

The ontogeny of intent-based normative judgments

Marina Proft | Hannes Rakoczy

Department of Developmental Psychology, University of Göttingen, Göttingen, Germany

Correspondence

Marina Proft, Georg-August-Universität Göttingen, 37073 Göttingen, Germany.
Email: marina.proft@psych.uni-goettingen.de

Abstract

When evaluating norm transgressions, children begin to show some sensitivity to the agent's intentionality around preschool age. However, the specific developmental trajectories of different forms of such intent-based judgments and their cognitive underpinnings are still largely unclear. The current studies, therefore, systematically investigated the development of intent-based normative judgments as a function of two crucial factors: (a) the type of the agent's mental state underlying a normative transgression, and (b) the type of norm transgressed (moral versus conventional). In Study 1, 5- and 7-year-old children as well as adults were presented with vignettes in which an agent transgressed either a moral or a conventional norm. Crucially, she did so either intentionally, accidentally (not intentionally at all) or unknowingly (intentionally, yet based on a false belief regarding the outcome). The results revealed two asymmetries in children's intent-based judgments. First, all age groups showed greater sensitivity to mental state information for moral compared to conventional transgressions. Second, children's (but not adults') normative judgments were more sensitive to the agent's intention than to her belief. Two subsequent studies investigated this asymmetry in children more closely and found evidence that it is based on performance factors: children are able in principle to take into account an agent's false belief in much the same way as her intentions, yet do not make belief-based judgments in many existing tasks (like that of Study 1) due to their inferential complexity. Taken together, these findings contribute to a more systematic understanding of the development of intent-based normative judgment.

KEYWORDS

intentionality, moral development, normativity, theory of mind

1 | INTRODUCTION

For adults, normative judgments are largely a function of the agent's will behind an action rather than merely of its outcome. When confronted with scenarios involving norm-violating actions they generally base their judgments more on the agents' underlying mental states, i.e., their beliefs, desires, and intentions that motivated the action, than the outcome of the action (Cushman, 2008; Guglielmo, Monroe, & Malle, 2009). For example, when confronted with two agents who bring about similar outcomes, adults judge a knowing, intentional transgressor as more blameworthy than an unknowing, unintentional one (e.g., Young, Cushman, Hauser, & Saxe, 2007).

Furthermore, when confronted with two agents who have the same motivational set, one of whom is "unlucky" and thus causes a negative outcome, adults again mostly blame the agent for negligence rather than the negative outcome per se (Young, Nichols, & Saxe, 2010). Consequently, for adults *why* something went wrong is a crucial basis for normative evaluations.

In contrast to these clear findings with adults, the developmental picture of such intent-based judgments is less clear. Pioneer work by Piaget (1932) suggested that children's moral judgments remain limited to outcome-based evaluations until middle childhood. Numerous studies since have aimed at investigating the emergence of intent-based judgments more systematically and straightforward

(Cushman, Sheketoff, Wharton, & Carey, 2013; Killen, Mulvey, Richardson, Jampol, & Woodward, 2011; Zelazo, Helwig, & Lau, 1996). The general emerging picture suggests that intent-based judgment clearly occurs at an earlier age than proposed by Piaget. That said, however, the specific findings still diverge massively and remain difficult to interpret. Some studies suggest that even 3-year-olds are capable of engaging in explicit intent-based judgments (e.g., Nobes, Panagiotaki, & Pawson, 2009; Vaish, Carpenter, & Tomasello, 2010). Other studies, in contrast, consistently fail to find such tendencies in even 5- and 6-year-old children (e.g., Helwig, Zelazo, & Wilson, 2001; Samland, Josephs, Waldmann, & Rakoczy, 2016).

When then does intent-based normative judgment develop? How can this complex picture of inconsistent findings be explained? One basic problem actually lies in the broad use of the term *intent-based*. Most commonly, it has been used as an umbrella term for all the situations where people do not only base their judgments on the outcome of the transgression but to some extent consider the agent's underlying motivation as well. Such a broad use, however, obscures the variability of unintentional transgressions. This can be most clearly illustrated by the difference between two groups of paradigmatic cases that have been used to elicit intent-based judgments.

The first group of cases includes situations in which the action that leads to the transgression was performed unintentionally. Accidents are a paradigmatic instance: for example, when someone stumbles and thereby accidentally destroys another person's valuable object. The second group of cases includes situations in which the action that leads to the transgression was performed intentionally, but not under the description of the outcome. Hence, while the action was intentional the outcome was not intended. Mistakes based on false beliefs are the paradigmatic instance: for example, when someone intentionally destroys another person's valuable object, though under the false assumption that it was a piece of trash.

Most importantly, due to their difference in underlying intentional structures (unintentional action in the first case vs. unintended outcome in the second case) the evaluations of the two cases clearly differ in complexity. For the case of the stumbling person, the information about the unintentionality of the action can directly feed into the normative evaluation. Only one reasoning step is required, i.e., from the unintentionality of the action to its normative status. For the case of a false-belief-based transgression, the fact that the action was performed intentionally requires an extra reasoning step on the part of the evaluator. That is, the evaluator has to first infer that bringing about the outcome under the description "destroying the valuable object" was not intended by keeping in mind the transgressor's epistemic state. Only then can this information feed into the normative evaluation. Thus, the inferential chain here is more complex and involves two reasoning steps: (a) from the false belief to the unintended outcome and (b) from the unintended outcome to the normative status of the action.

Given these differences in reasoning chains, from a developmental point of view, this licenses the expectation that intent-based judgment emerges earlier for accidental mistakes than for false-belief-based ones. Existing findings, in fact, seem to be

Highlights

- The development of intent-based normative judgments was investigated as a function of the type of norm violated and the agents' underlying mental states
- Results revealed two independent asymmetries in the degree of intent-based normative judgments
- All age groups included mental states more in moral than in conventional norms
- Children's (but not adults') normative judgments were more sensitive to the agent's intention than to her belief

compatible with such a pattern: Studies tapping children's sensitivity for unintentional actions tend to reveal intent-based judgments from around 3–4 years of age (e.g., Nelson, 1980; Núñez & Harris, 1998). For example, 3- to 4-year-olds (under certain conditions) have been found to consider the intentions of an actor as the most important factor when explicitly evaluating the permissibility of actions (Nobes et al., 2009). Moreover, some recent studies suggest that even infants and toddlers, in their early socio-moral evaluations, may distinguish between agents on the basis of their good and bad intentions (see, e.g., Choi & Luo, 2015; Dunfield & Kuhlmeier, 2010; Hamlin, 2013).

Studies tapping children's sensitivity for false-belief-based mistakes, in contrast, tend to indicate a later developing competence around 7 years (Helwig et al., 2001; Killen et al., 2011; Wang, Zhu, & Shi, 2011). For example, in a scenario in which an agent frightened another agent, only children from age 7 judged the action as more acceptable and less deserving of punishment when it was based on a false belief than the same action performed knowingly (Helwig et al., 2001). In addition, this basic developmental pattern of intention-based judgments preceding belief-based ones is also consistent with findings from the only study to date that has directly contrasted the agent's motives with the foreseeability of the outcome of the action (Yuill & Perner, 1988). In this study with 3- to 7-year-olds, younger children only distinguished between unforeseen accidents and intended outcomes, and only the older children also used foreseeability and thus the agent's subjective beliefs alone as a basis for normative judgment and blame.

Existing findings are thus compatible with a general asymmetry in the development of intent-based normative judgment such that intention-based judgments precede belief-based ones in development. However, whether such an asymmetry really exists and best explains developmental patterns, is still an outstanding question that needs to be systematically addressed—which is the first research aim of the present study.

If this asymmetry indeed holds, the next obvious question concerns its source. Does it reflect a difference in conceptual competence or merely a difference in external performance demands? In particular, do children's difficulties with belief-based evaluations reflect a competence limitation such that they lack the requisite

conceptual competence in principle? Or does it reflect a performance limitation such that children do have this conceptual competence but may have difficulties applying it due to extraneous task demands?

Given that belief-based judgments require a two-step inferential reasoning chain (belief→intention→normative status), the conceptual competence that is asked for is the ability to perform and combine these two reasoning steps. Conversely, a corresponding conceptual competence deficit may take different forms: first, children may have difficulties understanding the underlying intentional structure of false-belief-based actions and thus do not understand that, though performed intentionally, the action was not intentional under the description of the outcome. Second, children may have difficulties seeing a false belief as a mitigating factor and thus, though they understand that the outcome was not intended, do not see the false belief as an excuse for the immoral act.

In contrast, if children were generally able to perform both inferences when explicitly requested, but fail to do so spontaneously, this would speak for a mere performance limitation. In this case, children's outcome-based judgments most probably result from the complexity of the inferential chain itself and performance is limited by extraneous task demands rather than conceptual deficits.

In order to test these two possibilities (competence or performance limitation) against each other, children's ability to perform and combine the two required reasoning steps under various conditions need to be investigated.

The third research questions concerns not so much the *why* something went wrong (i.e., the underlying intentionality), but *what* went wrong, that is, the type of norm that was transgressed. Previous work has mainly focused on moral norm transgressions, especially moral harm norms. Our everyday norm psychology, however, involves different types of norms, each with specific characteristics. Take for instance, the distinction most commonly drawn in philosophy and psychology: moral versus conventional norms. There is ample developmental evidence that already preschoolers distinguish between the prototypical cases of the two types of norms: cognitively, they judge moral harm violations as more rule independent, context-insensitive and severe than paradigmatic social-conventional violations such as wearing pajamas to school (Nucci & Nucci, 1982; Smetana, 1981; Smetana & Braeges, 1990; Smetana et al., 2012). And emotionally, they show different characteristic reactions to these moral versus conventional norm transgressions (Hardecker, Schmidt, Roden, & Tomasello, 2016).

How does this distinction relate to the current research focus? From a theoretical point of view, the specific characteristics of the two types of norms might well have an impact on respective intent-based judgments. More specifically, we propose that the different reactive attitudes that come along with prototypical moral versus prototypical conventional norm transgressions influence the degree to which the agent's intentions are considered in normative evaluations. Moral norm transgressions elicit responses such as guilt and blame (Haidt, 2007; Nichols, 2002). Accordingly, the form of critique involved in moral transgressions is essentially related to the

blameworthiness of the agent. To evaluate whether someone is to be blamed, the key element is whether the agent had a guilty mind: was her will good or bad? Thus, when evaluating moral transgressions, the focus is on the agent herself and her underlying intentions. Her will is what counts. However, this might not necessarily be so for prototypical conventional norm transgressions. Consider, for example, a situation in which someone makes an inappropriate move in chess (say, moving the king two squares), a transgression of a social-conventional norm. Now this move gives the observer sufficient and good reason for normative critique ("No, one can only move the king one square!"). And importantly, this form of critique is equally justified when the agent in fact did not know the rules of chess. The reason lies in the form of critique involved here. It is not based on the blameworthiness of the agent but rather it expresses an evaluation, simply, of whether the move was correct or incorrect, i.e., legal or illegal according to the rules of chess. As a consequence, the focus shifts away from the agent to the outcome of the transgression. Therefore, the judgment of a conventional mistake should be less dependent on the underlying mental states that led to the transgression.

On this basis, a second asymmetry emerges, the degree of intent-based judgments should vary as a function of the kind of norm transgressed. Regarding the moral/conventional contrast specifically, there should be more intent-based judgments for prototypical moral harm than prototypical conventional violations.¹

Empirically, some recent studies suggest that these considerations are largely reflected in adult intuitions. For instance, within the area of moral norms, subjects tend to focus more on the agent's intent for harm norms than for purity or disgust norms (Russell & Giner-Sorolla, 2011; Young & Saxe, 2011). And for moral versus conventional norms more specifically, adults make stronger intent-based judgments for prototypical moral than for prototypical conventional norms (Giffin & Lombrozo, 2016, 2018).

From a developmental point of view, however, little is known about how these intuitions develop. So far, there is only one study that compared intent-based moral and conventional judgments in preschoolers (Josephs, Kushnir, Gräfenhain, & Rakoczy, 2016). This study with 3- and 4-year-olds used the child's protest as an indicator for normative understanding and found adult-like patterns in children's normative evaluations. When the agent was perceived as externally constrained and thus unfree in her choices, children refrained from criticizing the agent for inappropriate acts in moral situations, but still criticized conventional mistakes conducted under similarly constrained circumstances. So far, no study has tested for the differential influence of mental state information on children's explicit judgments of moral versus conventional transgressions.

Against the background of these open questions and issues, the aim of the present set of studies was to investigate the development of intent-based normative judgments more systematically. Our main interest were the two proposed asymmetries. First, concerning the type of mental state that led to the transgression, we expect more intent-based judgments for accidental than for false-belief-based mistakes. Second, concerning the type of norm transgressed, we

expect more intent-based judgments for moral than for conventional violations. Crucially, as we propose both asymmetries to derive from different cognitive foundations, we expect both factors to independently influence the degree of intent-based judgments.

To this aim, in a first step, we explored the two asymmetries in a systematic research design across different age groups (5-year-old and 7-year-old children and adults). In particular, we tested intent-based judgments as a function of (a) the type of unintentional transgression (accidents versus false-belief-based mistakes) and (b) the type of norm transgressed (moral versus conventional). In the two children's age groups, we found both predicted asymmetries: children made more intent-based judgments for accidents than for false-belief-based mistakes as well as for moral than for conventional transgressions. Adults' intent-based judgments, in contrast, only differed across the two types of norms (moral versus conventional) but not across the two types of mental states. In a second step, we thus explored the cognitive foundation of the mental state asymmetry more carefully, as it seems to undergo developmental changes into adulthood. Our main focus was to distinguish between competence and performance factors: does the asymmetry result from specific moral reasoning processes or from more general cognitive abilities?

The rationale for this approach was inspired by recent research in a different area: children's trait ascription (see Liu, Gelman, & Wellman, 2007). Much traditional research had shown that children until surprisingly late (7–8 years) fail tasks in which they have to predict an agent's future behavior (e.g., will she help her friends to tidy their room?) from her past and present behavior (helping old women carry bags, helping strangers across the road, etc.; see e.g., Rholes & Ruble, 1984). Traditionally, these findings were taken as evidence for a conceptual competence limitation to the effect that young children do not yet operate with trait concepts ("is helpful"; e.g., Heyman, 2009). Recent work, however, suggests that these findings may merely reflect performance limitations due to the inferential complexity of the task (Liu et al., 2007). In such tasks, subjects have to engage in a two-step inferential chain: (a) behavior (helps in situation A, B...) → trait (helpful); (b) trait (helpful) → future behavior (helps in novel situation C...). Liu et al. (2007) now showed that children were able to make each of the two inferences, and both together when the task was suitably simplified, but had a hard time making the two-step chain spontaneously. These findings thus clearly support a performance (rather than a competence) limitation account.

The present case, intent-based judgments of mistakes based on false beliefs, presents an analogous logical structure, with a two-step inferential chain: (a) false belief → unintended outcome; (b) unintended outcome → normative status. According to performance limitation accounts, children may well be able to make each inference, and to put them together in suitably simplified tasks, but fail to do so spontaneously. On this basis, inspired by the Liu et al. (2007) reasoning and approach, in a second step, we set out to test competence and performance limitation accounts in Studies 2a and 2b. Specifically, we tested for children's ability to operate each of these two reasoning steps separately and in combination, when explicitly asked to make the respective inferences. If children are capable of performing each inferential step

((a) false belief → unintended outcome and (b) unintended outcome → normative status) in isolation, as well as the two together in a sequence in suitably simplified tasks, this suggests they have the competence for belief-based normative judgment in principle. Their failure to apply this competence in the tasks of Study 1 would then reflect mere performance limitations.

2 | STUDY 1

In the first study, we asked 5- and 7-year-old children and adults to give explicit normative judgments in response to prototypical moral and conventional violations. In the moral scenarios, harm was done to a third person, whereas in the conventional scenarios an arbitrary rule was transgressed. In each story, two actors violated the same given norm, one of whom intended to do so while the other one did not intend the norm transgression—either because she did not act intentionally at all (unintentional action) or because she did act intentionally, yet based on a false belief regarding the outcome of her action (false belief; see Table 1). For example, in one moral norm story, Grandma's flowers were always destroyed by the use of the wrong fertilizer, either intentionally by Lisa versus unintentionally by Kathrin (who tripped, knocking down the wrong fertilizer from the shelf) or intentionally by Lisa versus unknowingly by Maja (who had a false belief about the content of one of the boxes with fertilizer). Crucially, by using the same scenarios, the transgressions were completely matched regarding all relevant factors such as the outcome of the action. As a consequence, a more negative evaluation of the intended than the unintended transgression is a clear indicator of intent-based judgments—as the only difference between the cases is the intentional structure of the action.

We included the adult population to compare children's performance against the "mature" adult intuition regarding both the unintentional versus false-belief-based mistake as well as the moral versus conventional norm contrasts. We tested 5- and 7-year-old children since previous research has shown that from age 5, children generally show a robust understanding of false beliefs. Such a general understanding of false beliefs is a prerequisite for making normative judgments on the basis of the beliefs ascribed to an agent. To ensure all children did in fact operate with an understanding of false beliefs, and to rule out that potential difficulties in belief-based normative judgment were due to delayed Theory of Mind, we administered standard false belief tasks to all children.

2.1 | Method

2.1.1 | Participants

The final sample consisted of 70 5-year-old children (60–71 months, $M = 64.97$, 34 girls), 70 7-year-old children (84–95 months, $M = 87.86$, 33 girls), and 48 adults (19–29 years, $M = 22.71$, 34 women). One additional 5-year-old child was tested but could not be coded due to a camera error.

TABLE 1 Schematic overview of the structure of two exemplary stories (moral flower and conventional picture)

Condition				Moral Flower story	Conventional Picture story
Main story				Grandma's beautiful, beloved flowers always need to get the yellow fertilizer from the shelf (the blue fertilizer would destroy them and thus make Grandma very sad).	There is a picture of an elephant lying on the table. The rule for the picture says that it is only allowed to be painted in blue.
Transgressor	Outcome	Action	Appearance in contrast		
Intended Outcome Transgressor	Intended	Intentional	Intention contrast & False Belief contrast	Lisa, who announces that she "wants to fertilize Grandma's flowers with the blue fertilizer!" and then fulfills her plan which leads to the death of Grandma's flowers.	Lena, who announces that she "wants to paint the elephant in green!" and fulfills her plan by putting green paint on the picture.
Unintentional Action Transgressor	Unintended	Unintentional	Intention contrast	Kathrin, who announces that she "wants to fertilize Grandma's flowers with the yellow fertilizer!" but trips on her way to the shelf and bumps against it. In doing so a box with blue fertilizer falls down on Grandma's flowers which consequently die.	Jenny, who announces that she "wants to paint the elephant in blue!" but trips on her way to the table and bumps against it. In doing so a can with green paint tilts over, directly on the elephant.
False Belief Transgressor	Unintended	Intentional	False Belief contrast	Maja, who has a false belief about the content of one of the yellow boxes with fertilizer (which in fact contains blue fertilizer). She announces that she "wants to fertilize Grandma's flowers with the yellow fertilizer!" and unknowingly takes the false content box to pour the fertilizer over the flowers. Grandma's flowers die.	Melanie, who has a false belief about the content of one of the blue cans of painting colors (which in fact has green paint in it). She announces that she "wants to paint the elephant in blue!" and unknowingly takes the false content can to put paint on the picture. The elephant gets green.

Children were from mixed socioeconomic backgrounds and recruited from a local database of parents, who had volunteered to participate in developmental studies. Adult participants were recruited on the campus of the University of Göttingen.

2.1.2 | Design

The present study followed a 2 (type of norm: moral versus conventional) × 2 (kind of contrast: intention versus false belief) × 3 (age group: 5-year-old and 7-year-old children, and adults) design with type of norm as a within-subject factor and kind of contrast and age as between-subjects factors.

Each participant received two moral and two conventional stories in blocks. In each story, they were presented with one contrast pair of norm transgressors—consisting of one character who intended the norm transgression and one character who did not intend the norm transgression—and were asked to evaluate each of the two transgressors separately. Depending on the condition, the contrast consisted of the intended outcome and the unintentional action

transgressor (intention contrast) or of the intended outcome and the false belief transgressor (false belief contrast, see Table 1). The differences in evaluation between the two transgressors within a contrast served as the main dependent variable. The order of the type of norm, the stories within a norm as well as the order of presentation of the two transgressors were counterbalanced across participants. Additionally, two false belief tasks were given to children to assess their Theory of Mind competence.

2.1.3 | Stories

In total, four different stories, two moral and two conventional ones, were used. Each story was accompanied by seven or eight colorful drawings, depending on contrast (seven in the intention contrast and eight in the false belief contrast; see the Appendix for a full description of the stories).

In the moral stories, a third person was harmed as a consequence of the transgression. In the flower version, Grandma's beloved yellow flowers were destroyed by the use of the wrong

fertilizer. In the cupcake version, Caro's self-made cupcake was destroyed.

In the conventional stories, the transgression violated an arbitrary rule. In the picture version, the rule stated that the picture was only supposed to be painted in blue, the mistake was to put green paint on it. In the aquarium version, the rule stated that only red fish were allowed in the aquarium, the mistake was to put green fish in it.

In each story, we contrasted one transgressor who intended the outcome with one who did not intend the outcome. Thus, each story consisted of three elements: Main story, Intended Outcome Transgressor and Unintended Outcome Transgressor (intention or false belief). The structure/elements of the stories and their appearance in the respective contrasts are illustrated in Table 1 for the moral flower story and the conventional picture story.

After the presentation of each character, participants answered three test questions in counterbalanced order: (a) Did XY just do something wrong? (b) Was XY just nice or mean? (c) Should XY get punished? (adapted from Cushman et al., 2013). Children answered with the help of a 4-point smiley scale (trained in the warm-up), adults were given a simple 4-point scale. Answers could range from (a) "completely right–a little right–a little wrong–completely wrong" (b) "really nice–a little nice–a little mean–really mean" and (c) "no punishment–small punishment–medium punishment–big punishment." After each story (i.e., after presenting both transgressors) children had to give one smiley face sticker to either of the two transgressors as an indirect reward/punishment measurement (see Vaish et al., 2010 for a similar method).

2.1.4 | Theory of Mind

Theory of Mind (ToM) competence was assessed with two standard false belief tasks, unexpected contents (Perner, Leekam, & Wimmer, 1987) and location change (Wimmer & Perner, 1983).

In the unexpected contents task, children were presented with a familiar box (a Smarties box) and asked what they thought was inside (the typical answer was "candy," "chocolate" or "Smarties"). Thereafter, the experimenter showed the child that there was actually a pen in the box. After she put back the pen she asked the child two test questions (a) "When I first showed you the box what did you think was inside the box?" and (b) "Your mum hasn't seen this box yet, right? If we now show it to her and ask her, what will she say in the box?". The test questions were followed by a control question ("What is really in the box?").

In the location change task, children were presented with a girl who put her car in a blue box. In her absence, her dad moved the car into an orange box. Children were then asked three control questions "Where did the girl put the car first?", "Where is it now?" followed by "Who put it there?". Finally, children were asked the standard first-order test question: "When the girl returns, where will she look for the car?" and, if they gave a correct answer to the 1st-order question, a 2nd-order question: "If we ask her 'Do you know where the car is?', what will she say? 'Yes, I know where it is' or 'No, I don't know where it is?'" (Perner & Howes, 1992).

2.1.5 | Procedure

Children were tested individually in the laboratory. Following a warm-up (including the smiley-scale training), the two story blocks (moral or conventional) were presented. In between blocks (i.e., after two stories), the ToM tasks were administered. Adults were tested on campus with a written paper–pencil version of the stories. The adult version only covered the stories and test questions (without sticker distribution or ToM tasks).

2.1.6 | Coding

The coding for the test questions followed the logic of the 4-point scale with 0 being "completely right/really nice/no punishment" and 3 being "completely wrong/really mean/big punishment." Since answers to the three test questions were highly correlated (all $r > 0.38$ and $p < 0.01$) we aggregated them in a sum score "evaluation of transgression" for each transgressor (range 0–9). In a second step, we then computed the crucial difference scores for each story by subtracting the score "evaluation of transgression" for the unintended outcome transgressor from the score of the intended outcome transgressor. These scores could range from –9 to 9. Crucially, a positive difference score indicates that the participant made an intent-based judgment: she judged the intended outcome transgressor more harshly than the unintended outcome transgressor (and hence included the agent's mental states). A score of zero or a negative difference score, in contrast, indicated that the participant did not form an intent-based judgment: she judged the unintended outcome transgressor equal or even worse than the intended outcome transgressor (and thus did not consider the agent's mental states at all). The means of these difference scores across the different trials served as the main dependent variable.

Note that three adults and one child did not give valid answers to one of the punishment questions. The child did not give any answer; the adults marked two conflicting answers (e.g., "no punishment" and "small punishment"). These cases were treated as missing values in computing the mean "evaluation of transgression" scores.

For the sticker allocation, a sum score for the amount of stickers given to the unintended outcome transgressor was computed (range 0–2/0–4 for difference between norms/mental states).

2.2 | Results

In this section (and in all following result sections), we report the main findings of interest for our hypotheses. The additional data and analyses concerning children's Theory of Mind understanding can be found in the supplementary material (see Table S1).

The mean difference scores as a function of type of norm and age group and as a function of mental states and age group are depicted in Figures 1 and 2, respectively.

First, we conducted a 2 (type of norm: moral vs. conventional) \times 2 (type of contrast: intention vs. false belief) \times 3 (age group: 5-year-olds vs. 7-year-olds vs. adults) ANOVA with repeated measures for type of norm (see Table 2). Recall that we were mainly

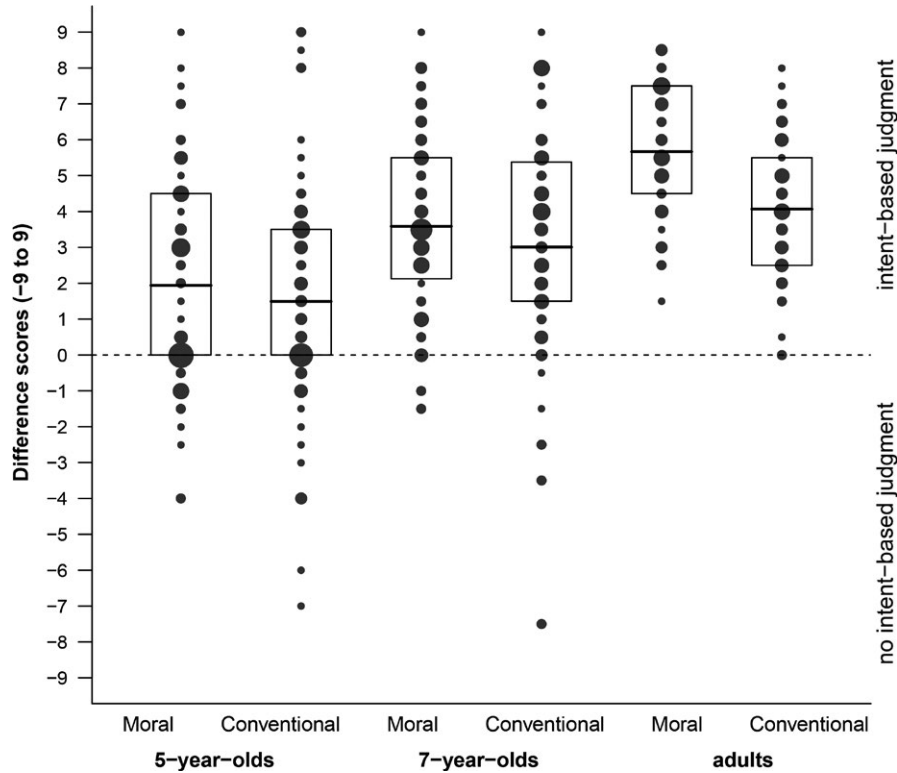


FIGURE 1 Mean difference score (-9 to 9) between the intended and unintended outcome transgressor as as function of the type of norm for the three age groups. The points depict the raw data for each participant with the size of the points indicating sample size (i.e., the number of identical observations) and as boxes the quartiles as well as the mean per condition. Note. A positive difference score indicates intent-based judgments, while a negative score or a score of zero indicates no consideration of the agent's mental states

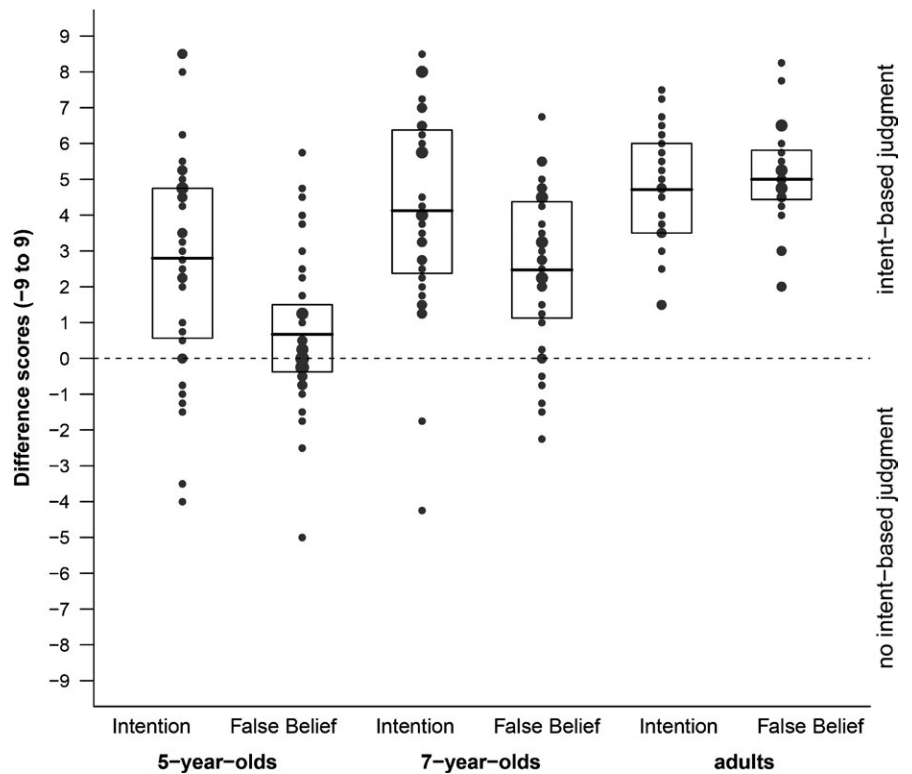


FIGURE 2 Mean difference score (-9 to 9) between the intended and unintended outcome transgressor as function of the type of contrast for the three age groups. The points depict the raw data for each participant with the size of the points indicating sample size (i.e., the number of identical observations) and as boxes the quartiles as well as the mean per contrast. Note. A positive difference score indicates intent-based judgments, while a negative score or a score of zero indicates no consideration of the agent's mental states

TABLE 2 Output from the main ANOVAs on the difference scores between intended and unintended outcome transgressors

	2 × 2 × 3 ANOVA (all ages)				2 × 2 ANOVA (only adults)				2 × 2 × 2 ANOVA (only children)			
	F	df	p	η_p^2	F	df	p	η_p^2	F	df	p	η_p^2
Type of norm	24.510	1,178	<0.001	0.121	30.002	1,43	<0.001	0.411	6.066	1,135	0.015	0.043
Type of contrast	10.146	1,178	0.002	0.054	0.340	1,43	0.563	0.008	17.987	1,135	<0.001	0.118
Age group	23.023	1,178	<0.001	0.206					12.359	1,135	0.001	0.084
Type of norm*Type of contrast	1.605	1,178	0.207	0.009	0.668	1,43	0.418	0.015	1.081	1,135	0.300	0.008
Type of norm*Age group	3.665	2,178	0.028	0.040					0.080	1,135	0.778	0.001
Type of contrast*Age group	3.568	2,178	0.030	0.039					0.282	1,135	0.596	0.002
Type of norm*Type of contrast*Age group	3.062	2,178	0.049	0.033					5.576	1,135	0.020	0.040

interested in the effects for type of norm and for type of contrast as well as the relation between the two factors. As predicted, results revealed both predicted main effects (difference scores moral > difference scores conventional and difference scores intention > difference scores false belief) but no interaction between the two. However, all effects of interest were influenced by age group: there were interactions between type of norm and age group, type of contrast and age group as well as a significant 3-way interaction between all three factors. Furthermore, we found a main effect for age group.

As a consequence, in a second step, we then conducted separate analyses for adults and children to further inspect the effects of interests in the two groups (see Table 2). In the adult group, the 2 (type of norm: moral vs. conventional) × 2 (type of contrast: intention vs. false belief) ANOVA only revealed a main effect for type of norm (difference scores moral > difference scores conventional) but no main effect for type of contrast anymore and again no interaction between the two factors.

For children, the 2 (type of norm: moral vs. conventional) × 2 (type of contrast: intention vs. false belief) × 2 (age group: 5- vs. 7-year-olds) ANOVA with the difference score as DV again yielded in all predicted effects: a main effect for type of norm (difference scores moral > difference scores conventional), a main effect for type of contrast (difference scores intention > difference scores false belief) but no interaction between the two factors. There was, however, again a significant 3-way interaction between type of norm, type of contrast and age group. Further inspection of this interaction revealed a difference between 5- and 7-year-olds, in that there was no interaction between norm and type of contrast for the 7-year-olds ($F(1, 68) = 0.72, p = 0.398, \eta_p^2 = 0.011$) but for the 5-year-olds ($F(1, 67) = 7.36, p = 0.008, \eta_p^2 = 0.099$). The interaction in 5-year-olds was specified by a difference between the two types of contrasts: 5-year-olds only made stronger intent-based judgments for moral than for conventional norms for the intention but not for the false belief contrast (intention contrast: $t(33) = 3.30, p = 0.002, r = 0.50$; false belief contrast: $t(34) = 0.661,$

$p = 0.513, r = 0.11$). This was due to the fact that 5-year-olds did not reliably distinguish between the intended and the unintended outcome transgressors in the false belief contrast at all, irrespective of the type of norm in question ($M = 0.67, SD = 2.16, t(34) = 1.84, p = 0.08, r = 0.30$). Additionally, the ANOVA revealed an overall main effect for age (difference scores 7-year-olds > difference scores 5-year-olds).

Children's sticker distribution mainly resembled their explicit ratings: children gave significantly more stickers to the unintended outcome transgressor in the intention contrast ($M = 3.39, SD = 0.11$) than in the false belief contrast ($M = 2.86, SD = 1.23, t(138) = 2.67, p = 0.008, r = 0.22$). Concerning the two types of norms children also tended to give more stickers to the moral unintended outcome transgressor ($M = 1.61, SD = 0.61$) in contrast to the conventional one ($M = 1.51, SD = 0.77, t(139) = 1.79, p = 0.075, r = 0.15$).

2.3 | Discussion

Study 1 investigated the development of intent-based normative judgments as a function of the type of the agent's mental state (false belief versus intention) and the type of norm transgressed (moral versus conventional). Results revealed that as soon as children started to make intent-based judgments for a certain mental state, both factors independently influenced normative evaluations. First, children (but not adults) made stronger intent-based judgments for mistakes that happened due to a lack of intentionality than for mistakes that happened because of a false belief. Second, across all ages, participants engaged more in intent-based judgments concerning moral than concerning conventional norm transgressions.

What this study shows is that factors that are unimportant for adults' intent-based judgments might well have an impact on children's judgments. While adults treated both types of unintentional transgressions equivalently, for children there seems to be a consecutive development of intention- before belief-based normative judgments. Thus, the two forms should be differentiated more clearly when it comes to children's normative evaluations rather

than merely categorized under one single broad term. What this study still leaves open is why the type of mental state leading to the unintentional transgression matters for children. Recall that for a belief-based normative judgment children need to make two inferences: (a) a false belief \rightarrow unintended outcome inference and (b) an unintended outcome \rightarrow normative status inference. In the following study, we aimed at examining whether children are capable of the two component processes. Do they fail to perform either of the two required reasoning steps and thus lack the conceptual competence to make belief-based judgments? Or does their focus on outcomes for belief-based mistakes reflect difficulties in spontaneously combining the two steps and thus mere performance limitations? To assess more directly which of the two options (competence vs. performance limitation) best explains the asymmetry, we conducted two further experiments.

3 | STUDY 2

The aim of the second set of experiments was to disentangle competence versus performance issues as the cognitive foundations of the mental state asymmetry. To this end, we tested 5-year-olds' ability to perform each of the two reasoning steps in isolation. In both studies, we focused on the 5-year-olds, the younger age group of Study 1. The underlying logic was to see whether even the younger children, under suitable circumstances, were capable of making the requisite inferences. And we focused on moral norm violations, since they were the ones to elicit a stronger focus on the underlying mental states in Study 1.

In Study 2a, we assessed 5-year-olds' ability to perform the first reasoning step: do they make the inference that a false-belief-based action leads to an unintended outcome? To this aim we presented children with agents who either intentionally or unknowingly (based on a false belief) transgressed a moral norm. However, instead of asking them to make a normative evaluation, children were asked whether the agent intentionally brought about the outcome (e.g., "Did Maja intentionally pour blue fertilizer over the flowers?").

In Study 2b, we assessed 5-year-olds' ability to perform the second reasoning step: when primed for the intentional structure of the action, do children include this information in their normative evaluation? To this aim children were tested in three conditions. In addition to the intention contrast and false belief contrast from Study 1, we administered a primed false belief contrast. In this condition children were presented with the same stories as in the original false belief contrast. However, before they were asked to make the normative evaluation, they were primed for the intentional structure of the action by asking whether the agent intentionally brought about the outcome as in Study 2a. In doing so, we guided them through the first reasoning step (false belief \rightarrow unintended outcome) so that they only had to perform the second reasoning step (unintended outcome \rightarrow normative status) on their own. The logic behind this was the following: if children generally do not see a false belief as a mitigating factor, the priming for the intentional structure should not influence

their normative evaluations. If, however, they only do not combine both reasoning steps spontaneously, helping them to perform the first one should enhance intent-based normative judgments.

4. | STUDY 2A

4.1 | Methods

4.1.1 | Participants

Thirty-two 5-year-olds (60–71 months, $M = 65.09$, 14 girls) participated in this study. Children were from mixed socioeconomic backgrounds and recruited from the same local database of parents as in Study 1. Children who had participated in Study 1 were excluded from recruitment of Study 2a.

4.1.2 | Design and stories

As material, the false belief contrasts of the two moral stories (Moral Flower and Moral Cupcake) from Study 1 were used. Thus, each child received two moral stories, in each story they were presented with two norm transgressors, one who intended the outcome (Intended Outcome Transgressor) and one who did not intend the outcome as she had a false belief (False Belief Transgressor). The order of the stories as well as the order of the presentation of the two transgressors was counterbalanced across participants. After the presentation of each character, participants were asked one test question: "Did XY intentionally do that?" For example, in the Moral Flower story after the presentation of the Intended Outcome Transgressor the question was "Did Lisa intentionally pour the blue fertilizer over the flowers?". As in Study 1, two false belief tasks were administered to control for children's Theory of Mind competence (unexpected contents and location change).

4.1.3 | Procedure

Children were tested individually in the lab. Following a warm-up, the two stories were presented. In between stories, the ToM tasks were administered.

4.1.4 | Coding

The number of "yes"-Answers to the test questions were summed across the two stories for each character (Intended Outcome and False Belief Transgressor), resulting in a score between 0–2.

4.2 | Results

Children rated the actions of the Intended Outcome Transgressor as more intentional ($M = 1.31$, $SD = 0.86$) than those of the False Belief Transgressor ($M = 0.66$, $SD = 0.87$, $t(31) = 4.49$, $p < 0.001$, $r = 0.63$, see Figure 3). Closer inspection of the intention ratings revealed that children's performance significantly differed from chance for both

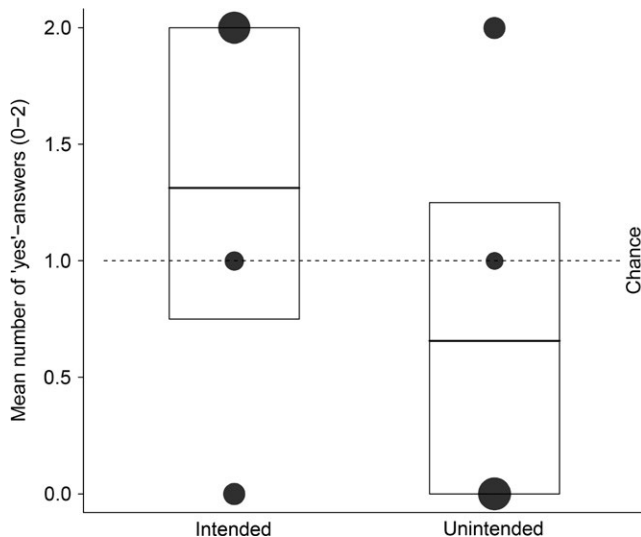


FIGURE 3 Mean sum score of “yes”-answers (0–2) to the test question “Did XY intentionally do that?”. The points depict the sum scores for each participant with the size of the points indicating sample size (i.e., the number of identical observations) and as boxes the quartiles as well as the mean per transgressor

characters (Intended Outcome Transgressor: $t(31) = 2.06$, $p = 0.048$, $r = 0.36$, False Belief Transgressor: $t(31) = 2.25$, $p = 0.032$, $r = 0.37$).²

5 | STUDY 2B

5.1 | Methods

5.1.1 | Participants

The final sample consisted of 74 5-year-olds (60–71 months, $M = 64.36$, 37 girls). Three additional children were tested but had to be excluded due to experimental errors. Children were from mixed socioeconomic backgrounds and recruited from the same local database of parents as in Study 1. Children who had previously participated in Study 1 or Study 2a were excluded from recruitment of Study 2b.

5.1.2 | Design and stories

Children were randomly assigned to either of the three conditions: intention contrast, false belief contrast or primed false belief contrast. The general design, counterbalancing and the three test questions as well as the sticker distribution were the same as in Study 1, except that children only received the two moral stories. The intention contrast and the false belief contrast were administered exactly as in Study 1, i.e., children in the intention contrast heard about the Intended Outcome and the Unintentional Action Transgressor while children in the false belief contrast heard about the Intended Outcome and the False Belief Transgressor (in counterbalanced order). In the primed false belief contrast children were also presented with the Intended Outcome and False Belief Transgressor,

however, before being asked the three test questions children were asked whether the agent intentionally brought about the outcome (“Did XY intentionally do that?”, as in Study 2a, for both the Intended Outcome and the False Belief Transgressor). Importantly, children were corrected when they gave an incorrect answer (e.g., “No, she didn’t know that there was blue fertilizer in the box. She did not do it intentionally.”) and reassured when they answered correctly (e.g., “Exactly, she didn’t know that there was blue fertilizer in the box. She did not do it intentionally.”). Again, we assessed children’s false belief competence with the two false belief tasks (unexpected contents and location change).

5.1.3 | Procedure

Children were tested individually in the lab or in a separated quiet room in their kindergarten. The general procedure was the same as in Study 2a.

5.1.4 | Coding

The coding for the test questions as well as the sticker allocation was performed as in Study 1 (correlations between test questions all $r > 0.31$ and $p < 0.01$).

5.2 | Results

The mean difference scores as a function of type of contrast are depicted in Figure 4. First, we conducted a one-way ANOVA to compare these difference scores across contrasts. The ANOVA yielded a main effect for type of contrast $F(1,71) = 6.15$, $p = 0.003$, $\eta_p^2 = 0.15$. Post-hoc pair-wise comparisons between the three contrasts showed the following results (all p -values Bonferroni corrected). Comparing the intention contrast and false belief contrast, we could replicate our results from Study 1 with bigger differences for the intention contrast than for the false belief contrast, $t(48) = 3.22$, $p = 0.006$, $r = 0.43$. In addition, difference scores were higher for the primed false belief than for the false belief contrast, $t(48) = 2.64$, $p = 0.033$, $r = 0.36$, but did not differ between the primed false belief contrast and the intention contrast, $t(46) = 1.00$, $p = 0.96$, $r = 0.15$.

Children’s sticker distribution resembled their explicit ratings: children gave significantly more stickers to the unintended outcome transgressor in the primed false belief ($M = 1.75$, $SD = 0.44$) than in the false belief contrast ($M = 1.35$, $SD = 0.75$, $t(46) = 2.305$, $p = 0.026$, $r = 0.32$), however, the amount of stickers given to the unintended outcome transgressor did not differ between primed false belief and intention contrast ($M = 1.83$, $SD = 0.38$, $t(46) = 0.7$, $p = 0.488$, $r = 0.10$).³

5.3 | Discussion

The aim of Study 2 was to investigate the cognitive foundation of the mental state asymmetry. Why do children consider the agent’s false belief less in contrast to her intentions in their normative judgments?

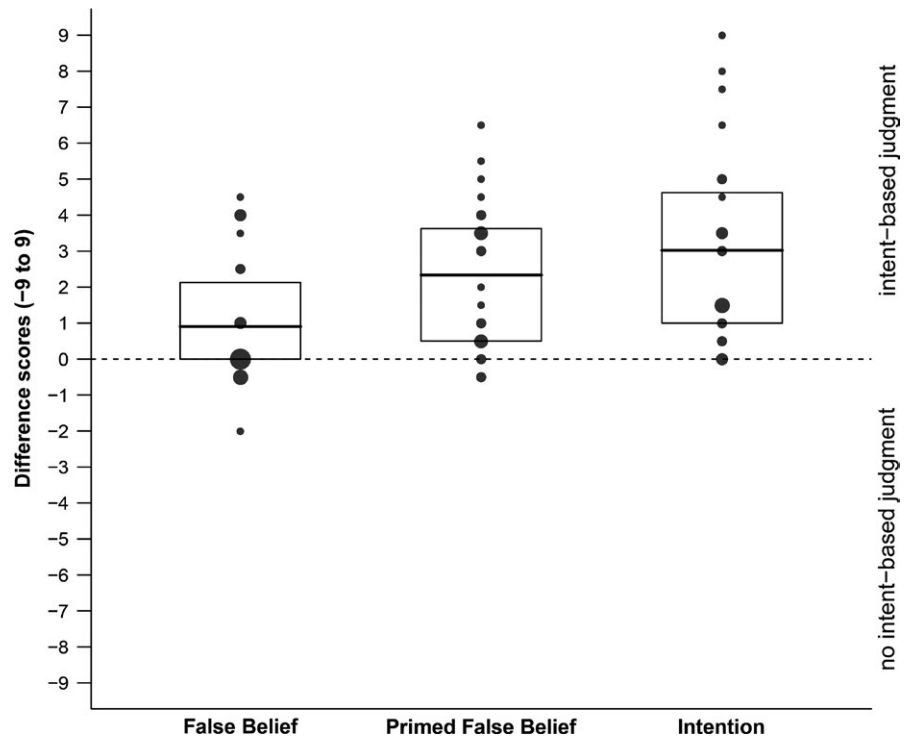


FIGURE 4 Mean difference score (–9 to 9) between the intended and unintended outcome transgressor as a function of type of contrast. The points depict the raw data for each participant with the size of the points indicating sample size (i.e., the number of identical observations) and as boxes the quartiles as well as the mean per contrast. *Note.* A positive difference score indicates intent-based judgments, while a negative score or a score of zero indicates no consideration of the agent’s mental states

The idea behind the two experiments was to assess children’s competence in each of the two required reasoning steps (false belief → unintended outcome → normative status) separately. If they failed to perform either or both steps, this would indicate a competence limitation. If, however, they were able to perform both steps separately, this would favor the idea of a performance limitation resulting from the inferential complexity to combine the two steps spontaneously (see Liu et al., 2007). And the latter is exactly what we found: children were able to process both steps separately and in a sequence, when explicitly asked to make the respective inferences (see Study 2a and 2b). However, they did not put the two steps together spontaneously, as indicated by their lack of ad hoc belief-based normative judgments (see Study 1 and 2b).

Concerning the first step, i.e., inferring that an action based on a false belief is not intended under the description of the outcome, we found that already 5-year-old children correctly made this inference. When asked about the intentional structure of the immoral actions, they classified the false-belief-based actions as unintentional and the intentional ones as intentional.

To assess children’s competence in the second reasoning step (from the unintended outcome to the normative status of the action), we first explicitly asked them about the intentional structure of the immoral action. Only after the correct(ed) answer, we then asked them for their normative evaluation of that action. Results revealed that after this priming, even the 5-year-olds considered the agent’s false belief in their normative judgments. In fact, their degree of intent-based judgments in these situations mirrored their degree of intent-based judgments for accidental mistakes. This pattern indicates that children are generally able to perform the second reasoning step as well as the combination of the two when explicitly guided through the reasoning step.

In combination, the results of the two experiments thus support the idea that the mental state asymmetry that became apparent in the literature and was explicitly found in the first study, mainly reflects children’s difficulty to spontaneously process the two inferences of belief-based judgments together. As children’s belief-based judgments could easily be enhanced when the reasoning chain was split into its components, the asymmetry seems to result from external performance factors rather than children’s moral reasoning capacities.

6 | GENERAL DISCUSSION

How do children form normative judgments in terms of the *why* something went wrong and the *what* went wrong? The present set of studies investigated this question concerning two crucial factors: the type of the agent’s underlying mental state (i.e., belief versus intention) and the type of norm transgressed (i.e., moral versus conventional norm).

Concerning the type of mental state, we found the predicted asymmetry in the sense that children (but not adults) made stronger intent-based judgments for accidental than for false-belief-based mistakes (Study 1). The two follow up studies highlighted inferential complexity as the most promising candidate to explain (away) children’s relative difficulty with including false beliefs. Even 5-year-olds were able to infer that a false belief leads to an unintended outcome (Study 2a). And most importantly, after being primed for this intentional structure of a false-belief-based action, the asymmetry disappeared (Study 2b). Concerning the type of norm transgressed we also found the second predicted asymmetry in the form that both

children and adults made more intent-based judgments for moral than for conventional norms (Study 1).

6.1 | The mental state asymmetry

The general asymmetry in children's consideration of intentions versus false beliefs in their normative evaluations nicely fits with and complements previous findings on the development of intent-based normative judgments (Killen et al., 2011; Nobes et al., 2009; Yuill & Perner, 1988). By exploring the cognitive foundation of this asymmetry, the insights from the current set of studies, however, go one important step further: the seemingly consecutive development—*younger children first consider only the intentionality of the action while only older children also consider the agent's false belief—most probably is not a product of moral development per se. Rather it seems to be a byproduct of their general socio-cognitive development: even younger children's intent-based judgments could easily be increased after reducing the inferential complexity of the two step reasoning process (false belief → unintended outcome → normative status).*

So, what are the external factors that potentially limit children's execution of spontaneous belief-based judgments? What competence do they acquire later on that helps them to form belief-based judgments spontaneously? From the current study, we cannot tell. Interestingly, however, from a theoretical perspective, the inferential complexity idea proposed here is much in line with several performance factor accounts of the development of moral reasoning. Such accounts assume that moral reasoning competence develops earlier than traditionally suggested, yet initially remains masked by performance factors such as a fragile working memory, limited executive functions or a restricted theory of mind (e.g., Chandler, Sokol, & Hallett, 2001; Killen et al., 2011; Margoni & Surian, 2016; Zelazo et al., 1996). But more specifically, how can these proposed external performance factors relate to the mental state asymmetry?

First, working memory capacities might directly influence the execution of the two step reasoning process: to include the agent's false belief in a normative judgment the child has to operate the first reasoning step (false belief → unintended outcome) and then use its output (unintended outcome) as the input of the second reasoning step. This process may be disrupted by a fragile working memory, particularly when one (or both) reasoning step(s) still require(s) cognitive effort (see Liu et al., 2007). In fact, children's weak, though significant performance on the first reasoning step (67%) may indicate that this inference step is still fragile and thus in need of cognitive resources. Second, executive functions might influence children's ability to suppress the predominant answer that is elicited by the negative outcome. Limited inhibitory control can impede this process, particularly when there are no visual cues for unintentionality, as it is the case for false-belief-based mistakes (Zelazo et al., 1996). Third, when making belief-based normative judgments children do not only have to be able to make a correct action prediction based on the false belief (the capacity measured by standard ToM tests as in the present studies). Rather, they may have to engage in more

sophisticated forms of perspective-taking such as reasoning about morally relevant false beliefs, an ability that might still be restricted until the late preschool years (MoToM, see Killen et al., 2011; Fu, Xiao, Killen, & Lee, 2014).

Future research will need to explore the role of inferential complexity and related performance factors on children's explicit intent-based judgments more directly and systematically.

Relatedly, the inferential complexity idea raises interesting new research questions for younger children. With more reduced performance factors, could even younger children engage in intent-based normative judgments? To date, several studies suggest that infants and toddlers in their social evaluations of prosocial and antisocial agents may be sensitive to the agent's underlying intent (see e.g., Choi & Luo, 2015; Fawcett & Liskowski, 2012; Hamlin, 2013; Hamlin, Wynn, & Bloom, 2007, 2010; Hamlin, Wynn, Bloom, & Mahajan, 2011). However, the robustness of these early competence findings as well as their interpretation (do they really tap precursors of moral evaluation?) remain disputed (see e.g., Cowell & Decety, 2015; Nighbor, Kohn, Normand, & Schlinger, 2017; Salvadori et al., 2015; Scarf, Imuta, Colombo, & Hayne, 2012). Future research should aim at systematically investigating the scopes, limits, and potential asymmetries of young children's early intent-based judgments in cognitively reduced cases of normative evaluations using more action-based methodologies.

6.2 | The asymmetry between normative evaluation of conventional vs. moral norm transgressions

While inferential complexity is an obvious and plausible candidate for explaining the asymmetry in belief-based and intention-based normative judgment, much less is currently known concerning the cognitive foundations of the asymmetry between intent-based judgments of moral versus conventional transgressions. From the current literature, two speculative possibilities emerge. First, as proposed in the introduction the asymmetry may be grounded in the different reactive attitudes elicited by transgressions of the two types of norms (Nichols, 2002, 2007). Moral norm transgressions (say, breaking someone's cup on purpose) call for reactive attitudes such as guilt and blame. And for the corresponding assessments of agents as guilty and acts as blameworthy, the agent's motivational set, her beliefs, desires, and intentions come into focus ("Did she know this was my cup? Did she break it knowingly and intentionally?"). Conventional norm violations (as putting green fish in the bowl for red fish), in contrast, rarely invoke such reactive attitudes and thus critique usually does not involve blame. Rather, the focus shifts away from the agent's motivational set and towards the outcome of the action. The consequence is mainly a behavior-based evaluation; was the outcome correct or incorrect ("No, there should only be red fish in this bowl"). It is true this reactive attitude possibility was not directly tested here. But in light of recent findings on children's different emotional reactions for moral versus conventional norms (see Hardecker et al.,



2016), it does seem at least plausible that similar reactions were invoked by the present conventional versus moral norm violation scenarios.

A second possibility is that the source of the moral-conventional asymmetry lies in the differential rule-dependency of conventional versus moral norms (as proposed by Giffin & Lombrozo, 2016; Turiel, 1983). The basic idea is this: moral norms are thought to be rule independent. That is, moral transgressions are seen wrong independently of whether or not a given rule applies in the context in question. Therefore, the sole and crucial basis for evaluating such a transgression is the will of the agent. Was the transgression intentional or unintentional? Conventional norms, in contrast, are thought to be rule dependent. A given action is not wrong in and of itself but wrong only relative to a rule prohibiting the behavior in question (Turiel, 1983). Thus, when evaluating conventional violations, there are two important sources of information: first, whether a given rule applied or not; and second whether the action was intentional. As a consequence, since the will of the agent is not conclusive, information about the mental states of the agent fades more into the background.

From the present design, we cannot tell which (if any) of the two possibilities applies. Crucially, the moral/conventional contrast used in the present study equally applied to both explanations: our cases equivalently differed both on the dimension of the reactive attitudes as well as on the dimension of rule dependency. However, this does not apply to all cases of moral and conventional norms. Moral norms can be rule-governed and conventional norms can have harm implications that might elicit similar reactive attitudes as moral norms (e.g., dress code at a funeral). Future research is needed that extends the present findings to these less clear-cut cases of social norms (see, e.g., disgust norms, Nichols, 2002). Such research would then help to disentangle the two options (and in fact, it might turn out that they are actually complementary rather than strictly incompatible) and to pinpoint whether the specific contrast used here generalizes to moral-conventional contrasts more broadly.

6.3 | Other factors influencing intent-based judgments/limitations of the present studies

In the present set of studies, we focused on two factors influencing intent-based normative judgments: the type of the agent's mental state and the kind of norm transgressed. However, the list of potential influences is certainly much longer so that our focus on the two might have led to the neglect of others. We want to highlight two additional factors here: the type of judgment to be made and negligence.

Concerning the type of judgment to be made, Cushman et al. (2013) already showed that, like adults, children focus more on the underlying mental state when asked for the wrongness of the action than for punishment judgments. Concerning negligence, previous work has shown that already 3-year-olds are sensitive to the carefulness with which an action has been performed (Nobes

et al., 2009). In the stories used in the present studies, we did not explicitly state that the unintentional transgressor acted carefully, so we cannot completely rule out that children found the characters to be negligent. However, it seems unlikely that such an impression can explain the present asymmetries. First, there was no indicators that children thought the false belief agent were more negligent than the accidental one. Second, even if that was the case, it seems unlikely that this impression was eliminated by the question about the intentional structure of the action in Study 2b, where children did not differentiate between accidental and false-belief-based mistakes anymore. Future research is necessary to see how these factors interact with the two factors under investigation in the present set of studies.

7 | CONCLUSION

The present set of studies documented two asymmetries in children's intent-based normative judgments. First, children's judgments—in contrast to adults'—are influenced by the kind of unintentional transgression: 5- and 7-year-olds made more intent-based judgments for accidents than for false-belief-based mistakes. However, this asymmetry seems to be mainly reducible to external performance factors, such as the inferential complexity of reasoning lines. Second, from age five, children, like adults, seem to share the intuition that the underlying intentional structure of an action matters more for moral than for socio-conventional norm violations. Future research will need to explore more systematically the cognitive foundations as well as early origins of these asymmetries.

ACKNOWLEDGEMENTS

Special thanks to Laura Schwerdt, Lisa Wenzel, Marlene Vahl, and Inga Bethan for help with testing, to Alice Graebener for the illustrations and to Marlen Kaufmann and Konstanze Schirmer for help with recruiting children and organizing the study.

ENDNOTES

¹See General Discussion section for further discussion on how the prototypical contrasts might (not) generalize to moral-conventional contrasts more broadly.

²The additional data and analyses concerning children's Theory of Mind understanding can be found in the supplementary material.

³The additional data and analyses concerning children's Theory of Mind understanding can be found in the supplementary material (see Table S2).

REFERENCES

- Chandler, M.J., Sokol, B.W., & Hallett, D. (2001). Moral responsibility and the interpretive turn: Children's changing conceptions of truth and rightness. In B. Malle, L.J. Moses & D. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 345–365). Cambridge, MA: Mit Press.

- Choi, Y.-J., & Luo, Y. (2015). 13-month-olds' understanding of social interactions. *Psychological Science*, 26(3), 274–283. <https://doi.org/10.1177/0956797614562452>
- Cowell, J.M., & Decety, J. (2015). Precursors to morality in development as a complex interplay between neural, socioenvironmental, and behavioral facets. *Proceedings of the National Academy of Sciences of the United States of America*, 112(41), 12657–12662. <https://doi.org/10.1073/pnas.1508832112>
- 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>
- Dunfield, K.A., & Kuhlmeier, V.A. (2010). Intention-mediated selective helping in infancy. *Psychological Science*, 21(4), 523–527. <https://doi.org/10.1177/0956797610364119>
- Fawcett, C., & Liszkowski, U. (2012). Infants anticipate others' social preferences. *Infant and Child Development*, 21(3), 239–249. <https://doi.org/10.1002/icd.739>
- 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>
- Giffin, C., & Lombrozo, T. (2016). Wrong or merely prohibited: Special treatment of strict liability in intuitive moral judgment. *Law and Human Behaviour*, <https://doi.org/10.1037/lhb0000212>
- Giffin, C., & Lombrozo, T. (2018). An actor's knowledge and intent are more important in evaluating moral transgressions than conventional transgressions. *Cognitive Science*, 42, 105–133. <https://doi.org/10.1111/cogs.12504>
- Guglielmo, S., Monroe, A.E., & Malle, B.F. (2009). At the heart of morality lies folk psychology. *Inquiry*, 52(5), 449–466. <https://doi.org/10.1080/00201740903302600>
- Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316(5827), 998–1002. <https://doi.org/10.1126/science.1137651>
- Hamlin, J.K. (2013). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition*, 128(3), 451–474. <https://doi.org/10.1016/j.cognition.2013.04.004>
- Hamlin, J.K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, 450(7169), 557–559. <https://doi.org/10.1038/nature06288>
- Hamlin, J.K., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science*, 13(6), 923–929. <https://doi.org/10.1111/j.1467-7687.2010.00951.x>
- Hamlin, J.K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences of the United States of America*, 108(50), 19931–19936. <https://doi.org/10.1073/pnas.1110306108>
- Hardecker, S., Schmidt, M.F.H., Roden, M., & Tomasello, M. (2016). Young children's behavioral and emotional responses to different social norm violations. *Journal of Experimental Child Psychology*, 150, 364–379. <https://doi.org/10.1016/j.jecp.2016.06.012>
- 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. <https://doi.org/10.1111/1467-8624.00266>
- Heyman, G.D. (2009). Children's reasoning about traits. *Advances in Child Development and Behavior*, 37, 105–143. [https://doi.org/10.1016/S0065-2407\(09\)03703-3](https://doi.org/10.1016/S0065-2407(09)03703-3)
- Josephs, M., Kushnir, T., Gräfenhain, M., & Rakoczy, H. (2016). Children protest moral and conventional violations more when they believe actions are freely chosen. *Journal of Experimental Child Psychology*, 141, 247–255. <https://doi.org/10.1016/j.jecp.2015.08.002>
- Killen, M., Mulvey, K.L., Richardson, C., Jampol, N., & Woodward, A. (2011). The accidental transgressor: Morally-relevant theory of mind. *Cognition*, 119(2), 197–215. <https://doi.org/10.1016/j.cognition.2011.01.006>
- Liu, D., Gelman, S.A., & Wellman, H.M. (2007). Components of young children's trait understanding: Behavior-to-trait inferences and trait-to-behavior predictions. *Child Development*, 78(5), 1543–1558. <https://doi.org/10.1111/j.1467-8624.2007.01082.x>
- Margoni, F., & Surian, L. (2016). Explaining the U-shaped development of intent-based moral judgments. *Frontiers in Psychology*, 7, 219. <https://doi.org/10.3389/fpsyg.2016.00219>
- Nelson, S.A. (1980). Factors influencing young children's use of motives and outcomes as moral criteria. *Child Development*, 51(3), 823–829. <https://doi.org/10.1111/1467-8624.ep12330467>
- Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment. *Cognition*, 84(2), 221–236. [https://doi.org/10.1016/s0010-0277\(02\)00048-3](https://doi.org/10.1016/s0010-0277(02)00048-3)
- Nichols, S. (2007). After incompatibilism: A naturalistic defense of the reactive attitudes. *Philosophical Perspectives*, 21(1), 405–428. <https://doi.org/10.1111/j.1520-8583.2007.00131.x>
- Nighbor, T., Kohn, C., Normand, M., & Schlinger, H. (2017). Stability of infants' preference for prosocial others: Implications for research based on single-choice paradigms. *PLoS ONE*, 12(6), e0178828. <https://doi.org/10.1371/journal.pone.0178818>
- 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(4), 382–397. <https://doi.org/10.1016/j.jecp.2009.08.001>
- Nucci, L.P., & Nucci, M.S. (1982). Children's social interactions in the context of moral and conventional transgressions. *Child Development*, 53(2), 403–412. <https://doi.org/10.2307/1128983>
- Núñez, M., & Harris, P.L. (1998). Psychological and deontic concepts: Separate domains or intimate connection? *Mind & Language*, 13(2), 153–170. <https://doi.org/10.1111/1468-0017.00071>
- Perner, J., & Howes, D. (1992). 'He thinks he knows': And more developmental evidence against the simulation (role taking) theory. *Mind & Language*, 7(1–2), 72–86. <https://doi.org/10.1111/j.1468-0017.1992.tb00197.x>
- 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(2), 125–137. <https://doi.org/10.1111/j.2044-835X.1987.tb01048.x>
- Piaget, J. (1932). *The moral judgment of the child*. Oxford, UK: Harcourt, Brace.
- Rholes, W.S., & Ruble, D.N. (1984). Children's understanding of dispositional characteristics of others. *Child Development*, 55, 550–560. <https://doi.org/10.2307/1129966>
- Russell, P.S., & Giner-Sorolla, R. (2011). Moral anger, but not moral disgust, responds to intentionality. *Emotion*, 11(2), 233–240. <https://doi.org/10.1037/a0022598>
- Salvadori, E., Blazsekova, T., Volein, A., Karap, Z., Tatone, D., Mascaro, O., & Csibra, G. (2015). Probing the strength of infants' preference for helpers over hinderers: Two replication attempts of hamlin and Wynn (2011). *PLoS ONE*, 10(11), e0140570. <https://doi.org/10.1371/journal.pone.0140570>
- Samland, J., Josephs, M., Waldmann, M.R., & Rakoczy, H. (2016). The role of prescriptive norms and knowledge in children's and adults' causal selection. *Journal of Experimental Psychology: General*, 145(2), 125–130. <https://doi.org/10.1037/xge0000138>
- Scarf, D., Imuta, K., Colombo, M., & Hayne, H. (2012). Social evaluation or simple association? Simple associations may explain moral reasoning in infants. *PLoS ONE*, 7(8), e42698. <https://doi.org/10.1371/journal.pone.0042698>
- Smetana, J.G. (1981). Preschool children's conceptions of moral and social rules. *Child Development*, 52(4), 1333–1336. <https://doi.org/10.2307/1129527>
- 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.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>
- Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, UK: Cambridge University Press.
- Vaish, A., Carpenter, M., & Tomasello, M. (2010). Young children selectively avoid helping people with harmful intentions. *Child Development*, 81(6), 1661–1669. <https://doi.org/10.1111/j.1467-8624.2010.01500.x>
- Wang, F., Zhu, L., & Shi, K. (2011). How do children coordinate information about mental states with social norms? *Cognitive Development*, 26(1), 72–81. <https://doi.org/10.1016/j.cogdev.2010.09.002>
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128. [https://doi.org/10.1016/0010-0277\(83\)90004-5](https://doi.org/10.1016/0010-0277(83)90004-5)
- Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences of the United States of America*, 104(20), 8235–8240. <https://doi.org/10.1073/pnas.0701408104>
- Young, L., Nichols, S., & Saxe, R. (2010). Investigating the neural and cognitive basis of moral luck: It's not what you do but what you know. *Review of Philosophy and Psychology*, 1(3), 333–349. <https://doi.org/10.1007/s13164-010-0027-y>

- Young, L., & Saxe, R. (2011). When ignorance is no excuse: Different roles for intent across moral domains. *Cognition*, 120(2), 202–214. <https://doi.org/10.1016/j.cognition.2011.04.005>
- 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>
- Zelazo, P.D., Helwig, C.C., & Lau, A. (1996). Intention, act, and outcome in behavioral prediction and moral judgment. *Child Development*, 67(5), 2478–2492. <https://doi.org/10.2307/1131635>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Proft M, Rakoczy H. The ontogeny of intent-based normative judgments. *Dev Sci*. 2018;e12728. <https://doi.org/10.1111/desc.12728>

APPENDIX

Moral Flower Story

Main story

These are grandma's beautiful flowers. Grandma really loves her flowers and she will be really sad if they die. Next to the flowers is the fertilizer for the flowers. Do you know what a fertilizer is? (explained if necessary) As the flowers are yellow, they always need to get the yellow fertilizer. When they get the blue fertilizer they will die and Grandma will be really sad. Now, can you show me the boxes that can be used for the flowers?

Intended Outcome Transgressor

This is Lisa. Lisa says: "I want to fertilize Grandma's flowers with the blue fertilizer." Lisa goes to the shelf, takes the blue fertilizer in the blue box and pours the fertilizer on the flowers. Look, Grandma's beautiful flowers died.

Unintentional Action Transgressor

This is Kathrin. Kathrin says: "I want to fertilize Grandma's flowers with the yellow fertilizer." Kathrin goes to the shelf. Look, on the way to the shelf she trips and bumps against the shelf. In doing so the box with the blue fertilizer falls down, directly on Grandma's flowers. Look, Grandma's beautiful flowers died.

False Belief Transgressor

Look at this yellow box of fertilizer. I have to tell you something about it. It looks as if there was yellow fertilizer in it. However, there is blue fertilizer in it. But one cannot see that. It looks as if there was yellow

fertilizer in it. This is Maja. Maja says: "I want to fertilize Grandma's flowers with the yellow fertilizer." Maja goes to the shelf and takes the yellow box. Maja does not know that it contains blue fertilizer. She thinks there is yellow fertilizer in there. Maja takes the box and pours the fertilizer on the flowers. Look, Grandma's beautiful flowers died.

Moral Cupcake Story

Main story

This is Caro. And here is Caro's delicious cupcake. Caro made it all by herself and she is so excited to eat it later.

Intended Outcome Transgressor

This is Simon. Simon says: "I now want to eat Caro's cupcake." Simon goes to the table, takes the cupcake and eats it up. Now Caro's cupcake is gone.

Unintentional Action Transgressor

This is Tim. Tim says: "I want to put Caro's cupcake in the shelf, so it does not get broken." Tim takes the cupcake and walks to the shelf. Look, on the way to the shelf he trips and the cupcake falls on the ground. Now Caro's cupcake is broken.

False Belief Transgressor

Caro puts her cupcake in this brown paper bag, because she wants to take it to her friend later where she wants to eat it. She tightly closes the paper bag so that one cannot see that the cupcake is in there. This is Moritz. Moritz wants to help his Mum tidying up. He does not

know that the cupcake is in the paper bag. He thinks there is garbage in the bag. He says: "Look, here is the paper bag with the garbage. I want to throw the garbage in the garbage can." He goes to the table, takes the bag and throws it in the garbage can. Now Caro's cupcake is gone.

Conventional Picture Story

Main story

On this table, there is the picture of an elephant. And here are many different cans with painting colors. For the picture there is a rule. The rule goes like this: It is only allowed to paint the elephant in blue. Can you show me the cans of colors that are allowed to be used for the picture?

Intended Outcome Transgressor

This is Lena. Lena says: "I now want to paint the elephant in green." She goes to the table, takes a can with green paint and puts the paint on the picture. Now the elephant is green.

Unintentional Action Transgressor

This is Jenny. Jenny says: "I now want to paint the elephant in blue." Jenny goes to the table. Look, on her way to the table she trips and bumps against the table. In doing so a can with green color tilts over, directly on the elephant. Now the elephant is green.

False Belief Transgressor

Look at this blue can of painting colors. It looks as if there was blue paint in it. However, there is green paint in it. But one cannot see that. It looks as if there was blue paint in it. This is Melanie. Melanie says: "I now want to paint the elephant in blue." Melanie goes to the

table and takes the blue can. Melanie does not know that it contains green paint. She thinks there is blue paint in it. She takes the can and puts the paint on the picture. Now the elephant is green.

Conventional Aquarium Story

Main story

Here is an aquarium. On the shelf next to the aquarium are lots of boxes with fish. For the aquarium there is a rule. The rule goes like this: It is only allowed to put red fish in the aquarium. Can you show the boxes with fish that are allowed to be put in the aquarium?

Intended Outcome Transgressor

This is Max. Max says: "I now want to put green fish in the aquarium." Max goes to the shelf, takes a green box with green fish and dumps them in the aquarium. Now there are green fish in the aquarium.

Unintentional Action Transgressor

This is Peter. Peter says: "I now want to put red fish in the aquarium." Peter goes to the shelf. Look, on the way to the shelf he trips and bumps against the shelf. In doing so a box with green fish falls down, directly into the aquarium. Now there are green fish in the aquarium.

False Belief Transgressor

Look at this red box. It looks as if there were red fish in there. However, there are green fish in it. But one can't see that. It looks as if there were red fish in there. This is Leo. Leo says: "I now want to put red fish in the aquarium." Leo goes to the shelf and takes the red box. Leo does not know that it contains green fish. He thinks that there are red fish in there. He takes the box and dumps it in the aquarium. Now there are green fish in the aquarium.