; Smetana, 1981; Smetana & Braeges, 1990). The hypothetical stimuli consisted of eight 8.5 × 11-in. (21.59 × 27.94-cm) colored drawings depicting familiar moral transgressions. The moral transgressions were hitting, shoving, teasing another child, and calling another child names (children were also asked about four conventional items, but these responses are not discussed here). We employed male and female versions of the stimuli, with the gender of the story characters matched to the child's gender. The items were presented in varying order.

For each stimulus item, children responded to the following questions in a fixed order: (a) "Is it OK or not OK for the child to __?" assessing *nonpermissibility*; (b) "What if the teacher didn't see him/her ___? Would it be OK to ____ then?" assessing *authority independence*; (c) "What if no one ever told him/her it was wrong to ___. Would it be OK to ____ then?" assessing *rule independence*; (d) "What if all the teachers got together and said that kids could ___. Would it be OK then?" assessing *rule nonalterability*; (e) "Now let's think about a different situation. Let's say the child was at home or another school. Would it be OK or not OK to ___ at home?" assessing *generalizability*; and (f) "Should

[the transgressor] get in trouble?" and if yes, "A little bit or a lot?" assessing *deserved punishment.*

Except for deserved punishment, the social judgment questions were scored categorically, with moral responses (e.g., that the act is not permissible, wrong independent of rules and authority, generalizably wrong, and that rules are not alterable) coded as 1 and nonmoral responses coded as 0. Responses regarding deserved punishment were scored on a 3-point scale ranging from 1 (*no punishment*) to 3 (*a lot*). For each question, mean responses to the four stimulus items were obtained for each participant.

**Theory of mind tasks.** Based on Wellman and Liu (2004), we administered five standard ToM tasks assessing diverse desires, diverse beliefs, false beliefs (both contents and location), and belief-emotion relationships. Because of the length of the interviews, we did not assess knowledge access or an understanding of real-apparent emotions (which has been found to develop at older ages than those sampled here). Thus, the tasks we selected spanned a broad range of difficulty and were appropriate for the age range of the sample. As described below, we used Wellman & Liu's (2004) scoring in each of the tasks.

**Diverse desires.** Children were first shown pictures of two snacks, a carrot and a cookie, and asked to choose their favorite, assessing their *own desires*. Next they were shown a doll (matched to the child's gender) and told that the doll prefers the choice opposite from their own. The *target question* asked which snack the doll would choose. Responses were coded as 1 (correct) if the child indicated that the doll would choose the snack different from his or her own desires. Incorrect responses were coded as 0.

**Diverse beliefs.** Children were shown a doll and a picture of a garage with bushes next to it. They were told that the doll wants to find his/her missing cat and that the cat might be hiding in the bushes or the garage. First, they were asked where they thought the cat might be hiding (in the bushes or in the garage), assessing their *own beliefs*. Then they were told that that the character thought the cat was in the opposite location. The *target question* asked them where the character would look for the cat. Responses indicating that the character would look in the location opposite from the child's own beliefs but consistent with the character's beliefs were coded as 1 (correct), and incorrect responses were coded as 0.

**False beliefs.** Two tasks were used to assess false beliefs. The first was an "unexpected contents false belief task" (Gopnik & Astington, 1988; Perner, Leekam, & Wimmer, 1987). As done by Gopnik and Astington (1988) in their "Smarties Task," children were first shown a box (in our implementation, a Band-Aid box) and asked what they thought was in the box. If they said something other than Band-Aids, they were corrected and asked again until they responded correctly. Then the child was allowed to examine the contents of the box, which had something unexpected inside (a toy, rather than Band-Aids). The child was shown a doll and told that the doll had never seen inside the box. We assessed children's understanding of false beliefs by asking the child two *target questions*: what the doll thought was inside the box and what was actually in the box. (Between the two target questions, children were asked a second memory check, whether the doll had seen inside the container. Incorrect responses were corrected before proceeding to the second target question.) At Wave 2, most children did not initially guess that the Band-Aid box held a toy, so the same materials as in Wave 1 were used. Because some children at Wave 2 did remember the manipulation from the previous waves,

however, the task was changed at Wave 3 to a Cheez-Its box with a toy inside. We followed the same procedure as before. Correct responses to each target question were coded as 1; responses indicating that at least one question was missed were coded as 0.

Participants also were administered a standard "change of location" ("Maxi and the chocolate") false beliefs task (Wimmer & Perner, 1983). They were shown cutout pictures of a doll named Maxi and a set of drawers. They were told that Maxi had put some chocolate in the drawer, which was then moved. The task was modified to indicate that a neighbor, represented by an additional cutout of a doll with a backpack, had taken the chocolate and put it in his or her backpack while Maxi had gone outside to play. We conducted a memory check by asking where Maxi originally put the chocolate. Incorrect responses were corrected to ensure that the child understood the story. Children were then asked two *target questions*: where the chocolate currently was and where Maxi would look for the chocolate. As in the previous false belief task, children had to answer both target questions correctly to be coded as passing the task.

**Belief-emotion.** Children were shown a doll and a Goldfish crackers box that, unknown to the child, contained rocks. They were first asked what they thought would be inside the box. The expected answer was Goldfish; children who gave an incorrect response were corrected. The doll then stated that he or she loves Goldfish and that it is his or her favorite snack. The experimenter then put away the doll, opened the box, showed the contents to the child, and remarked that the box contained rocks, not Goldfish. We conducted a memory check to confirm that the child understood that the box contained rocks and not crackers and that the doll's favorite snack was Goldfish. Again, children who failed the memory check were corrected before proceeding. Next, children were told that the doll was returning for snack time. They were asked whether the doll would feel happy or sad when he or she got the box (the *target question*). Then the box was opened, the doll was made to look into the box, and the child was asked what the doll saw. If they answered incorrectly, they were corrected before being asked the next question, which was how the doll felt (happy or sad) when he or she saw what was inside (the *emotion-control question*). To be coded as correct (scored as 1), children had to respond that the doll would feel happy when receiving the box and sad after looking in the box. All other responses were coded as 0.

Responses to each of the tasks were summed to form a total theory of mind score, which could range from 0 to 5.

## Procedures

Children for whom consent was obtained were interviewed in their daycare centers (or at home, for two children) at a quiet corner or in a separate room. A trained research assistant was present at each interview to record children's responses, including correct or incorrect ToM responses, on a checklist. Because of its length, the interview was administered in two sessions (and sometimes three, for the youngest children). Both the Social Rules Interview and the ToM tasks were divided into two parts and administered in separate sessions. The two interviews and the order of the moral items within the interview were counterbalanced. The ToM tasks were always administered in the same order within each interview. Based on Wellman & Liu's (2004) recommendation, easier tasks were administered first, but the same tasks

were used in each interview. Sessions always began with a warm-up task involving reading a book together.

## Results

### Descriptive Analyses

Means and standard deviations for the different judgments are shown in Table 1. As can be seen, on average, children passed half or less of the five tasks at Wave 1 and three to four of the five tasks by Wave 3. Preliminary analyses revealed no significant gender differences in ToM understanding or moral judgments at each wave. Correlations between ToM and moral judgments are shown in Table 1. As can be seen, ToM was moderately correlated with different moral judgments. Correlations among the five moral evaluations were moderate, ranging from .48 to .68 at Wave 1, from .34 to .58 at Wave 2, and from .43 to .55 at Wave 3.

### Across-Time Analyses of Theory of Mind and Moral Judgments

To examine our process model, we employed path analysis within a SEM framework (e.g., Kline, 2005). Model fit was evaluated using three recommended fit indices (McDonald & Ho, 2002). We employed the chi-square goodness of fit ($\chi^2$) and the Bentler comparative fit index (CFI; Bentler, 1990). Finally, we also examined the standardized root-mean-square residual (SRMR) as an indicator of residual fit and error of approximation, which may be more robust than other indices, to low sample size and low degrees of freedom (e.g., Kenny, Kaniskan, & McCoach, 2011). Nonsignificant chi-square statistics, CFIs greater than .90, and SRMRs less than .10 indicate acceptable model fit.

We also screened the data for univariate skewness and kurtosis. Large values of these statistics indicate non-normality of data for analysis variables, which may prove to be problematic in SEM analyses by introducing bias in the standard errors. To account for the possible effects of non-normality, standard errors for model parameters were derived using bootstrap procedures in AMOS. Bootstrapping techniques (Zhu, 1997) create multiple subsamples from the original sample to derive a sampling distribution, which is not limited by assumptions of normality. This procedure yields more accurate standard error estimations of model parameters (West, Finch, & Curran, 1995). Models were run requesting a maximum likelihood (ML) bootstrap on 500 samples with a 90% bias-corrected confidence interval; significance of model pathways was determined using bootstrapped standard errors. As AMOS bootstrapping procedures require complete data, the SPSS missing values program utilizing the expectation maximization algorithm (EM) estimator was used to impute missing values for the 2.5% of the data lost at the Time 3 assessment (Dempster, Laird, & Rubin, 1977). To examine whether the SPSS estimation was robust, we ran the model analyses deleting the five cases with missing data at Wave 3 to determine if there were differences between the full analyses with 70 cases, including estimated data, and the 65 complete cases. Model parameters were similar in magnitude and direction; no changes in results occurred with the use of the EM estimator. Thus, we use the imputed data in the final analysis and retained all 70 cases.

Table 1

*Means, Standard Deviationss, and Correlations for Theory of Mind and Moral Judgments*

| Variable | M | SD | W1 Age | W1 ToM | W2 ToM | W3 ToM |
|---|---|---|---|---|---|---|
| W1 Theory of Mind | 2.60 | 1.29 | .47** | 1.00 | .59** | .53** |
| W2 Theory of Mind | 3.16 | 1.44 | .60** | | 1.00 | .66** |
| W3 Theory of Mind | 3.49 | 1.41 | .38** | | | 1.00 |
| W1 Nonpermissibility | .93 | .20 | .22† | .22† | .17 | .15 |
| W2 Nonpermissibility | .96 | .12 | .14 | .05 | .07 | .11 |
| W3 Nonpermissibility | .98 | .08 | .15 | .15 | .13 | −.02 |
| W1 Authority Independence | .80 | .27 | .30* | .25* | .36** | .20 |
| W2 Authority Independence | .86 | .26 | .06 | −.04 | .02 | .15 |
| W3 Authority Independence | .92 | .21 | .08 | .40** | .28* | .01 |
| W1 Rule Independence | .87 | .25 | .09 | .09 | .06 | .13 |
| W2 Rule Independence | .86 | .26 | .07 | −.19 | −.03 | −.01 |
| W3 Rule Independence | .90 | .20 | .02 | .00 | .01 | −.01 |
| W1 Rules Not Alterable | .69 | .38 | .01 | .08 | .03 | .10 |
| W2 Rules Not Alterable | .65 | .36 | −.07 | −.33** | −.19 | −.06 |
| W3 Rules Not Alterable | .63 | .41 | −.24† | −.37** | −.45** | −.35** |
| W1 Generalizability | .81 | .30 | .04 | .14 | .09 | .11 |
| W2 Generalizability | .86 | .22 | .15 | .06 | .01 | .01 |
| W3 Generalizability | .87 | .25 | .08 | .04 | −.06 | .08 |
| W1 Deserved Punishment | 2.52 | .53 | .04 | −.10 | .02 | .09 |
| W2 Deserved Punishment | 2.50 | .62 | −.15 | −.16 | −.10 | −.22† |
| W3 Deserved Punishment | 2.52 | .52 | −.01 | −.16 | .04 | −.15 |

*Note.* W = Wave. Theory of mind (ToM) was scored on a 6-point scale ranging from 0 to 5; Punishment was scored on a 3-point scale ranging from 1 to 3.
† $p < .10$.  * $p < .05$.  ** $p < .01$.

Figure 1 shows all of the pathways estimated within each model. First, we included covariance pathways between ToM and moral judgment constructs at each time point. Second, autoregressive pathways within each construct were included from Wave 1 to Wave 2 and from Wave 2 to Wave 3 assessments. Within the context of the present study, autoregressive parameters demonstrate the level of stability across time. Path coefficients less than 1 in autoregressive paths reflect a change in the rank ordering of individuals on particular variables over time. Third, we included cross-lag pathways between ToM and moral judgments for the Wave 1–Wave 2 and Wave 2–Wave 3 assessments. Cross-lag paths between predictor variables and variables occurring later in time (e.g., ToM at Wave 1 and moral judgment at Wave 2) model how earlier occurring con-

structs may predict change over time in individual rank-ordering in outcome variables. Overall, the measures showed significant stability across time.

Finally, our sample was diverse with respect to age. To control for the potential impact of age on model processes, we included child age at Wave 1 as a covariate with Wave 1 ToM and moral judgment assessments as a predictor at Wave 2 and Wave 3 ToM and moral judgment assessments in all analyses. Although, for the sake of clarity, the effects for age are not depicted in the models, child age was uniquely associated only with Wave 2 ToM; older children had higher ToM at this wave. However, as a stringent test of the possible impact of age on transactional processes studied here, we control for child age in all model analyses. Thus, results presented are those that are significant over and above any effect
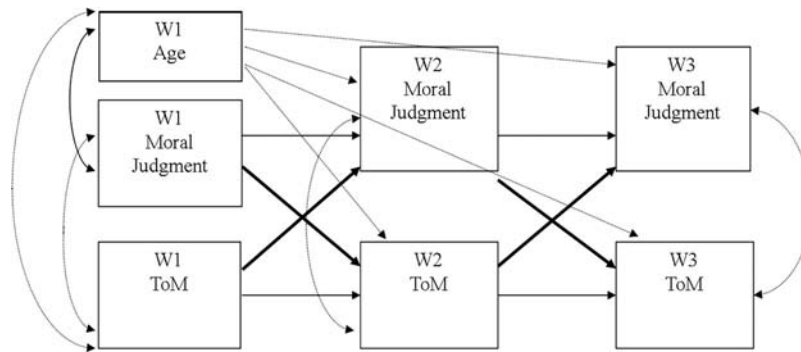


*Figure 1.* Conceptual model tested in path analyses. Bold lines indicate cross-lag predictor paths. Solid lines indicate auto-regressive paths. Curved lines indicated covariances between within-time estimates of moral judgment and theory of mind (ToM). Dotted lines indicate predictor pathways and covariances included to control for the effect of age on model pathways.

of child age on process pathways. The results are discussed separately for each moral judgment dimension.

Finally, given the potential moderating role of child gender in developmental models of children's ToM and moral judgments, we also examined whether the parameters in our transactional path model differed as a function of child gender. To test the moderating role of gender, we estimated the model presented in Figure 1 simultaneously for boys and girls using a multiple-group analysis. Our first model allowed all parameters to freely vary across boys and girls. We compared the fit of this model with a nested model in which the autoregressive pathways (e.g., Wave 1 ToM predicting Wave 2 ToM) and the cross-lag pathways (e.g., Wave1 ToM predicting Wave 2 moral judgment) were constrained to be equal across boys and girls. The change in model fit from the freely estimated model to the constrained model was examined using chi-square difference tests (i.e., Loehlin, 1998). Significant degradation in fit from the freely estimated model to the constrained model suggests that the pathways differ by gender. Model comparisons across the six different moral judgment criteria revealed only one significant gender difference out of a total of 40 possible ToM-moral judgment pathways. Thus, model comparisons suggested that pathways did not vary by child gender; all subsequent analyses were performed using the full sample.

**Nonpermissibility.** The model for nonpermissibility, shown in Figure 2, provided an excellent fit to the data, $\chi^2(4) = 2.96$, $p = .57$, CFI = 1.00, SRMR = .03. Controlling for the effects of children's age at Wave 1, the autoregressive paths, and the covariance between measures, there were no significant associations between Wave 1 and Wave 2 ToM and judgments of nonpermissibility. However, consistent with hypotheses, children who evaluated moral transgressions as more nonpermissible at Wave 2 had more advanced ToM at Wave 3.

**Authority independence.** The model for authority independence provided a good fit to the data, $\chi^2(4) = 3.20$, $p = .42$, CFI = .93, SRMR = .03. As shown in Figure 3, there were significant paths from moral judgments of authority independence to ToM 6 months later. That is, both from Wave 1 to Wave 2 and from Wave 2 to Wave 3, children who treated moral transgressions as more wrong independent of authority demonstrated more mature theory of mind understanding 6 months later.

**Rule independence.** The model for rule independence provided a good fit to the data, $\chi^2(4) = 7.18$, $p = .13$, CFI = .97, SRMR = .05. As shown in Figure 4, there were significant paths from ToM to rule independence, but the direction of the associations differed at the two lags. More specifically, children with more advanced ToM at Wave 1 judged moral transgressions as *less* independent of rules at Wave 2, but more advanced ToM at Wave 2 led to judgments of moral transgressions as *more* independent of rules at Wave 3.

**Rule alterability.** The model, presented in Figure 5, provided an excellent fit to the data, $\chi^2(4) = 1.87$, $p = .76$, CFI = 1.00, SRMR = .02. At both intervals, a more mature understanding of ToM led to judgments of moral rules as more alterable 6 months later.

**Generalizability.** The model for generalizability provided a good fit to the data, $\chi^2(4) = 6.55$, $p = .16$, CFI = .98, SRMR = .04, but there were no significant associations between moral judgments and ToM.

**Deserved punishment.** The model estimating paths for ToM and ratings of deserved punishment provided an adequate fit to the data, $\chi^2(4) = 8.46$, $p = .08$, CFI = .97, SRMR = .05. As shown in Figure 6, there was a significant path between Wave 1 ToM and Wave 2 ratings of deserved punishment. This showed that young children who had a more mature understanding of ToM at Wave 1 rated moral transgressors as less deserving of punishment 6 months later.

## Discussion

Research examining intersections between children's developing theory of mind and moral judgments typically has assumed that children's mental state knowledge, or their understanding of others' beliefs, desires, or intentions, informs complex moral thinking. Bidirectional influences rarely have been examined, and even in longitudinal studies (e.g., Lane et al., 2010), causal influences have not been tested using longitudinal designs. To identify whether these domains develop independently from one another or whether transactional processes are at play, the present study employed a three-wave longitudinal design across 1 year to more precisely identify the potential links between preschool children's
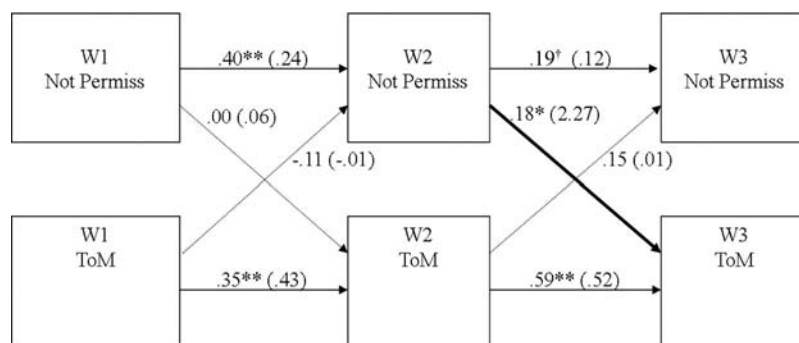


*Figure 2.* Path model of theory of mind (ToM) and moral nonpermissibility. Bolded lines represent significant paths. Unstandardized coefficients are in parentheses. For ease of interpretation, covariances and age effects are not presented; however, they are available from the first author upon request. $\chi^2(4) = 2.96$, $p < .57$, comparative fit index = 1.00, standardized root-mean-square residual = .03. $^\dagger p < .10$. $^* p < .05$. $^{**} p < .01$.
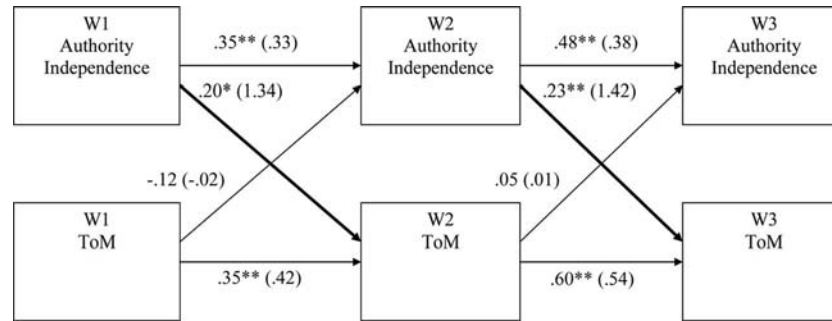
*Figure 3.* Path model of theory of mind (ToM) and authority independence. Bolded lines represent significant paths. Unstandardized coefficients are in parentheses. For ease of interpretation, covariances and age effects are not presented; however, they are available from the first author upon request. $\chi^2(4) = 3.20$, $p < .95$, comparative fit index = 1.00, standardized root-mean-square residual = .03. $^* p < .05.$ $^{**} p < .01.$

developing ToM, as assessed on five standard tasks, and their moral understanding of familiar, prototypical moral transgressions as assessed along multiple dimensions. To our knowledge, this is the first study to explicitly examine transactional associations between ToM and moral judgment criteria in early child development.

Toward this goal, we found significant longitudinal associations between children's developing understanding of ToM and different moral judgment criteria. Interestingly, links both to and from ToM and moral judgments appeared to operate in different ways. Significant paths from moral judgments to mental state understanding were found and suggested that moral judgments led to a more mature understanding of ToM. However, significant paths from ToM to moral judgments also suggested that better mental state understanding led to more flexible evaluations of hypothetical moral transgressions.

## Moral Judgments as Influencing Theory of Mind

First, the results indicated that young children who treated moral transgressions as more wrong independent of authority dictates (i.e., wrong even if a teacher did not see the act) had a more advanced understanding of others' mental states 6 months later. This finding was obtained from Wave 1 to Wave 2 and again from Wave 2 to Wave 3. Likewise, we found that children who evalu-

ated moral transgressions as more impermissible (less permissible) at Wave 2 had a more mature understanding of ToM 6 months later.

These findings suggest that children's attempts to understand and evaluate social relationships and events influenced their understanding of others' mental states. This conclusion may seem at odds with some previous research and theorizing suggesting that psychological understanding of others' mental states informs or constrains moral judgments (Killen et al., 2011; Wainryb & Brehl, 2006; Wainryb & Ford, 1998). However, our study differs from this previous research in that we examined straightforward, prototypical moral transgressions, whereas these other studies focused on more complex situations that required psychological inferences for correct conclusions about the moral situations to be drawn.

Recent research suggests that moral evaluations develop very early in childhood. Based on studies showing that preverbal infants prefer individuals who help rather than hinder others, as assessed by looking time (Hamlin & Wynn, 2011; Hamlin, Wynn, & Bloom, 2007), some researchers have claimed that morality emerges in infancy. Observational studies, however, show that children's awareness of moral and social rules emerges between the first and second year of life (Dunn, 2006). Studies explicitly assessing young children's judgments of different moral criteria have focused on children from 2.5 years of age on (Smetana &
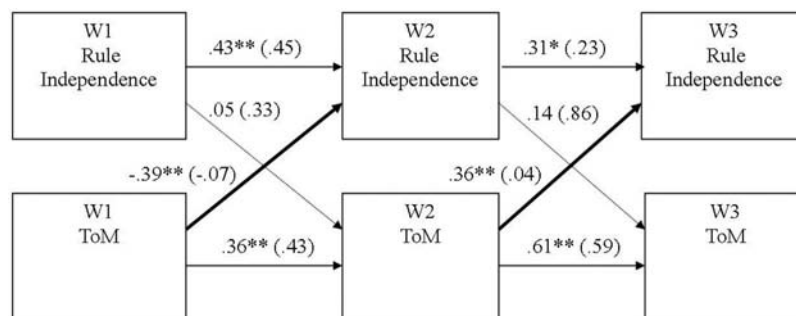


*Figure 4.* Path model of theory of mind (ToM) and moral rule independence. Bolded lines represent significant paths. Unstandardized coefficients are in parentheses. For ease of interpretation, covariances and age effects are not presented; however, they are available from the first author upon request. $\chi^2(4) = 7.18$, $p = .13$, comparative fit index = .97, standardized root-mean-square residual = .05. $^* p < .05.$ $^{**} p < .01.$
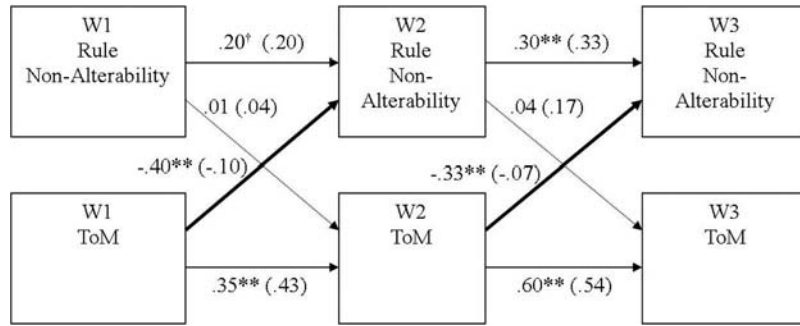
*Figure 5.* Cross-domain model of theory of mind (ToM) and moral rule nonalterability. Bolded lines represent significant paths. Unstandardized coefficients are in parentheses. For ease of interpretation, covariances and age effects are not presented; however, they are available from the first author upon request. $\chi^2(4) = 1.87$, $p = .76$, comparative fit index = 1.00, standardized root-mean-square residual = .02. $^\dagger p < .10.$ $^* p < .05.$ $^{**} p < .01.$

Braeges, 1990), and this research indicates that young children make rudimentary moral evaluations about prototypical transgressions well before the ages at which they evidence much mental state understanding.

This may be because some moral evaluations may not require mental state understanding. For instance, children can make judgments about the severity of acts (for themselves as well as for other children) by observing the effects of moral transgressions on others. Likewise, children can make judgments of generalizability by considering whether actions are prohibited in different contexts; indeed, other research has suggested that an understanding of generalizability may develop earlier than other moral criterion judgments (Smetana & Braeges, 1990). In other words, children can make these rudimentary moral judgments by observing the world without considering mental states, or "fitting the world to the mind" (Searle, 1983; Sokol, Chandler, & Jones, 2004).

However, judgments of authority independence demand somewhat more complex judgments that involve thinking about the consequences of moral acts under different conditions (with and without the presence of authority, for instance). They do not require drawing psychological inferences about wrongdoers' intentions and motivations and transgressors' and victims' desires, but they may spur that type of thinking. In other words, it may help children fit the mind to the world (Searle, 1983; Sokol et al., 2004). Our results suggest that some thinking about familiar, prototypical

moral transgressions may facilitate the movement from a "copy" to a representational ToM (Perner, 1991). The moral judgments obtained here regarding everyday, familiar prototypical transgressions also can be seen as an early step toward more advanced moral judgments regarding more complex or multifaceted situations that may also involve judgments of others' mental states.

The types of moral transgressions studied here are ubiquitous in young children lives (Dunn, 2006; Smetana, 2006). Therefore, it is not surprising that children develop mental state understanding in the context of these everyday social experiences. The finding that moral judgments spur ToM are consistent with longitudinal studies showing that family discourse, discussion of feelings, and shared pretense all predict differences in children's understanding of false beliefs (Dunn, Brown, & McGuire, 1995; Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991; Wellman & Miller, 2008). They are also consistent with recent theorizing from a constructivist perspective suggesting that children's mental state understanding develops from social relationships and communicative interactions with others (Carpendale & Lewis, 2004) and with research demonstrating that parental responses that focus the child on the victim's feelings in the context of moral transgressions are associated with more advanced false belief understanding (Ruffman et al., 1999). Further research is needed to examine both parents' responses and the characteristics of children's social and
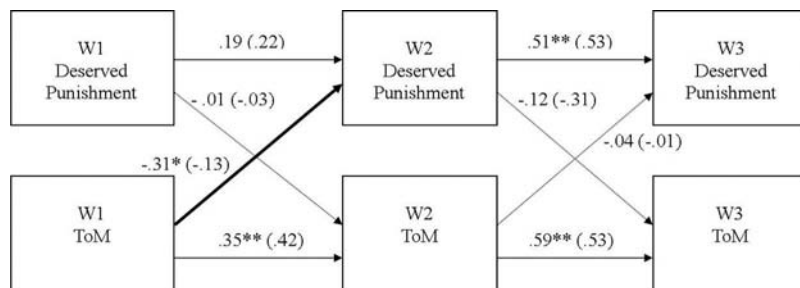


*Figure 6.* Model of theory of mind (ToM) and ratings of deserved punishment. Bolded lines represent significant paths. Unstandardized coefficients are in parentheses. For ease of interpretation, covariances and age effects are not presented; however, they are available from the first author upon request. $\chi^2(4) = 8.46$, $p = .08$, comparative fit index = .97, standardized root-mean-square residual = .05. $^* p < .05.$ $^{**} p < .01.$

moral interactions that facilitate more advanced mental state understanding.

## Theory of Mind as Influencing Moral Judgments

Our results were not consistently in the direction of moral judgments influencing children's psychological understanding of others' mental state, however. We also obtained some evidence for the effects of ToM on moral judgments, although these results were contrary to the expectation that advances in one realm would correspond to advances in the other. Instead, we found that a more developmentally advanced understanding of others' mental states led to less prototypically moral judgments. That is, although one criterion for defining morality is that moral rules are considered to be unalterable (Smetana, 2006; Turiel, 1983), we found that at two successive waves, children's more advanced understanding of others' mental states led to judgments of moral rules as more alterable 6 months later. In addition, children who initially had better mental state understanding treated moral violators as less deserving of punishment at Wave 2. It is important to note, however, that even though more advanced ToM led to greater flexibility in some moral judgments, the changes were associated with changes in individuals' rank-ordering on these variables over time, not overall mean levels of moral judgments. Children still viewed moral transgressions as highly independent of rules and authority, generalizably wrong, and deserving of punishment.

These results differ from previous research indicating that mental state understanding leads to a more nuanced understanding of moral transgressions (Killen et al., 2011). It is possible that as children develop a better understanding of others' mental states, they make what appear to be less prototypically moral judgments because their judgments become more flexible and less absolutistic, particularly in evaluating adults' ability to influence and change moral rules. Thus, children who are better able to understand others' desires, beliefs, and intentions may be able to see the world in a more complex way and attempt to coordinate their understanding of the moral nature of the misdeed with possible psychological attributions for why it occurred. For instance, a better understanding and appreciation for the psychological experience of others may lead children to entertain the possibility that moral transgressions may be more acceptable if the transgressor did not know that there was a rule. Likewise, as children's understanding of others' mental states develops, they may consider whether there were mitigating factors that led the transgressor behave as they did, leading to lowered ratings of deserved punishment (Zelazo et al., 1996).

The findings for rule independence are consistent with this speculation, as we found that the direction of the associations between ToM and judgments of rule independence differed over time. Children with a better initial understanding of ToM evaluated moral rules as less rule-independent at Wave 2, but a more advanced understanding of ToM at Wave 2 led to moral evaluations of greater rule independence later on. These findings, as well as the fact that the link between ToM and ratings of less deserved punishment was found only at the first 6-month interval, suggest that the greater flexibility found here in some judgments of straightforward moral transgressions may be a temporary perturbation in the development of more mature moral understanding. That is, children may initially overgeneralize their ToM understanding to situations that do not require it. For instance, children may evaluate another child as deserving less punishment for hitting once they grasp that individuals may have varied reasons for hitting, even though they realize that hitting is almost always wrong. This is consistent with recent research suggesting that transitions in moral thinking are sometimes characterized by what appear to be regressions in thinking, reflecting instability and periods of coordination as new forms of thinking emerge (Nucci & Turiel, 2009).

## Theory of Mind and Moral Judgments as Reciprocal, Interactive Processes

The two distinct patterns of results observed here are consistent with the hypothesis that ToM and early moral judgments develop as bidirectional, transactional processes. Children's experiences of moral transgressions provide the types of rich social interactions that facilitate an awareness of others' mental states, but as this awareness develops, children become better able to grapple with and understand more complex moral situations or dilemmas that entail conflicting beliefs or intentions that do not match the outcomes of actions (e.g., Phillips et al., 1998; Schult, 2002; Sokol et al., 2004; Zelazo et al., 1996). Thus, although preschool children have been found to distinguish morality from social conventions in their judgments along the dimensions studied here (Nucci & Turiel, 1978; Smetana, 1981; Smetana & Braeges, 1990), children may initially evidence more absolutistic and inflexible judgments about hypothetical prototypical moral transgressions than they do as their ToM understanding develops. It is important to note that the associations we obtained here are robust, as our modeling strategy provides a conservative test of transactional models (Burt, Obradović, Long, & Masten, 2008) in that we controlled for previous levels in examining change over time. In the present study, we explicitly examined bidirectional paths while controlling for autoregressive paths and associations across constructs at each time.

Finally, because participants in our study ranged in age from 2.5 to 4 years when they entered the study, we examined the potential impact of children's age at entry into the study on model processes. Age was found to have few significant effects on the pathways in the model; the significant associations between ToM and moral evaluations found here were obtained over and above effects for age. It is interesting that child age was not a significant covariate in our process models, and this warrants some discussion against the backdrop of the many studies that have identified critical age periods for the influence of ToM on moral judgments. Our findings suggest that although mean levels of ToM and moral judgments are associated with age, as identified in previous research, the transactional processes underlying the development of ToM and moral judgments are not age-dependent. Rather, children's developing ToM and moral judgments influence their further development in these domains regardless of children's age. Consideration of the underlying processes, as opposed to simple age differences, suggests that the field may benefit from considering more transactional frameworks that focus on the specific dynamic processes influencing the development of children's theory of mind and moral reasoning.

## Study Limitations and Future Directions

Although our study is the first to identify transactional processes across domains, there are limitations to our research. First, although our sample was homogeneous with respect to income and demographics, it is possible that a third, unmeasured variable, such as IQ or language ability, may account for findings. A recent meta-analysis indicates that young children's language ability influences false belief understanding (Milligan, Astington, & Dack, 2007). Research also has shown that children with more advanced language abilities have a better understanding of distinctions between conventional and moral transgressions, particularly among very young children (2- to 3-year-olds; Smetana & Braeges, 1990). Although this hypothesis should be investigated further, it seems unlikely that developing language (or more broadly, differences in children's intellectual abilities) can fully account for present results and particularly, the bidirectional associations between moral judgments and mental state understanding.

Second, although gender differences in the pathways between moral judgments and ToM were examined here, the sample size for the multiple-group analyses was small. Therefore, it is possible that significant gender differences in the associations between ToM and moral judgments exist but were not detected here. Third, our sample was homogeneously middle class and largely European American. Therefore, further research with larger and more diverse samples is needed to more fully explore gender, ethnicity, and socioeconomic status effects.

Finally, the present study did not assess all possible components of ToM and moral domains. From a developmental standpoint, an important strength is that we employed a set of ToM tasks that have been shown to form a developmental sequence in children's mental state understanding. In addition, the moral stimuli in this study focused exclusively on physical and psychological harm or distress, as children apply moral criteria to events involving physical harm more consistently and at earlier ages than unfairness (Smetana, 1981). Nevertheless, it remains for future research to determine if these different components of children's developing ToM (Wellman & Liu, 2004; Wellman et al., 2011) have varying associations with the different moral criteria studied here.

In summary, our results empirically document that as children grow older, they put their understanding of theory of mind to use in interpreting more complex moral situations. Children's morally relevant experiences and interactions may influence mental state understanding, but this, in turn, appears to be important in understanding more complex, multifaceted moral situations. More research is needed to fully articulate these processes and developmental pathways. Furthermore, it would be useful to examine the effects of moral judgments as well as other aspects of social relationships and interactions on young children's understanding of ToM to determine their relative effects. Nevertheless, the results of the present study demonstrate that children's early moral judgments and their developing understanding of other minds are not only interrelated but have causal influences. Through social interactions, children actively attempt to understand their social world, including both other minds and what they and others ought to do.

## References

Baird, J., & Sokol, B. (2004). *Connections between theory of mind and sociomoral development.* San Francisco, CA: Jossey-Bass.

Bentler, P. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246. doi:10.1037/0033-2909.107.2.238

Burt, K. B., Obradović, J., Long, J. D., & Masten, A. S. (2008). The interplay of social competence and psychopathology over 20 years: Testing transactional and cascade models. *Child Development, 79,* 359–374. doi:10.1111/j.1467-8624.2007.01130.x

Carpendale, J. I. M., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences, 27,* 79–96. doi:10.1017/S0140525X04000032

Chandler, M. J., Sokol, B. W., & Wainryb, C. (2000). Beliefs about truth and beliefs about rightness. *Child Development, 71,* 91–97. doi:10.1111/1467-8624.00122

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39,* 1–39.

Denham, S. A. (1998). *Emotional development in young children.* New York, NY: Guilford Press.

Dunn, J. (2006). Moral development in early childhood and social interaction in the family. In M. Killen, & J. Smetana (Eds.), *Handbook of moral development* (pp. 329–350). Mahwah, NJ: Erlbaum.

Dunn, J., Brown, J. R., & Maguire, M. (1995). The development of children's moral sensibility: Individual differences and emotion understanding. *Developmental Psychology, 31,* 649–659. doi:10.1037/0012-1649.31.4.649

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. *Child Development, 62,* 1352–1366. doi:10.2307/1130811

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In N. Eisenberg (Ed.), W. Damon, & R. M. Lerner (Series Eds.), *Handbook of child psychology. Vol. 3: Social, emotional, and personality development* (6th ed., pp. 646–718). New York, NY: Wiley.

Flavell, J. H., Mumme, D. L., Green, F. L., & Flavell, E. R. (1992). Young children's understanding of different types of beliefs. *Child Development, 63,* 960–977. doi:10.2307/1131247

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development, 59,* 26–37. doi:10.2307/1130386

Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development, 26,* 30–39. doi:10.1016/j.cogdev.2010.09.001

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluations by preverbal infants. *Nature, 450,* 557–559. doi:10.1038/nature06288

Harris, P. L. (2006). Social cognition. In D. Kuhn, R. S. Siegler (Eds.), & W. Damon (Series Ed.), *Handbook of child psychology. Vol. 23: Cognition, perception, and language* (6th ed., pp. 811–858). New York, NY: Wiley.

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2011). *The performance of RMSEA in models with small degrees of freedom.* Unpublished manuscript, University of Connecticut.

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

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed). New York, NY: Guilford Press.

Knobe, J. (2005). Theory of mind and moral cognition: Exploring the connections. *Trends in Cognitive Sciences, 9,* 357–359. doi:10.1016/j.tics.2005.06.011

Knobe, J. (2010). The person as moralist account and its alternatives. *Behavioral and Brain Sciences, 33,* 353–365. doi:10.1017/S0140525X1000230X

Kohlberg, L. (1971). From is to ought: How to commit the naturalistic

fallacy and get away with it in the study of moral development. In T. Mischel (Ed.), *Cognitive development and epistemology* (pp. 151–235). New York, NY: Academic Press.

Lane, J. D., Wellman, H. N., Olson, S. L., LaBounty, J., & Kerr, D. C. R. (2010). Theory of mind and emotion understanding predict moral development in early childhood. *British Journal of Developmental Psychology, 28,* 871–889. doi:10.1348/026151009X483056

Leslie, A., Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect: Theory of mind and moral judgment. *Psychological Science, 17,* 421–427. doi:10.1111/j.1467-9280.2006.01722.x

Little, T. D., Card, N. A., Preacher, K. J., & McConnell, E. (2009). Modeling longitudinal data from research on adolescence. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (3rd ed., Vol. 2, pp. 15–54). Hoboken, NJ: Wiley. doi:10.1002/9780470479193.adlpsy001003

Loehlin, J. C. (1998). *Latent variable models: An introduction to factor, path, and structural analysis.* Mahwah, NJ: Erlbaum.

McDonald, R. P., & Ho, M. H. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7,* 64–82.

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78,* 622–646. doi:10.1111/j.1467-8624.2007.01018.x

Nucci, L. P., & Turiel, E. (1978). Social interactions and the development of social concepts in preschool children. *Child Development, 49,* 400–407. doi:10.2307/1128704

Nucci, L. P., & Turiel, E. (2009). Capturing the complexity of moral development and education. *Mind, Brain, and Education, 3,* 151–159. doi:10.1111/j.1751-228X.2009.01065.x

Perner, J. (1991). *Understanding the representational mind.* Cambridge, MA: MIT Press.

Perner, J., Leekam, S., & Wimmer, H. (1987). Three-year olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5,* 125–137. doi:10.1111/j.2044-835X.1987.tb01048.x

Phillips, W., Baron-Cohen, S., & Rutter, M. (1998). Understanding intention in normal development and in autism. *British Journal of Developmental Psychology, 16,* 337–348. doi:10.1111/j.2044-835X.1998.tb00756.x

Ruffman, T., Perner, J., & Parkin, L. (1999). How parenting style affects false belief understanding. *Social Development, 8,* 395–411. doi:10.1111/1467-9507.00103

Sameroff, A. (2009). The transactional model. In A. Sameroff (Ed.), *The transactional model of development: How children and contexts shape each other* (pp. 3–21). Washington, DC: American Psychological Association. doi:10.1037/11877-001

Schult, C. (2002). Children's understanding of the distinction between intentions and desires. *Child Development, 73,* 1727–1747. doi:10.1111/1467-8624.t01-1-00502

Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind.* Cambridge, England: Cambridge University Press.

Smetana, J. (1981). Preschool children's conceptions of moral and social rules. *Child Development, 52,* 1333–1336. doi:10.2307/1129527

Smetana, J. (2006). Social-cognitive domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (pp. 119–153). Mahwah, NJ: Erlbaum.

Smetana, J. G., & Braeges, J. L. (1990). The development of toddlers' moral and conventional judgments. *Merrill-Palmer Quarterly, 36,* 329–346.

Smetana, J. G., Rote, W., Jambon, M., Tasopoulos-Chan, M., Villalobos, M., & Comer, J. (in press). Developmental changes and individual differences in young children's moral judgments. *Child Development.*

Sokol, B. W., Chandler, M. J., & Jones, C. (2004). From mechanical to autonomous agency: The relationship between children's moral judgments and their developing theories of mind. In J. Baird & B. W. Sokol (Eds.), *Connections between theory of mind and sociomoral development* (pp. 19–36). San Francisco, CA: Jossey-Bass.

Turiel, E. (1983). *The development of social knowledge: Morality and convention.* Cambridge, England: Cambridge University Press.

Turiel, E. (2006). The development of morality. In N. Eisenberg (Ed.) & W. Damon (Series Ed.), *Handbook of child psychology. Vol. 3: Social, emotional, and personality development* (6th ed., pp. 789–857). New York, NY: Wiley.

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. doi:10.1037/a0014322

Wainryb, C., & Brehl, B. (2006). I thought she knew that would hurt my feelings: Developing psychological knowledge and moral thinking. In R. Kail (Ed.), *Advances in child development and behavior* (pp. 131–171). New York, NY: Elsevier.

Wainryb, C., & Ford, S. (1998). Young children's evaluations of acts based on beliefs different from their own. *Merrill-Palmer Quarterly, 44,* 484–503.

Wellman, H. M. (1990). *The child's theory of mind.* Cambridge, MA: MIT Press.

Wellman, H. M., Fang, F., & Peterson, C. C. (2011). Sequential progressions in a theory-of-mind scale: Longitudinal perspectives. *Child Development, 82,* 780–792. doi:10.1111/j.1467-8624.2011.01583.x

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

Wellman, H. M., & Miller, J. (2008). Including deontic reasoning as fundamental to theory of mind. *Human Development, 51,* 105–135. doi:10.1159/000115958

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). Thousand Oaks, CA: Sage.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13,* 103–128. doi:10.1016/0010-0277(83)90004-5

Zelazo, P. D., Helwig, C. C., & Lau, A. (1996). Intention, act, and outcome in behavioral prediction and moral judgment. *Child Development, 67,* 2478–2492. doi:10.2307/1131635

Zhu, W. (1997). Making bootstrap statistical inferences: A tutorial. *Research Quarterly for Exercise and Sport, 68,* 44–55.