STUDIA UBB PSYCHOL.-PAED., LXVIII, 1, 2023, p. 29-60
(Recommended Citation)
DOI:10.24193/subbpsyped.2023.1.02

The relationship between theory of mind and children's moral judgment: A scoping review

Daniela Teodora SEUCAN[1], Laura VISU-PETRA[1,2*]

ABSTRACT. Theory of mind is the ability that allows us to decipher human behavior by understanding the desires, beliefs, intentions, and motivations of others and our own. It also permits us to predict others' behavior and model our own in order to act morally toward others. Moral judgment (the ability to appraise whether an intention or the outcome of an action is good or bad) involves mental state understanding, especially when detecting other's intentions. Because the relationship between theory of mind and moral judgment in early and middle childhood is far from being clearly understood, we conducted a scoping review that allowed us to advance valuable avenues for future research. Of the 34 papers initially identified in five databases (Google Scholar, Web of Science, Scopus, PsycInfo, and ResearchGate), 20 studies were eligible and analyzed concerning their main results and the assessments used for both theory of mind and moral judgment. Most studies reported a direct relationship (one study identifying even bidirectional links) between the two variables of interest. A discussion regarding the implications and future directions is put forward to advance research in children's understanding of their worlds (internal and external) in a socio-moral way.

Keywords: theory of mind, moral judgment, children, autism, scoping review

[1] Department of Psychology, Babeş-Bolyai University, 400015, Cluj-Napoca, Romania, E-mail address: danielaseucan@psychology.ro
[2] Department of Social and Human Research, Romanian Academy, 400088, Cluj-Napoca, Romania
* Corresponding author. Address for correspondence: Research in Individual Differences and Legal Psychology Lab (RIDDLE), Babeş-Bolyai University, 37 Republicii Street, 400015, Cluj-Napoca, Romania, Telephone: +40-264-590967; E-mail address: laurapetra@psychology.ro

Introduction

Very early on, as they enter preschool, children are faced with the complex nature of social interactions and have plenty of opportunities to witness moral transgressions (e.g., hitting, exclusion from play, unwillingness to share, rule-breaking, and teasing). In order to make accurate moral judgments, representing evaluations if someone's intentions or their actions and outcomes are wrong or right) and to flexibly respond to transgressions, they need to correctly assess complex situational information. Children must consider both transgressor's and victim's mental states, the outcomes, and even the level of negligence behind the action (Margoni & Surian, 2020; Nobes et al., 2009).

Recent research offers a more specific and differentiated theory of morality than Kohlberg's global stage model (1969), highlighting the importance of investigating the socio-cognitive components of moral behavior. This model proposes that moral development emerges from a self-oriented, pre-moral level (stages 1 and 2), and advances through a group-conventional level (stages 3 and 4) until it reaches a justice level (stages 5 and 6). A theory that complements Kohlberg's model is the Social Domain Theory (Nucci, 2001; Smetana, 2006; Turiel, 2006) which proposes that morality consists of three domains of social knowledge (moral, societal, and psychological) and is built out of social interactions. Empirical research has demonstrated that these three types of knowledge coexist within individuals and are essential tools for evaluating straightforward and complex situations (Killen & Rutland, 2011; Smetana, 2011; Turiel, 2006). Thus, the evaluation and interpretation of social events involve three types of reasoning: moral reasoning (physical and psychological harm; issues of fairness, equality, and justice), societal reasoning (concerns about conventions and social institutions, group norms and group functioning, traditions, and cultural rituals), and psychological reasoning (concerns with identity and autonomy, personal goals and individual prerogatives; Nucci, 2001).

The current review focuses only on the first type (moral reasoning) and its relationship with theory of mind (ToM; the ability to infer other people's desires, beliefs, and intentions; Wellman & Liu, 2004) since recent research has highlighted the importance of mental states understanding in moral cognition (Chakroff & Young, 2015; Young & Tsoi, 2013). The aim of this scoping review is to highlight the complex nature of the relationship between moral judgment and ToM and to propose future research directions that could fill the existing gaps in the literature.

Children's moral judgment progressively improves from age 3 onward (for a review, see Smetana et al., 2018), as children undergo rapid cognitive development. The ability to coordinate acts, intentions, and outcomes is

sustained by cognitive gains in children's information-processing skills that are visible between 3 and 7 years of age (Zelazo et al., 1996). After the age of 7, a more advanced ToM form develops, interpretive theory of mind (iToM, understanding that people's beliefs represent interpretation of information; Carpendale & Chandler, 1996; Lalonde & Chandler, 2002) that allows children to have a more profound social understanding. A child with iToM should be able to recognize both the RECIPIENT (the one needing help) and the AGENT (the one in a position of offering help) interpret the context from their individual perspective, and there might be a conflict between these two points of view. This complex understanding has great implications for moral judgment but the link between iToM and children's moral judgment needs further investigation (Harari & Weinstock, 2021).

Two essential components of moral reasoning are the ability to distinguish intentional from accidental actions and the ability to differentiate between positively and negatively valenced outcomes. A typical moral judgment task requires children to evaluate the acceptability / wrongness of an act ("When [AGENT] offered [RECIPIENT] the box, was [AGENT] doing a good thing, a bad thing, or just okay?") performed with intent or by accident and then assign an amount of punishment or reward to account for that action ("Should [AGENT] get in a lot of trouble, a little trouble, or no trouble?" Ochoa et al., 2022a). Punishment judgment refers to the act of imposing consequences on individuals who cause harm or deviate from societal norms (Cushman, 2008). Children expect wrongdoers to be penalized (Kenward & Osth, 2012) and they punish selfish peers both when they are directly affected and when a another has suffered injustice (Vaish et al., 2011). When children's moral judgment focuses mostly on intentions, their punishment judgments also become substantially more intent-based (Nobes et al., 2016).

In order to make accurate moral judgments, children need to be able to differentiate between *intentional harm* (when someone causes a negative outcome having a negative intention), *accidental harm* (when someone causes a negative outcome having a neutral or positive intention), and *attempted harm* (when someone causes nothing even though had the intention to cause harm). In order to be successful at this task, children need to manage to inhibit the outcome information and to focus primarily on the intention information. During preschool, children's moral judgments develop at an accelerated pace, with an important shift occurring from outcome-based moral judgment to intent-based moral judgment (Zelazo et al., 1996; see Margoni & Surian, 2016b for a review). In outcome-based moral judgments, children's evaluations rely on the consequences of an action, such as the victim's emotional distress cues in the case of harmful actions. Conversely, in intent-based moral judgments,

evaluations rely on the agent's mental states, which allows for distinguishing between intentional and accidental actions, and also between types of motives for performing an action (having good or bad intentions). For example, 3-year-olds tend to condemn the prosocial agent who accidentally performed an action that resulted in a negative outcome, suggesting that they focus more on the consequences of that act, rather than on the absence of a bad intention (Cushman et al., 2013; Margoni & Surian, 2017).

According to the dual-process model (Cushman, 2008, 2013; Cushman et al., 2013) act evaluations are mostly generated by the intent-based process (relying especially on mental states information), while the punishment evaluations are generated by both intent-based and outcome-based processes (relying on both mental states information and other consequences factors). Thus, these two types of moral evaluations (act and punishment judgments) are supported by two distinct underlying processes, rather than a developmental replacement of the outcome-based evaluations by the intent-based evaluations. When intentions and outcomes are in direct opposition to one another (good intentions - negative outcomes) children within this age range often struggle to coordinate such competing information (D'Esterre et al., 2019). This difficulty could stem from the fact that the outcome of an action - unlike the intentions - is immediately visible and does not need to be inferred. Supporting this hypothesis are results showing that young children are able to take intent into account when the agent's mental states are explicitly presented (Baird & Astington, 2004; Nelson, 1980). In addition, children with a higher ToM ability evaluate the moral quality of an action more accurately (if the actor has a good or a bad intention) (Dunn et al., 2000; Fu et al., 2014; Killen et al., 2011). Developmental changes in ToM are associated with intent-based moral judgment (Killen et al., 2011). These results indirectly suggest that ToM could be an important developmental socio-cognitive mechanism that supports the shift from outcome-to-intent in young children's sociomoral evaluations.

ToM skills involve progressive abilities that enable inference of other's mental states at different levels of recursive thinking (*first-order ToM*, Lisa thinks X; *second-order ToM*, Lisa thinks Anna thinks X), helping to understand the complexity of human social interactions (Rakoczy, 2022). In addition, a more complex ToM understanding, also called ***"morally-relevant" ToM*** (MoToM; Killen et al., 2011) enables children to intersect their mental states understanding and their moral reasoning in complex social and moral contexts. In the seminal study of Killen et al. (2011), a MoToM task was developed to see if and when children detect the intentions of an accidental transgressor (a child wanting to help the teacher clean the classroom throws away a classmate's bag, thinking it

contained garbage, when instead it contained a highly desirable item - a classmate's cupcake). The participants are asked (1) a contents false belief question ("What did the teacher's helper think was in the bag?"); (2) an intention question ("When the teacher's helper threw out the bag, did he/she think he/she was doing something that was all right or not all right?''); (3) a justification question (''Why?''); (4) an act evaluation question (''When the teacher's helper threw out the bag, do you think he/she was doing something that was all right or not all right?''; and (5) a justifications question (''Why?''). The next three questions referred to the actions of the victim: (6) a location false belief question (''Now the classmate wants to eat the cupcake that they brought in from home. Where will he/she look for his/her cupcake?''); (7) attributions of the emotional state of the victim (''How will the classmate feel about losing his cupcake?''); and (8) attributions of the victim emotion towards the accidental transgressor (''How will the classmate feel about the teacher's helper?''). The authors found that only the children who passed the false belief tasks in this morally-relevant scenario were able to dismiss the accidental transgressor from the negative outcome and assign less punishment than children who did not demonstrate a false belief understanding in this complex moral setting.

Subsequent research showed that children's MoToM serves as a strong predictor of their moral judgments (D'Esterre et al., 2019; Fu et al., 2014; Li et al., 2017). Moreover, MoToM was a better predictor for moral judgment than both age and classical ToM (false belief understanding; D'Esterre et al., 2019). A recent study (Glidden et al., 2021) found that MoToM competence also mediated between children's group membership and their moral judgments (intention evaluation and social exclusion decision). Nonetheless, even though research has shown that ToM informs and constrains moral judgment (Killen et al., 2011; Wainryb & Brehl, 2006), there are studies that found ToM not always required when children judged whether causing harm to another person was wrong (Zelazo et al., 1996) or when evaluating moral issues in different contexts (Smetana, Jambon, et al., 2012). Given these mixed results, the potential relationship between ToM and children's moral judgment requires deeper and further theoretical and methodological consideration.

Moreover, examining moral judgment in ***atypical development*** is another direction worth pursuing and possibly capable of elucidating this relationship. Of particular interest is autism (Autism Spectrum Disorder - ASD; APA, 2013), a developmental disability characterized by communication difficulties and severe social impairment. Since ToM is essential for moral judgment, and because ToM is impaired in autism, the investigation of the relationship between moral judgment and ToM in this atypical population could

bring valuable insights into children's moral judgments. The difficulties that define autism are often linked with difficulties in mental state understanding (Baron-Cohen et al., 2013; Peterson et al., 2012; Surian & Leslie 1999). Studies on moral reasoning in autism pertain to two lines of investigations (for a review see Margoni & Surian, 2016a). The first line examined if these children can differentiate between moral and conventional norms and found just a little delay in acquiring this knowledge as compared to typically developed (TD) children (Rogers et al. 2006; Shulman et al. 2012; Zalla et al. 2011), even though ASD children rely more on external factors (emotional cues; Margoni & Surian, 2016a) in their judgments. The second line investigated children's ability to decipher mental states before making a moral evaluation and the results are mixed and not straightforward (Buon et al. 2013; Grant et al. 2005; Salvano-Pardieu et al. 2016; Steele et al. 2003), ranging from completely relying on outcomes (Grant et al., 2005) to an immature intent-based moral judgment (Hamilton, 2009). Thus, further research with ASD children could bring more insight about the role of ToM in children's moral judgments.

Method

The current scoping review aimed to 'map' the emerging relevant literature at the intersection of the moral and the socio-cognitive domains and identify research gaps concerning the relationship between ToM and children's moral judgment. We started by using the five stages of the methodological framework for scoping reviews developed by Arksey and O'Malley (2005) and also followed the scoping review guidelines proposed by Tricco et al. (2018). First, we formulated the research question, and then we employed a search strategy that enabled us to identify and select relevant literature. Second, we charted and summarized the data and finally, reported the results and critically discussed them suggesting future research directions.

The computer-based search was conducted on high-impact databases in the fields of psychology (PsychInfo), education (ERIC), and interdisciplinary databases (Google Scholar, Web of Science, Science Direct, Scopus, ERIH PLUS) in order to identify relevant scientific papers. The keywords that we used were: theory of mind, false belief understanding (FB), mentalizing, first-order ToM, second-order ToM, moral judgment, moral evaluation, wrongness evaluation, act acceptability evaluation, and punishment attribution. The Boolean operator AND was used in addition to quotes for compound terms.

Inclusion-exclusion criteria

According to our inclusion-exclusion criteria (see Table 1), we excluded studies in which theory of mind and/or moral judgment was measured only by psychophysiological/imaging methods (such as fMRI and other brain scans), and those using indirect measures of these variables. Also, we selected only papers written in English and we limited them to the ones published in the last 25 years because of the methodological refinements brought to the moral judgment tasks in this interval).

Table 1. *Selection Stages of Inclusion and Exclusion Criteria*

Selection Stages	Inclusion Criteria	Exclusion Criteria
Stage 1 - papers retrieved based on titles and abstracts (a single researcher performed this stage) and stored in Sciwheel®	Papers investigating the relationship between ToM and moral judgment (MJ)	Addressing other constructs in moral reasoning (i.e., societal reasoning and psychological reasoning).
	Papers published in the last 25 years	
	Empirical studies published in English	Reviews, theoretical studies, and book chapters.
	The papers are available in full-text	The full text was not available or was not published in scientific journals
Stage 2 - papers selected based on the reading of the full texts (two independent researchers performed this stage) organized in a shared GoogleDrive®	Preschoolers and school-aged children with typical and atypical development	College students and adults
	Including ToM, MoToM, and prototypical moral judgment tasks.	Measuring other constructs pertaining to moral judgment (blame, "side-effect" effect).

Data collection and analysis procedure

Concerning data collection, for uniform access among researchers and for systematizing the general information, the Sciwheel® reference management program was used (Stage 1). GoogleDrive® was next used for storing and

organizing the relevant papers suitable for the descriptive analysis (Stage 2). As for the analysis procedure, the co-authors noted in a spreadsheet when and who performed the analysis, and the papers were subdivided into terms of theoretical concepts, aims, and objectives of the study, hypotheses, design, used measurements for the variables of interest, results, and discussion of results. The author with more expertise in the field mediated disagreements at each stage.

After identifying, evaluating, and eliminating the duplicates according to the inclusion-exclusion criteria, 19 papers were eligible for further analysis. The summary of the results of the included studies was organized into three sections. Therefore, the first section presents a brief description of the analyzed papers. The second section portrays a general overview of the main findings. The third section portrays the ToM-MJ relationship. Finally, we highlight the limitations, implications, and future directions in studying children's moral judgment.

Results

The search was conducted from March to April 2023 and a total of 78 papers were initially identified in the databases. After using filters set by date, age range, and after reviewing the research aims of the found studies, 44 studies were excluded from the initial sample, thus remaining 34 papers for full-text reading. Eighteen of these papers met all the inclusion criteria, the others (n = 16) being excluded for the next considerations: 1) not measuring ToM (n = 4) explicitly; 2) measuring other forms of moral reasoning (societal and psychological; n = 5); 3) investigating the ToM-moral judgment relationship in TD adults (n = 4) and 4) ASD adults (n = 3).

Table 2 presents a summary of the 20 studies investigating the relationship between ToM and moral judgment in TD and ASD children. The first study, dating from 2000, was conducted in the United Kingdom, and the more recent one is from Romania and is currently under review. The majority of the studies were conducted by authors located in the United States of America (n = 10), followed by China (n = 2), Spain (n = 2) the United Kingdom (n = 1), Italy (n = 1), Indonesia (n = 1), Brazil (n = 1), Israel (n = 1), Romania (n = 1). The participants were preschoolers and school-aged TD and ASD children. Some of the studies also included adult samples for comparison reasons. Age ranged from 2.5 to 15 years.

Concerning instruments for ***ToM assessment***, the majority of studies used the classical *false belief task* (first-order ToM; n = 10), the *second-order ToM task* (n = 3), an *interpretive ToM task* (n = 1), a *deception task* (n = 1), a *strange stories* task (Happé, 1994; cognitive and affective ToM; n = 1). For the

moral judgment assessment, the tools were more diverse, such as: three studies used *motive-based moral stories* (identifying the different motives behind identical actions; Baird & Astington, 2004); another three studies used *moral transgressions stories* (pushing someone off the swing; Smetana, 2006); one study used a *moral judgment of unintentional and intentional false claims* measure. The authors also employed *interview-based tasks* (the Social events interview, n = 2; The Morally-relevant belief vignettes, n = 2; the Moral interview, n = 1; Moral dilemmas, n = 1) (see Table 2 for more details). An interesting observation was that the studies reviewed used the morally-relevant ToM task sometimes for measuring children's ToM (n = 3) and sometimes for measuring children's moral judgment (n = 6).

The majority of the studies were quantitative, and cross-sectional, whereas two were longitudinal studies (Seucan et al., under review; Smetana et al., 2012). The analyses mainly consisted of correlational and intra- and inter-group inferential studies, using variables related to age, gender, act evaluation, punishment attribution, and scores obtained in the ToM tasks (both overall scores and scores obtained in subtasks: contents false belief, location false belief, second-order ToM, etc.). Five studies investigated the relationship between ToM and moral judgment with other variables of various constructs, such as language and IQ (predominantly used as control measures); empathy, emotion understanding, deception detection, friendship, and group membership. Moreover, one study (Glidden et al., 2021) investigated if ToM could be a mediator between children's group biases and their moral judgment (intention evaluation and peer exclusion decision).

Main findings

General overview

The majority of the studies reviewed in the current study identified a statistically significant association between ToM and children's moral judgment, even after controlling for other potentially relevant constructs, such as age, inhibitory control, verbal intelligence. There was only one study (Loureiro & Souza, 2013) to report a non-significant association between ToM-moral judgment. However, the study had only 24 TD children, potentially providing insufficient power to detect a relation. All the other studies reported a significant ToM-MJ relationship, but nonetheless, the direction of the relationship varied. For example, even though the majority of the studies found a positive association between the two variables of interest, there was one study (Hao & Liu, 2016)

that found that ToM and deontological moral judgments (moral rules should be abided by regardless of the consequences; Kant, 1959) were negatively correlated for children (8- to 10-year-olds), adolescents (13- to 15-year-olds), and older adults (60- to 70-year-olds) but positively correlated for younger adults (19- to 24-years old).

Another study (Glidden et al., 2021), found an indirect relationship between ToM and moral judgment, such that ToM mediated between children's group membership and their intent evaluations and the decision to exclude another (ingroup vs. outgroup). In this study, ToM understanding was investigated using a MoToM task and MJ was examined using an attribution of intention question. The authors found that when children had to attribute intention in an advantageous condition, morally-relevant ToM mediated the relation between group membership and moral judgment (see Table 2 for the summary of the papers). The single study measuring interpretive ToM (Harari & Weinstock, 2021), found that this particular ToM form, and not ToM false belief ToM understanding, enabled children to take the empathic perspective in prosocial moral conflicts and make more accurate evaluations.

Concerning the developmental milestones, after 3 ½ years of age, children's ToM (first-order ToM, contents false belief understanding, location false belief understanding) develops and helps them appropriately evaluate prototypic transgressions (Ball et al., 2017; Dunn et al., 2000; Fu et al., 2014). Later, between 4 and 6 years of age, first-order ToM contributes more to children's moral judgment in intentional transgressions and less in accidental transgressions (Kuntoro et al., 2018). Ochoa et al. (2022a) showed that 5-year-olds with higher ToM rate agents with false beliefs as more positively intentioned in good intent trials (even though the outcome was bad) than in bad intent trials (even though the outcome was good). Nonetheless, 4-year-olds with higher ToM did not manage to integrate their false belief understanding with their moral judgments any better than same-age children with low FB understanding, suggesting that before the age of 5, children cannot integrate beliefs that do not correspond to reality in their moral judgments even though they have a false belief understanding. Beginning with 6-7 years of age, when second-order ToM develops, children make accurate evaluations of the accidental transgressor's intention (Fu et al., 2014). Moreover, once children integrate their false belief understanding in a morally-relevant context (MoToM), this understanding predicts their moral judgment above and beyond age and false belief understanding (D'Esterre et al., 2019; Killen et al., 2011).

Table 2. *Summary of Papers Included in the Scoping Review*

Paper	Sample's origin and size	Participants	Aims of the study	ToM measure(s)	MJ measure(s)	Relationship ToM-MJ
Baird and Astington (2004)	USA (*N* = 42)	4-, 5- and 7-year-olds typically developed and with a behavior disorder	investigate children's ability to consider mental states in evaluating the moral quality of others' actions	-first-order false-belief tasks (Perner et al., 1987) -second-order false-belief tasks (Homer & Astington, 2001)	-motives-based moral reasoning task (moral stories without specifying the outcomes of the actions)	-children's ToM was significantly correlated with the act evaluation and the punishment attribution both for the typically developed and those with a behavior disorder
Ball et al. (2017)	USA (*N* = 108; 56 girls)	3 ½ years old	investigate associations among preschoolers' empathy, false belief understanding, and moral judgments	- unexpected contents false belief task -change of location false belief task (Wellman & Liu, 2004)	-Social events interview (Smetana, 1985) - moral stories without specifying the consequences of transgressions	-greater ToM associated with more mature moral judgments; -ToM associated with judgments of physical harm but not psychological harm; -significant interaction between ToM and empathy predicted judgments about psychological harm -preschoolers with low ToM did not distinguish between moral and conventional issues in judgments of deserved punishment
D'Esterre et al. (2019)	USA (*N* = 122; 62 girls)	4- to 10-year-olds	investigating ToM's role in children's ability to differentiate between intentional and unintentional false claims regarding resources.	- a content false belief task (Gopnik & Astington, 1988). - a morally relevant false belief ToM (MoToM; Killen et al., 2011)	-a moral judgment of unintentional -intentional false claims measure (developed by the authors)	- ToM predicted more favorable evaluations of the unintentional transgressor; -MoToM predicted children's responses for all of the assessments above and beyond age and false belief understanding.
Dunn et al. (2000)	UK (*N* = 128; 63 girls)	3- to 4-year-olds	investigating relations between children's views on the permissibility of transgressions involving friends and justification for such views, and their ToM	-seven theory of mind tasks (involved predicting, explaining or recalling a false belief; Cutting & Dunn, 1999) --a deception task (Sodian & Frith, 1992)	-Moral interview (Slomkowski & Killen, 1992)	-children's ToM was positively correlated with act permissibility and moral justification
Fadda et al. (2016)	Italy (*n* = 30 TD children and *n* = 30 ASD children; all boys)	10- to 12-year-olds	investigating whether ToM might foster children's autonomous MJ achievement.	-a second-order ToM task (Perner & Wimmer, 1985)	-MJ task (Piaget, 1932)	-children with ASD lacking ToM abilities judged guilty the protagonists of the two versions of morally appropriate behavior and focused more on outcomes than intentions
Fu et al. (2014)	China (*N* = 79; 39 girls)	4- to 7-year-olds	investigating the interrelationships between children's moral judgments	-content false belief task; (Wellman & Liu, 2004)	-MoToM task (Killen et al., 2011) -The prototypic	-second-order ToM and MoToM both played an important role in young children's accurate attributions of the accidental

Paper	Sample's origin and size	Participants	Aims of the study	ToM measure(s)	MJ measure(s)	Relationship ToM-MJ
			of accidental and prototypic transgressions and first-order and second-order ToM	-location false belief task (Wimmer & Perner, 1983) -two second-order false belief task (Astington et al., (2002) and Sullivan et al., (1994))	moral transgression story (Smetana et al., 2014)	transgressor's intention. -first-order ToM helps children appropriately evaluate prototypic transgressions
Garcia-Molina et al. (2019)	Spain (*N* = 60) *n* = 30 TD; 4 girls *n* = 30 ASD, 4 girls	7- to 12-year olds	investigating the link between ToM and moral judgment	-first-order ToM (De Villiers & De Villiers, 2012)	MJ task (Molina et al., 2019)	-ASD children had difficulties in moral judgments in a context involving deception; -ASD children based their justifications less on mental states that TD children
Garcia-Molina et al. (2020)	Spain (*N* = 62) *n* = 32 TD; 7 girls *n* = 30 ASD, 5 girls	7- to 12-year olds	investigating the ToM-MJ relationship in TD and ASD children	-4 'Faux Pas' stories of accidental situations (Baron-Cohen et al., 1999)	-4 stories of intentional situations (Garcia-Molina et al., 2016)	-ASD children had difficulties in understanding the moral transgressions when the action directly affected another person and not an object;
Glidden et al. (2021)	USA (*N* = 120; 64 girls)	4- to 7-year olds	investigating the relationship between ToM, group membership, and MJ.	-MoToM question (false belief evaluation)	-Attribution of Intention (AoI) question (acceptability of intention evaluation)	-MoToM mediated the relations between group membership and attribution of intentions in an morally-relevant advantageous condition, but not when an advantage was a straightforward moral transgression. -MoToM was predictive of person judgments and varied based on ingroup/outgroup status of the target.
Gönülta s et al. (2021)	USA (*N* = 117; 81 girls)	3- to 8-year olds	investigating whether children consider victim negligence when making judgments and whether children's ToM influences their assessments	-a false-belief contents task (Wellman & Liu, 2004)	-two moral transgression stories (MoToM; Killen et al., 2011; Nobes et al., 2009).	-children with higher ToM were more likely to consider victim's negligence when making moral judgments. -children with higher ToM were more likely to use negligence information in their punishment judgments for the transgressor and victim in both conditions (negligent vs. careful). -children with lower ToM did not differentiate their punishment judgments for the transgressor and victim.
Harari & Weinsto ck (2020)	Israel (*N* = 225; 120 girls)	7- to 11-year olds	investigating if iToM would predict better than ToM prosocial moral reasoning	-a iToM task (Lalonde & Chandler, 2002)	-three prosocial moral dilemmas (Eisenberg-Berg & Hand, 1979)	-iToM, but not ToM, predicted empathic and internalized values of prosocial moral reasoning, even when controlling for age, inhibitory control and emotion understanding.
Hao & Liu	China (*N* =	8- to 10-year olds	investigated whether ToM was	-the strange stories	-the moral dilemmas	ToM and deontological moral judgments were negatively

Paper	Sample's origin and size	Participants	Aims of the study	ToM measure(s)	MJ measure(s)	Relationship ToM-MJ
(2016)	204; *n* = 48; 21 girls; n = 45; 26 girls; n = 62; 36 females ; n = 49; 24 females)	13- to 15-year-olds 19- to 24-years 60- to 70-years	consistently positively associated with MJ from middle childhood to late adulthood.	(cognitive and affective ToM; Happé, 1994)	(Hauser et al., 2007)	correlated for children, adolescents, and older adults but positively correlated for younger adults.
Killen et al. (2011)	USA Study 1 (*N* = 162; 90 girls) Study 2 (*N* = 46; 25 girls)	3.5-, 5.5- to 7.5- year-olds	investigating children's false belief ToM in a morally relevant context	Study 1 & 2 -contents false belief task -location false belief task (Wellman & Liu, 2004)	Study 1 & 2 -MoToM task (Killen et al., 2011) -prototypical moral transgression task (Smetana, 2006)	Study 1 -children with higher ToM were more likely to attribute good intentions to the accidental transgressor than children who didn't pass the false belief tasks; Study 2 -children with higher ToM were less likely to punish the accidental transgressor than those who didn't pass the false belief tasks.
Kuntoro et al. (2018)	Indonesia (*N* = 122)	4- to 6-year-olds	investigating the contribution of children's ToM in morally relevant situations and moral judgment	-five ToM tasks (diverse desires, diverse beliefs, knowledge access, false beliefs, and, hidden emotions; Wellman & Liu, 2004)	-MoToM (Killen et al., 2011) -a scale of prototypical moral transgressions (Smetana, 2006)	-ToM contributes towards children's moral judgment in intentional moral transgressions but not in accidental transgressions
Lane et al. (2010)	USA (N = 128; 68 girls	3.5- to 5.5-year-old	longitudinally investigating how ToM and emotion understanding concurrently and prospectively predict children's MJ	-false belief understanding (Bartsch & Wellman, 1989) -appearance-reality emotion understanding (Harris et al., 1986)	-moral judgment task (Eisenberg-Berg & Roth, 1980)	-a more advanced ToM predicted greater use of psychological-needs reasoning; -ToM and emotion understanding jointly predicted moral judgment
Loureiro and Souza (2013)	Brazil (*N* = 24; 13 girls)	4- to 6-year-olds	investigating the relation between ToM and moral judgment (based on intention and based on motive)	-4 ToM tasks (Wellman & Liu, 2004)	-a MJ task based on motive (Baird & Astington, 2004); -a MJ task based on intention (Nelson-Le Gall, 1985)	- correlations between ToM and scores in the moral development tasks were not significant
Ochoa et al. (2022a)	USA Study 1 (*N* = 64; 28 girls) Study 2 (*N* = 109; 55 girls; *n* = 42 adults)	4- to 5-year-olds	investigating the influence of false belief understanding on moral judgment Study 1	Study 1 -standard false belief task (Leslie et al., 2005)	Study 1 -Morally-relevant Belief Vignettes (2 stories with true belief and 2 stories with false belief) -standard moral transgression task (Smetana,	Study 1 - 5-year-olds with higher FB rated agents with false beliefs as more positively intentioned in good intent trials (even though the outcome was bad) than in bad intent trials (even though the outcome was good); - 4-year-olds with higher FB did not manage to integrate their false belief understanding with their

Paper	Sample's origin and size	Participants	Aims of the study	ToM measure(s)	MJ measure(s)	Relationship ToM-MJ
					2006)	moral judgments any better than same age children with low FB understanding. - all had difficulties assigning punishment based on intent.
			Study 2 - reduced task demands, simplified design; punishment and reward evaluation	Study 2 -Morally-relevant Belief Vignettes Agent belief ("What does [AGENT] think is in the container?")	Study 2 -Morally-relevant Belief Vignettes Agent intention evaluation ("When [AGENT] handed [RECIPIENT] the box, was [AGENT] trying to be nice, mean, or just okay?"); Agent consequence evaluation ("Why should [AGENT] get [assigned consequence]) Assign punishment/reward	Study 2 -4-year-olds with higher FB made appropriate intent judgments; - children of all ages had difficulties assigning punishment based on intent.
Ochoa et al. (2022b)	USA (*N* = 61; 24 girls)	5- to 7-year-olds	investigating relations between false belief understanding and moral judgments	-Morally-relevant Belief Vignettes Agent belief ("What does [AGENT] think is in the container?")	-Morally-relevant Belief Vignettes Agent intention evaluation ("When [AGENT] handed [RECIPIENT] the box, was [AGENT] trying to be nice, mean, or just okay?")	- 5-year-olds consistently rated agents with false beliefs as better intentioned in a good intent condition (even though the outcome was bad) than in a bad intent condition (even though the outcome was good).
Seucan et al. (under review)	Romania (N = 92; 43 girls)	3- to 5-year-olds	investigating relations between ToM, emotion understanding and moral judgment	-3 ToM tasks (Wellman & Liu, 2004)	- 3 moral stories (Baird & Astington, 2004)	-MJ was associated with ToM and with emotion understanding; -punishment evaluation was predicted only by emotion understanding
Smetana et al. (2012)	USA (N = 70; 37 girls)	2.5- to 4-year-olds	investigating associations between children's ToM and judgments of prototypical moral transgressions	-five ToM tasks (diverse desires, diverse beliefs, false beliefs (both contents and location), and belief-emotion relationships)	-Social rules interview (Smetana & Braeges, 1990)	-early MJ and ToM develop as reciprocal, bidirectional processes: -children who evaluated moral acts as more wrong independent of authority had more mature ToM 6 months later; -judgments of moral transgressions as less permissible at Wave 2 also led to more advanced ToM at Wave 3; -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

Discussing the association between ToM and moral judgment

Act evaluation

Research aims to unravel the role that understanding other minds plays in children's ability to decide if an action with a good or a bad outcome was performed with good or bad intentions. The studies included in the current scoping review found a direct relationship between ToM and moral judgment, yet this relationship differed in terms of ToM type (first-order ToM, second-order ToM, morally-relevant ToM) and in terms of moral transgression type (prototypical standard transgression, morally-relevant transgression). As such, Baird and Astington (2004) found that 5- to 7-year-olds' ability to consider people's intentions to make moral distinctions between identical actions was facilitated by children's false belief understanding. Using the same task, but on a younger sample (3- to-5-year-olds), Seucan et al. (under review) showed that ToM enables such young children to make accurate moral judgments, correctly identifying the intention (good vs. bad) behind identical actions. Building on previous findings (Lane et al., 2010), in their study, alongside ToM, emotion understanding (Pons et al., 2004) was also a predictor of such young children's moral judgments (and even of punishment attributions), indicating that ToM - although necessary in young children's moral judgments - may not be sufficient to enable them.

A somewhat similar result was obtained by Ball et al. (2016) who investigated the association between 3½-year-olds' moral judgments and their ToM and empathy. They found that both ToM and empathy were associated with moral judgment in related but distinct ways, according to the type of harm and related judgment. Children in their study evaluated physical harm acts more negatively than acts of unfairness. Even though they condemned unfairly taking another's resources, preschoolers evaluated inflicting physical harm on another as more serious. Ball et al. (2016) found that higher ToM was associated with children's view of moral transgressions as invariably wrong, independent of authority mandates or rules. In their study, the consequences of the transgressions were not specified, so children needed to infer them before making their judgments.

In addition, higher ToM was associated with more mature judgments about psychological harm, but only for less empathic preschoolers. Therefore, low-empathy preschoolers might be more inclined to apply their mental states understanding to infer psychological harm. Conversely, Ball et al. (2016) found that for preschoolers lower in ToM, higher empathy was associated with more mature moral judgments. This result suggests that children with less advanced

ToM may rely more on affective information such as the perceived distress of the victim when evaluating moral acts (Arsenio & Ford, 1985). Another interesting result was that empathy and moral judgment were not linked in preschool children with higher ToM scores. Thus, children that were able to detect others' mental states may not need empathy to comprehend that conduct which produces psychological harm is morally wrong and negatively affects the victim.

The findings of the two previous studies emphasize the importance of considering in tandem the affective and cognitive processes in order to fully understand the specific and potential reciprocal relationship between young children's evolving moral judgments and the underlying socio-cognitive mechanisms. Another important variable that needs to be considered when moral transgressions take place is *negligence*, both of the transgressor's and of the victim's. For example, if a child is on a swing (accidental transgressor) and another child being inattentive (victim) passes in front of the child on the swing and gets hit and falls down, the transgressor should be judged less harshly since the negligence of the victim accidentally contributed to the harm inflicted. Building on previous results of Mulvey et al. (2020) that showed that children consider both transgressors' and victims' negligence when evaluating moral-relevant situations, Gönültas et al. (2021) demonstrated that indeed, children consider both transgressors' and victims' negligence when making moral judgments in property damage and physical harm stories, and this ability is facilitated by false belief understanding.

Investigating the acceptability of transgressions across conditions (victim careful/transgressor negligent or victim negligent/transgressor careful) the authors (Gönültas et al., 2021) extended earlier research (Nobes et al., 2009) by emphasizing the importance of also considering the victims' negligence when making moral judgments. Children with higher ToM were more likely to consider the victim's negligence, especially in the physical harm story, compared with children with lower ToM. In the property damage story, children with higher ToM evaluated more negatively the negligent transgressor compared with children with lower ToM. This finding suggests that ToM might enable children to consider both the victim's and transgressor's intentions and actions when making moral judgments.

Another worth-mentioning finding of Gönültas et al. (2021) was that ToM accounted for children's *alternative actions proposals* (for both the transgressor and victim) in morally-relevant situations. Their results showed that children with more advanced ToM were more likely than children with lower ToM to suggest ways in which both the transgressor and the victim could have acted differently. This result suggests that ToM may favor the flexible consideration of alternative actions by envisioning different possible outcomes for different types of actions.

Interestingly, there was no difference in the acceptability judgments of the transgressor's negligence between children with higher ToM and children with lower ToM in the physical harm story. A possible explanation was that the negative outcome of the physical harm was more easily detected by children, irrespective of their ToM ability, compared to the intent detection in the property damage story, confirming studies showing that young children are particularly sensitive to harm infliction (e.g., Helwig et al., 1995, 2001). This result suggests that ToM may contribute only to some types of moral judgments (psychological harm, fairness judgments, property damage) and not to others (physical harm, rule-breaking). Moreover, different types of ToM (first-order ToM, second-order ToM, morally-relevant ToM) may have different contributions according to moral transgression type (prototypical standard transgression, morally-relevant transgression). Indeed, Fu et al. (2014) investigated the interrelationships between children's moral judgments of accidental and prototypic transgressions and first-order and second-order ToM. They showed that children's moral judgments were related to their first-order ToM. Moreover, children who performed better in a MoToM task made more accurate judgments about the intention of the accidental transgressor (e.g., throwing away a bag containing another's preferred object) than those who could not correctly attribute false beliefs to the victim and to the transgressor.

Fu et al. (2014) were the first to show that, even after partialling out the age effect, it was not first-order ToM but rather second-order ToM that helped children to accurately evaluate the accidental transgressor's intention. Children that were able to engage in this level of recursive thinking (e.g., that the victim thought that the transgressor thought that there was trash inside the bag) were more likely to judge the intention of the transgressor less negatively. This result is in concert with previous studies that found that children's second-order ToM also counts for responsibility attribution (Yuill & Perner, 1988). Fu et al. (2014) showed that second-order ToM was useful in situations where the children needed to coordinate multiple perspectives in order to make accurate judgments of wrongdoing (accidental transgression). However, when the intention and the action were not in conflict (prototypical transgression; pushing someone off the swing), second-order ToM was not significantly related to children's moral judgments.

Another study which showed that ToM predicted more favorable evaluations of the accidental transgressor is that of D'Esterre et al. (2019). The authors investigated ToM's role in children's ability to differentiate between intentional and unintentional false statements regarding claims to resources. Moreover, they found that MoToM predicted children's responses for all of the assessments above and beyond age and false belief understanding. Similarly, another study (Ochoa et al., 2022 a, b) reported significant results concerning

children's MoToM and their performance on moral judgment tasks. The authors used a morally relevant belief vignettes task as a dual assessment of false belief understanding and moral reasoning in a sample of 4- to 5-year-olds and found important developmental changes, compared to previous findings (Killen et al., 2011) indicating that it was only from the age of 7-8 years that children could integrate false belief understanding in morally-relevant contexts. Moreover, children performed worse on a morally-relevant vignette task than on a standard moral transgression task. They had difficulties with the intention and the consequences questions (true belief, bad intent). One possible explanation could be the complexity of the morally-relevant task (e.g., an undesirable object hidden in a box) as compared to the standard task (e.g., someone hurting their knee).

In the previous study, the vignettes featured two characters, one of whom (the agent) discovered in an opaque container a pleasant animal (kitten) and in another, an unpleasant animal (skunk). The agent decided to share (prosocial vs. antisocial behavior) one of these containers with another, after the animals switched their containers. For two stories the agent knew the contents of the container (true belief condition) and for other two stories, the agent had a mistaken belief about the contents (false belief condition). On the one hand, in the *true belief condition*, where the agent handed what they believed they would hand, a desirable pet (kitten; good intent) versus an undesirable pet (skunk; bad intent) all children performed well in evaluating intention and consequence. They evaluated the agent with the bad intention as being meaner and assigned more punishment compared with the good intent condition, irrespective of their ToM level. Moreover, the punishment evaluation matched the moral judgment in this true belief condition. On the other hand, in the *false belief condition*, where the agent unknowingly didn't manage to hand what they believed they would hand (offering a skunk even though they thought that in the box is the kitten they had put inside) children had difficulties with the intention question, especially those with low ToM. This result extends previous findings (Killen et al., 2011), showing that ToM enables children to make a more accurate intent evaluation in morally-relevant situations.

Nonetheless, only 5-year-olds with higher ToM were able to integrate intention in morally-relevant contexts (rating agents with false beliefs as more positively intentioned in good intent trials (even though the outcome was bad) than in bad intent trials (even though the outcome was good) as compared with 4-year-olds also with high ToM. Irrespective of their ToM level (low vs. high), 4-year-olds didn't manage to integrate intention in these contexts where intention and outcome were in conflict. 4-year-olds with higher ToM were able to detect that the agent had a mistaken belief about the contents of the box but were unable to use this knowledge to evaluate the agent's intention.

Punishment attribution

Punishment represents a penalty or a retribution directed toward those who inflict harm or violate social norms (Cushman, 2008). Even though prior research showed that moral judgment and deserved punishment attribution are highly correlated (Smetana, Jambon, et al., 2012), Ball et al. (2016) found a low to moderate correlation between these two ratings. Similarly, Ochoa et al. (2022a) found that moral judgment and punishment attribution were correlated but only in the true belief condition (as compared with the false belief condition, where even the 5-year-olds with higher ToM performed no better than chance). This result is in accordance with previous findings that showed that young children, especially those without false belief understanding, often focus more on outcomes when assigning consequences (Cushman et al., 2013; Zelazo et al., 1996). Moreover, Cushman et al. (2013) observed a lag between integrating false beliefs in intent judgments and doing so in punishment judgments.

The findings of Ochoa et al. (2022a) support Cushman et al.'s (2013) view and Zelazo et al.'s findings (1996), which showed that when processing demands were reduced, thus children could focus on relevant information, both 4-year-olds and 5-year-olds with higher ToM were able to correctly rate agents' intentions compared with 4 and 5-year-olds with lower ToM. Nonetheless, in Ochoa et al. (2022a) the simplification of the task did not have an impact on punishment attribution, with children being unable to integrate false beliefs into punishment judgments. Interestingly, all children made reference more often to mental states than outcomes when justifying punishment attribution. However, children with low ToM that incorrectly determined the agent's intention seemed to match the intent of the agent with the outcome in their justifications, thus still focusing on the outcome when evaluating deserved punishment. Another study (Ball et al., 2016) found that preschoolers with relatively low ToM did not distinguish between moral and conventional issues in judgments of deserved punishment. One possible explanation for this result could be the fact that both the moral and the conventional violations were presented without specifying the outcomes. Because both these violations are legitimately punishable, children low in ToM might have had difficulties in considering the supplemental information that needs to be analyzed in case of the moral violations in order to make more nuanced punishment attributions to these kinds of violations as compared with the conventional ones.

The previous results suggest that other factors besides ToM might influence children in their punishment evaluations. Ball et al. (2016) found that higher *empathy* was associated with ratings of greater deserved punishment for fairness violations but not for physical harm or, unexpectedly, for psychological harm. Thus, punishment may be particularly important in helping children

attend to and enforce issues of fairness. Correspondingly, as more empathic preschoolers tended to judge fairness violations as more serious than their less empathic peers, they also assigned more punishment. Empathy was not related to severity judgments regarding physical harm. One possible explanation could be that the child didn't need to take the perspective of the victim since the consequences were immediately evident. Thus, in the contexts where the outcomes are salient and the child sees the inflicted harm and doesn't need to infer it, empathy may not be necessary to identify and judge the harm as being punishable. Also pertaining to the affective processing when calculating deserved punishment, Seucan et al. (under review) showed that punishment was associated with the ability to understand other people's emotions. In their study, even though no information was offered about how the victim in the situation may be feeling, 3- to 5-year-olds who had a higher ability to infer other people's emotions correctly assigned more punishment to the bad character and less punishment to the good character. It is possible that having this ability allowed children to simulate whether the victim in the situation would be upset or happy if the agent carried on with the action (e.g., *'The brother will be sad if his sand castle gets wrecked by his sister.'*) and assign an appropriate punishment.

Apart from taking the affective perspective of the victim, children might also consider the negligence of the agent before deciding the amount of deserved punishment. Similar to Nobes, Panagiotaki, and Bartholomew (2016) and Nobes et al. (2017), the findings of Gönültas et al. (2021) showed that children judged negligent actions as more punishable than careful actions. The authors showed that when children had to assign punishment, children with higher ToM scores were more likely to use both the transgressor's and the victim's negligence information in their judgments whereas children with lower ToM scores did not differentiate between their punishment judgments for the transgressor and victim. This finding indicates that ToM may enable children to shift between intention and outcome information when making moral judgments. Moreover, the punishment was differentiated based on ToM and story. In the property damage story, children with higher ToM scores were less likely to assign punishment when the victim was negligent. A possible explanation might be that children with higher ToM might have considered that having a special cupcake thrown away is punishment enough or that the victim simply forgot to label the bag containing the cupcake and that forgetfulness is not punishable. The fact that the same difference was not visible in punishment evaluations of the victim in the physical harm story, suggests that children might consider the type of harm elicited in different moral contexts before making their moral evaluations. In other words, the possible interaction between negligence and ToM might be diminished by the salience of the severity of the act.

Summary and future directions

The aim of the current scoping review was to investigate the relationship between theory of mind and children's moral judgment. The majority of the analyzed studies showed ToM to be directly related to moral judgment; one study found ToM to be a mediator between group membership and yet another found bidirectional links between ToM and children's moral judgment. Smetana et al. (2012) designed a 1-year three-wave longitudinal design to identify potential links between preschool children's ToM and their moral judgments in different contexts. They found significant bidirectional longitudinal associations between the two, with moral judgment leading to a more mature ToM, and also, with ToM enabling a more accurate evaluation of moral transgressions.

Concerning ToM influence on moral judgment, Smetana et al. (2012) found that children more advanced in ToM, although they considered moral rules as more alterable, they attributed less punishment. One possible explanation is that as children become more proficient in understanding the mental states underlying a complex moral action (e.g., coordinating their evaluation of the misdeed with possible psychological motives for why it happened), their moral judgments become more flexible and less absolutistic. Also, children with a higher ToM could take into account other factors (e.g., negligence of the transgressor and of the victim; Zelazo et al., 1996) that determined the transgressor to behave the way it did (e.g., the victim forgetting to label the bag containing the special cupcake), leading to lowered ratings of deserved punishment (Gönültas et al., 2021).

Concerning moral judgment's influence on ToM, it seems that ToM is influenced by children's efforts to understand and evaluate complex social relationships. Smetana and Braeges (1990) showed that starting with 2.5 years of age children can make accurate judgments in prototypical straightforward moral transgressions, even though at this age they do not have yet a first-order ToM understanding. Before the age of 4, when it is documented that children have a full-fledged ToM, they might evaluate the severity of acts by observing the effects of moral transgressions on others without inferring mental states (Sokol, Chandler, & Jones, 2004). This finding, that moral judgments enhance ToM is similar to results obtained in observational and longitudinal studies showing that family discourse (Dunn, 2006), and parental responses that focus the child on the victim's feelings in the context of moral transgressions (Ruffman et al., 1999) predict differences in children's ToM. These findings are also in line with the constructivist theory (Carpendale & Lewis, 2004) that highlights that children's ToM develops from communicative interactions and social relationships with others.

Concerning the reciprocal links between ToM and moral judgment, children's day-to-day experiences with moral transgressions (e.g., hitting, pushing, exclusion from play, unwillingness to share, rule-breaking, and teasing) supply them with a rich socio-moral environment that encourages and facilitates the development of mental state understanding. But as this skill develops, children become better at analyzing and understanding complex moral settings where intentions do not match outcomes, or other factors need to be considered before making a moral decision (e.g., group biases, negligence; emotion understanding; Glidden et al., 2021; Gönültas et al., 2021; Seucan et al., under review).

The studies with ASD children also add valuable information regarding the relationship between the two variables of interest. Garcia-Molina et al. (2019) showed that ASD children had greater difficulty than TD children in recognizing and explaining intentionality and action-morality in basic moral stories. Significant differences were found between the two groups in their justifications, showing that ASD children failed to apply the required ToM ability to understand and judge the intention and morality of the act. In a later study, investigating also preadolescents with ASD (7- to 12-years old), Garcia-Molina et al. (2022) found that ASD children as compared with TD children had difficulties in understanding the moral transgressions when the action directly affected another person (e.g., having the desire to take revenge on another person) but not when it affected an object (e.g., having the desire to obtain an object belonging to someone else by stealing it). The difference between these two types of judgments could stem from the fact that in the former case, the child needs to understand the mental states of the transgressor, with his specific desires and beliefs. Second, there is the understanding that the victim, who is affected by the action, also has a mind and does not know the information about the transgressor's desires.

Regardless of whether the cases are ambiguous or unambiguous, to resolve them correctly, the mental states of the victim and of the transgressor need to be integrated, thus imposing a greater complexity for ASD individuals (Moran et al., 2001; Zalla et al., 2011; Zalla & Leboyer, 2011) and even for TD children at younger ages (Killen at la., 2011). Also, when they had to evaluate the intention and the agent's morality, ASD children rated the agent in faux-pas scenarios as "bad" even if the intention was previously rated as "good". Their responses might be influenced by the bad outcome (e.g., 'she was hurt by his remark') even when the agent's intention was understood (e.g., 'he wanted to help her'). Thus, a discrepancy arises when the moral context is ambiguous, ASD children base their judgment on the outcome and not on the intention, suggesting that autistic individuals could have ToM-related deficits as a stumbling block in their moral judgment performance (Margoni & Surian, 2016a).

The studies reviewed seem to indicate that both conceptual changes and information processing improvements are likely implicated in the integration of ToM within moral judgments. Considering the processing demands, numerous studies have shown that executive functioning impacts ToM development (e.g., Carlson & Moses, 2001; Devine & Hughes, 2014). Specifically, as executive functioning is necessary but not sufficient for ToM, it may be that they are not sufficient for integrating ToM within moral reasoning. Even when processing requirements were reduced and the methodological adjustments did improve 4-year-olds' evaluations of intention, punishment judgments were not improved (Ochoa et al., 2022). Even when processing requirements were reduced and the methodological adjustment did improve 4-year-olds' evaluations of intention, punishment judgments were not improved (Ochoa et al., 2022a). Thus, as Cushman et al. (2013) have argued, the incorporation of ToM into punishment evaluations may require a further conceptual advance. According to their results, they conclude that two systems (outcome-based vs. intent-based) operate in determining punishment judgments. Across development, a conceptual reorganization occurs such that the intent-based system increasingly constrains, but does not fully override outcome-based punishment judgments. These arguments of Cushman et al. (2013) found recent support in Ochoa et al. (2022a) study results but need further investigation to portray a clearer picture of children's moral act evaluation and punishment attribution.

Even though the studies included in this scoping review have a valuable contribution to the field of moral development, there are several ***limitations*** to be considered. The main limitation of the studies included refers to the correlational, cross-sectional nature, reducing the possibility of making inferences about causal influences or reciprocal links between ToM and moral judgment. Longitudinal research would be ideal for investigating this relationship more thoroughly. Another limitation is the limited types of moral transgression (e.g., property damage, physical harm) investigated in studies trying to untangle children's moral reasoning. Future research could include moral situations involving resource allocation, social exclusion, and deception. In addition, the fixed order of the questions might prime children in their subsequent responses given that Nobes et al. (2016) showed that children's answers are influenced by this methodological adjustment. Also, to better understand how the first-, second-order ToM, and MoToM interact with each other to influence children's moral evaluations, explicit questions evaluating second-order ToM could be embedded within the story (e.g., "What does the victim think that the transgressor thought was inside the bag?").

Future directions are also worth advancing in order to inform the researchers which aspects regarding children's moral reasoning need further

clarification. Future studies could include a more widespread age range, evaluating school-aged children and adolescents' moral reasoning to better capture any variation in the maturation of moral judgment. Also, longitudinal designs are needed to shed light on the mechanisms related to social-cognitive abilities and children's moral judgments in social interactions as the relations among these processes may vary across different developmental periods. Smetana et al.'s (2012) longitudinal results not only highlighted that ToM and moral judgment are interrelated but also that they have causal influences on each other. Future studies should further investigate these links and also, see if these links exist between ToM and the other two forms of moral reasoning from Social Domain Theory (societal and psychological; Turiel, 2006).

In order to determine what features of the complex situation envisioned by the accidental transgression were difficult to process and evaluate by children, a wider range of potential transgressions should be used. Moreover, developing new MoToM tasks for different types of transgressions would be helpful for future research. Also, varying other aspects of the context such as the agent's familiarity (e.g., friends, siblings, strangers), or the nature of the misdeed. Moreover, the development of early moral concepts may also be influenced by moral emotions. For instance, Hoffman (2000) has suggested and Kochanska et al. (2002) have demonstrated that guilt following indiscretions enables children to better understand moral contexts. Future research should examine how moral emotions like guilt or shame shape children's harm infliction evaluations.

The study by Glidden at al. (2021) showed that MoToM was not a mediator between intention evaluation and ingroup bias in the intentional unfair condition, but only in the other two conditions (intentional fair and unintentional unfair), suggesting that children's MoToM ability might be context-dependent. Future research could investigate when and why children rely on their morally-relevant ToM. Also, including tasks that assess children's interpretive ToM could enable us to see how children interpret a morally-relevant situation based on the characters' emotions and their position in the event (Ross, 2006), thus opening new avenues for research. Investigating other cognitive factors, such as language (Dunn et al., 2000; Milligan et al., 2007) and executive functioning might explain the mechanism behind the relationship between ToM and moral judgments (Buon et al., 2016). A great body of research, both cross-sectional and longitudinal, documented the tight links between ToM and executive functions (see Devine & Hughes, 2014, for a meta-analysis), as interindividual variation in ToM was found to be due to executive functioning (Devine & Hughes, 2014). Thus, future research should investigate if executive functions (e.g., inhibitory control, shifting) may be an explanatory factor in the role of individual variation of ToM in children's moral judgments of transgressions.

These results have important ***implications*** for moral education, as teachers might facilitate young children's understanding of moral contexts by requiring them to consider all aspects of a transgression and help them take the perspective of both the transgressor and victim into account when making moral judgments and possible punishment decisions. In complex moral interactions, educators could help children integrate intention and outcomes while also considering additional information (e.g., negligence, group biases). Furthermore, these findings also have important implications for understanding how social-cognitive abilities interact with information about the victim and the transgressor's intention (and negligence) to influence children's decision-making processes in different socio-moral contexts.

Conclusion

ToM is an essential prerequisite for children's moral judgment. It is more predictive than age when it comes to making accurate moral reasoning. Moreover, second-order ToM and MoToM are better predictors than both age and first-order ToM (D'Esterre et al., 2019), showing that children need to be capable of complex recursive thinking when judging complex moral settings. When compared with classical ToM measures (e.g., location false belief task; content false belief task), children's MoToM abilities are better at predicting children's moral understanding and reasoning. One possible explanation might be that MoToM represents a more ecologically valid measure by requiring children to ascribe mental states in socially and morally multifaceted complex scenarios. Nonetheless, future studies should develop new MoToM tasks, as current results point to the fact that MoToM ability might be context-dependent.

Acknowledgment

This research was supported by Unitatea Executivă Pentru Finanţarea Învăţământului Superior, a Cercetării, Dezvoltării şi Inovării (UEFISCDI), project number PN-III-P1-1.1-TE-2021-1051.

REFERENCES

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology: Theory and Practice, 8*(1), 19 –32. https://doi.org/10.1080/1364557032000119616

Arsenio, W., & Ford, M. E. (1985). The role of affective information in social-cognitive development. Children's differentiation of moral and conventional events. *Merrill-Palmer Quarterly, 31*, 1-17.

Astington, J. W., Pelletier, J., & Homer, B. (2002). Theory of mind and epistemological development: The relation between children's second order false belief understanding and their ability to reason about evidence. *New Ideas in Psychology, 20*, 131–144. doi:10.1016/S0732- 118X(02)00005-3

Baird, J. A., & Astington, J. W. (2004). The role of mental state understanding in the development of moral cognition and moral action. *New Directions for Child and Adolescent Development*, *103*, 37–49. https://doi.org/10.1002/cd.96

Ball, C. L., Smetana, J. G., & Sturge-Apple, M. L. (2016). Following my head and my heart: Integrating preschoolers' empathy, theory of mind, and moral judgments. *Child Development*, *88*(2), 597–611. https://doi.org/10.1111/cdev.12605

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders, 29*(5), 407–418. https:// doi.org/10.1023/A:1023035012436

Baron-Cohen, S., Tager-Flusberg, H., & Lombardo, M. (Eds.). (2013). Understanding other minds: Perspectives from developmental social neuroscience. Oxford: Oxford University Press. https:// doi.org/10.1093/acprof:oso/9780199692972.001.0001.

Bartsch, K., & Wellman, H. M. (1989). Young children's attribution of action to beliefs and desires. *Child Development, 60*, 946–964

Buon, M., Dupoux, E., Jacob, P., Chaste, P., Leboyer, M., & Zalla, T. (2013). The role of causal and intentional judgments in moral reasoning in individuals with high functioning autism. *Journal of Autism Developmental Disorders, 43*, 458–470. https://doi. org/10.1007/s10803-012-1588-7

Buon, M., Seara-Cardoso, A., & Viding, E. (2016). Why (and how) should we study the interplay between emotional arousal, Theory of Mind, and inhibitory control to understand moral cognition? *Psychonomic Bulletin and Review*, *23*(6), 1660–1680. https://doi.org/10.3758/s13423-016-1042-5

Carlson, S. M., Mandell, D. J., & Williams, L. (2004). Executive function and theory of mind: Stability and prediction from ages 2 to 3. *Developmental Psychology*, *40*(6), 1105–1122. https://doi.org/10.1037/0012-1649.40.6.1105

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032–1053. https://doi.org/10.1111/1467-8624.00333

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*(1), 79–151. https://doi.org/10.1017/S0140525X04000032

Carpendale, J. I., & Chandler, M. J. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development, 67*(4), 1686–1706. https://doi.org/10.2307/1131725

Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development, 70*(4), 853–865. https://doi.org/10.1111/1467-8624.00061

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, *108*(2), 353–380. https://doi.org/10.1016/j.cognition.2008.03.006

Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition*, *127*(1), 6–21. https://doi.org/10.1016/j.cognition.2012.11.008

Devine, R. T., & Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: A meta-analysis. *Child Development*, *85*(5), 1777–1794. https://doi.org/10.1111/cdev.12237

De Villiers, P. A., & De Villiers, J. G. (2012). Deception dissociates from false belief reasoning in deaf children: Implications for the implicit versus explicit theory of mind distinction. British Journal of *Developmental Psychology, 30*(1), 188–209. https://doi.org/10.1111/j.2044-835X.2011.02072.x.

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

Dunn, J., Cutting, A. L., & Demetriou, H. (2000). Moral sensibility, understanding others, and children's friendship interactions in the preschool period. *British Journal of Developmental Psychology, 18*(2), 159–177. https://doi.org/10.1348/026151000165625

Eisenberg-Berg, N., & Roth, K. (1980). Development of young children's prosocial moral judgment: A longitudinal follow-up. *Developmental Psychology, 16*, 375–376.

Fadda, R., Parisi, M., Ferretti, L., Saba, G., Foscoliano, M., Salvago, A., et al. (2016). Exploring the role of theory of mind in moral judgment: The case of children with autism spectrum disorder. *Frontiers in Psychology, 7*, 523. https://doi.org/10.3389/fpsyg .2016.00523.

Fu, G., Xiao, W. S., Killen, M., & Lee, K. (2014). Moral judgment and its relation to second-order theory of mind. *Developmental Psychology, 50*(8), 2085–2092. https://doi.org/10.1037/a0037077

Garcia-Molina, I., Clemente-Estevan, R.A., Andrés-Roqueta, C., & Rodríguez, L. (2016). Intervención mediante historias complejas de teoría de la mente. Meteduras de pata, juicios morales e ironías. *Psique, 17*. Castelló de la Plana: Universitat Jaume I. https://doi.org/10.6035/Psique.2016.17

Garcia-Molina, I., & Clemente-Estevan, R.-A. (2019). Moral reasoning in autistic individuals: A theoretical review. *Studies in Psychology, 40*(2), 312–337. https://doi.org/10.1080/02109395.2019.1589083

Garcia-Molina, I., & Rodríguez-Clavell, P. (2022). Moral judgements among neurotypical children, autistic children and adults with intellectual disability. *International Journal of Developmental Disabilities*, 1-8. 10.1080/20473869.2022.2156666.

Glidden, J., D'Esterre, A., & Killen, M. (2021). Morally-Relevant Theory of Mind Mediates the Relationship between Group Membership and Moral Judgments. *Cognitive Development*, *57*, 100976. https://doi.org/10.1016/j.cogdev.2020.100976

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*(1), 26–37. https://doi.org/10.2307/1130386

Gönültaş, S., & Mulvey, K. L. (2022). Theory of mind as a correlate of bystanders' reasoning about intergroup bullying of Syrian refugee youth. *Frontiers in Psychology, 13,* 815639. doi: 10.3389/fpsyg.2022.815639. PMID: 35432123; PMCID: PMC9005638.

Grant, C., Boucher, J., Riggs, K., & Grayson, A. (2005). Moral understanding in children with autism. *Autism, 9*, 317–331. https://doi.org/10.1177/1362361305055418

Hamilton A. F. (2009). Goals, intentions and mental states: challenges for theories of autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *50*(8), 881–892. https://doi.org/10.1111/j.1469-7610.2009.02098.x

Hao, J., & Liu, Y. (2016). A mind-reader does not always have deontological moral judgments and prosocial behavior: A Developmental Perspective. *Frontiers in Psychology*, *7*, 1261. https://doi.org/10.3389/fpsyg.2016.01261

Happé, F. G. E. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders, 24*(2), 129–154. https://doi.org/10.1007/BF02172093

Harari, Y., & Weinstock, M. (2021). Interpretive theory of mind and empathic prosocial moral reasoning. *The British Journal of Developmental Psychology*, *39*(1), 78–97. https://doi.org/10.1111/bjdp.12345

Harris, P. L. (1986). Social cognition. In Damon W, Lerner RM, Kuhn D, Siegler R, editors. *Handbook of child psychology: Volume 2: Cognition, perception, and language.*(pp. 811–858). New York: John Wiley & Sons.

Hauser, M., Cushman, F., Young, L., Kang-Xing Jin, R., & Mikhail, J. (2007), A dissociation between moral judgments and justifications. *Mind and Language, 22*, 1-21. https://doi.org/10.1111/j.1468-0017.2006.00297.x

Helwig, C. C., Zelazo, P. D., & Wilson, M. (2001). Children's judgments of psychological harm in normal and noncanonical situations. *Child Development*, *72*(1), 66–81. https://doi.org/10.1111/1467-8624.00266

Hepach, R., Vaish, A., & Tomasello, M. (2012). Young children are intrinsically motivated to see others helped. *Psychological Science*, *23*(9), 967–972. https://doi.org/10.1177/0956797612440571

Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice.* Cambridge University Press. https://doi.org/10.1017/CBO9780511805851

Homer, B. D., & Astington, J. W. (2001). Children's representation of second order beliefs in self and other [Unpublished manuscript]. New York University.

Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Developmental Psychology*, *43*(6), 1447–1459. https://doi.org/10.1037/0012-1649.43.6.1447

Kant, I. (1959). Foundation of the Metaphysics of Morals. Transl. by L. W. Beck.Indianapolis, IN: Bobbs-Merrill.

Kenward, B., & Osth, T. (2012). Enactment of third-party punishment by 4-year-olds. *Frontiers in Psychology*, *3*, 373. doi: 10.3389/fpsyg.2012.00373

Killen, M., & Rutland, A. (2011). *Children and social exclusion: Morality, prejudice, and group identity.* Wiley Blackwell. https://doi.org/10.1002/9781444396317

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

Kochanska, G., Gross, J. N., Lin, M. H., & Nichols, K. E. (2002). Guilt in young children: Development, determinants, and relations with a broader system of standards. *Child Development*, *73*(2), 461–482. https://doi.org/10.1111/1467-8624.00418

Kohlberg, L. (1969). Stage and sequence: The cognitive-development approach to socialization. In D. A. Goslin (Ed.), Handbook of socialization theory and research (pp. 347- 480). Chicago: Rand McNally.

Kuntoro, I., Dwiputri, G., & Adams, P.. (2017). The contribution of parenting style and theory of mind to the understanding of morally relevant theory of mind in Indonesian children. *Diversity in Unity: Perspectives from Psychology and Behavioral Sciences.* https://doi.org/10.1201/9781315225302-11.

Lalonde, C. E., & Chandler, M. J. (2002). Children's understanding of interpretation. *New Ideas in Psychology, 20*(2-3), 163–198. https://doi.org/10.1016/S0732-118X(02)00007-7

Lane, J. D., Wellman, H. M., Olson, S. L., LaBounty, J., & Kerr, D. C. R. (2010). Theory of mind and emotion understanding predict moral development in early childhood. *The British Journal of Developmental Psychology*, *28*(Pt 4), 871–889. https://doi.org/10.1348/026151009x483056

Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology, 50*(1), 45–85. https://doi.org/10.1016/j.cogpsych.2004.06.002

Loureiro, C., & Souza, D. D. (2013). The relationship between theory of mind and moral development in preschool children. *Paidéia, 54*, 93-101. doi: http://dx.doi.org/10.1590/1982-43272354201311

Li, L., Rizzo, M. T., Burkholder, A. R., & Killen, M. (2017). Theory of mind and resource allocation in the context of hidden inequality. *Cognitive Development, 43,* 25–36. https://doi.org/10.1016/j.cogdev.2017.02.001

Malti, T., & Latzko, B. (2010). Children's moral emotions and moral cognition: towards an integrative perspective. *New Directions for Child and Adolescent Development*, *2010*(129), 1–10. https://doi.org/10.1002/cd.272

Margoni, F., & Surian, L. (2016a). Mental state understanding and moral judgment in children with autistic spectrum disorder. *Frontiers in Psychology, 7*, 1478. https://doi.org/10.3389/fpsyg .2016.01478.

Margoni, F., & Surian, L. (2016b). Explaining the U-shaped development of intent-based moral judgments. *Frontiers in Psychology, 7*, 219. https://doi.org/10.3389/fpsyg.2016.00219.

Margoni, F., & Surian, L. (2017). Children's intention-based moral judgments of helping agents. *Cognitive Development*, *41*, 46–64. https://doi.org/10.1016/j.cogdev.2016.12.001

Margoni, F., & Surian, L. (2020). Conceptual continuity in the development of intent-based moral judgment. *Journal of Experimental Child Psychology*, *194*, 104812. https://doi.org/10.1016/j.jecp.2020.104812

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*(2), 622–646. https://doi.org/10.1111/j.1467-8624.2007.01018.x

Moran, J. M., Young, L., Saxe, R., Lee, S. M., O'Young, D., Mavros, P. L., et al. (2011). Impaired theory of mind for moral judgment in high-functioning autism. *Proceedings of the National Academy of Sciences, 108*, 2688–2692. https://doi.org/10.1073/pnas.10117 34108.

Nelson, S. A. (1980). Factors influencing young children's use of motives and outcomes as moral criteria. *Child Development*, *51*(3), 823. https://doi.org/10.2307/1129470

Nelson-Le Gall, S. A. (1985). Motive-outcome matching and outcome foreseeability: Effects on attribution of intentionality in moral judgments. *Developmental Psychology, 21*(2), 332-337. https://doi.org/10.1037/0012-1649.21.2.332

Nobes, G., Panagiotaki, G., & Pawson, C. (2009). The influence of negligence, intention, and outcome on children's moral judgments. *Journal of Experimental Child Psychology, 104,* 382-97. https://doi.org/10.1016/j.jecp.2009.08.001.

Nobes, G., Panagiotaki, G., & Bartholomew, K. J. (2016). The influence of intention, outcome and question-wording on children's and adults' moral judgments. *Cognition*, *157*, 190–204. https://doi.org/10.1016/j.cognition.2016.08.019

Nucci, L. P. (2001). *Education in the moral domain.* Cambridge University Press. https://doi.org/10.1017/CBO9780511605987

Ochoa, K. D., Rodini, J. F., & Moses, L. J. (2022). False belief understanding and moral judgment in young children. *Developmental Psychology*, *58*(11), 2022–2035. https://doi.org/10.1037/dev0001411

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

Perner, J., & Wimmer, H. (1985). "John *thinks* that Mary *thinks* that…": Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology, 39*(3), 437–471. https://doi.org/10.1016/0022-0965(85)90051-7

Peterson, C. C., Wellman, H. M., & Slaughter, V. (2012). The mind behind the message: Advancing theory-of-mind scales for typically developing children, and those with deafness, autism, or Asperger syndrome. *Child Development, 83*, 469–485. https://doi. org/10.1111/j.1467-8624.2011.01728.x.

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

Pons, F., Harris, P. L., & de Rosnay, M. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology*, *1*(2), 127–152. https://doi.org/10.1080/17405620344000022

Rakoczy, H. (2022). Foundations of theory of mind and its development in early childhood. *Nature Reviews Psychology, 1*, 223–235. https://doi.org/10.1038/s44159-022-00037-z

Ross, H. (2006). Theory of Mind Goes Social [Review of *How Children Develop Social Understanding*, by J. Carpendale & C. Lewis]. *Human Development*, *49*(6), 375–379. https://www.jstor.org/stable/26763913

Rogers, J., Viding, E., Blair, J., Frith, U., & Happé, F. (2006). Autism spectrum disorder and psychopathy: Shared cognitive underpinnings or double hit? *Psychological Medicine, 36*, 1789– 1798. https://doi.org/10.1017/s0033291706008853.

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

Seucan. D. T., Szekely-Copîndean, R. D., & Visu-Petra, L. (under review). Preschoolers' moral judgment and punishment tendency relies on their understanding of other people's minds and emotions. *Social Development.*

Shulman, C., Guberman, A., Shiling, N., & Bauminger, N. (2012). Moral and social reasoning in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*, 1364–1376. https:// doi.org/10.1007/s10803-011-1369-8.

Slomkowski, C. L., & Killen, M. (1992). Young children's conceptions of transgressions with friends and nonfriends. *International Journal of Behavioral Development, 15*(2), 247–258. https://doi.org/10.1177/016502549201500205

Smetana, J. G. (1985). Preschool children's conceptions of transgressions: Effects of varying moral and conventional domain-related attributes. *Developmental Psychology, 21*(1), 18–29. https://doi.org/10.1037/0012-1649.21.1.18

Smetana, J. G., & Braeges, J. L. (1990). The development of toddler's moral and conventional judgments. *Merrill-Palmer Quarterly, 36*(3), 329–346.

Smetana, J. (2006). Social domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen, & J. Smetana (Eds.), *Handbook of moral development* (pp. 119-154). Hillsdale, NJ: Erlbaum.

Smetana, J. G. (2011). *Adolescents, families, and social development: How teens construct their worlds.* Wiley-Blackwell.

Smetana, J. G., Jambon, M., Conry-Murray, C., & Sturge-Apple, M. L. (2012). Reciprocal associations between young children's developing moral judgments and theory of mind. *Developmental Psychology*, *48*(4), 1144–1155. https://doi.org/10.1037/a0025891

Smetana, J., Jambon, M., & Ball, C. (2014). The social domain approach to children's moral and social judgments. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (2nd ed., pp. 23–45). New York: Psychology Press.

Smetana, J. G., Jambon, M., & Ball, C. L. (2018). Normative changes and individual differences in early moral judgments: A constructivist developmental perspective. *Human Development*, 1–17. https://doi.org/10.1159/000492803

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*(2), 683–696. https://doi.org/10.1111/j.1467-8624.2011.01714.x

Sodian, B., & Frith, U. (1992). Deception and sabotage in autistic, retarded and normal children. *Child Psychology and Psychiatry and Allied Disciplines, 33*(3), 591–605. https://doi.org/10.1111/j.1469-7610.1992.tb00893.x

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. *New Directions for Child and Adolescent Development*, *103*, 19–36. https://doi.org/10.1002/cd.95

Steele, S., Joseph, R., & Tager-Flusberg, H. (2003). Brief report: Developmental change in theory of mind abilities in children with autism. *Journal of Autism and Developmental Disorders, 33*, 461–467. https://doi.org/10.1023/a:1025075115100.

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology, 30*, 395– 402. doi:10.1037/0012-1649.30.3.395

Surian, L., & Leslie, A. M. (1999). Competence and performance in false belief understanding: A comparison of autistic and normal 3-year-old children. *British Journal of Developmental Psychology, 17*, 141–155. https://doi.org/10.1348/026151099165203.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., Lewin, S., … Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

Turiel, E. (2006). The development of morality. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 789–857). John Wiley & Sons, Inc.

Vaish, A., Missana, M., & Tomasello, M. (2011). Three-year-old children intervene in third-party moral transgressions. *The British Journal of Developmental Psychology*, *29*(1), 124–130. doi: 10.1348/026151010X532888

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

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

Young, L., & Tsoi, L. (2013). When mental states matter, when they don't, and what that means for morality. *Social and Personality Psychology Compass*, *7*(8), 585–604. https://doi.org/10.1111/spc3.12044.

Yuill, N., & Perner, J. (1988). Intentionality and knowledge in children's judgments of actor's responsibility and recipient's emotional reaction. *Developmental Psychology, 24*(3), 358–365. https://doi.org/10.1037/0012-1649.24.3.358

Zalla, T., Barlassina, L., Buon, M., & Leboyer, M. (2011). Moral judgment in adults with autism spectrum disorders. *Cognition, 121*, 115–126. https://doi.org/10.1016/j.cognition.2011.06.004

Zalla, T., & Leboyer, M. (2011). Judgment of intentionality and moral evaluation in individuals with high functioning autism. *Review of Philosophy and Psychology, 2*, 681–698. https://doi.org/10.1007/ s13164-011-0048-1.

Zelazo, P. D., Helwig, C. C., & Lau, A. (1996). Intention, act, and outcome in behavioral prediction and moral judgment. *Child Development*, *67*(5), 2478. https://doi.org/10.2307/1131635