

Moral Complexity in Middle Childhood: Children's Evaluations of Necessary Harm

Marc Jambon and Judith G. Smetana
University of Rochester

We assessed 5- to 11-year-olds' ($N = 76$) judgments of straightforward moral transgressions (prototypical harm) as well as their evaluations of complex, hypothetical scenarios in which an actor transgresses in order to prevent injury (necessary harm). The nature of the actor's transgression (psychological or physical harm) varied across participants. Moral judgments and justifications, knowledge of the actor's psychological experience, and their associations were examined. At all ages, children negatively evaluated prototypical harm; judgments of necessary harm became increasingly more forgiving with age as justifications pertaining to the actor's harm decreased. References to the actor's positive actions and children's tendency to coordinate conflicting concerns increased with age, but only when evaluating psychological harm. Across conditions, older children viewed transgressors as holding increasingly more positive attitudes toward their own actions, and this was uniquely associated with more forgiving moral judgments and justifications of necessary but not prototypical harm. Findings are discussed in relation to the emergence of more flexible and nuanced moral evaluations during middle childhood.

Keywords: moral judgments, middle childhood, social domain theory

Supplemental materials: <http://dx.doi.org/10.1037/a0032992.supp>

Children as young as 3–4 years of age judge intentional acts of physical and verbal aggression, such as hitting or hurting another's feelings, to be morally wrong and deserving of punishment, based on the harmful consequences for others (Smetana & Braeges, 1990; Smetana, Rote, et al., 2012). Although this early developing moral concern is maintained across the life span, research from the social domain perspective (see Smetana, 2006; Turiel, 2006, for reviews) suggests that an important aspect of normative social development involves an understanding that many situations entail conflicts between different legitimate, yet competing moral concerns (Helwig, 1995; Nucci & Turiel, 2009). For instance, although norms against harming others appear to be universal (Turiel, 2006), intentionally hurting others may be morally justifiable when aimed at preventing greater harm, such as in cases of self-defense or the protection of loved ones. Furthermore, this type of well-intentioned, necessary harm represents a common feature of children's and adolescents' everyday moral experiences (Wainryb, Brehl, & Matwin, 2005). Examining age differences in children's reasoning about necessary harm would contribute to our understanding of moral development.

During the early school years, children begin to make rudimentary distinctions between malevolent transgressions and those entailing mitigating circumstances (Darley & Shultz, 1990). Children as young as 6 years of age, along with adolescents and adults, are more forgiving of positively motivated acts of aggression compared with selfish harm, with few age differences in judgments evident beyond the early elementary school years (Darley, Klosson, & Zanna, 1978; Rule, Dyck, McAra, & Nesdale, 1975; Rule, Nesdale, & McAra, 1974). However, prior studies aimed to establish the earliest ages at which children could use information about an actor's intentions (i.e., deliberate vs. accidental) and motives (i.e., positive vs. negative). Little attention was paid to the specific nature and consequences of the moral transgressions themselves, and children's reasoning was not assessed.

For instance, Darley et al. (1978) combined acts such as one child throwing water on another to stop them from playing with matches with more prototypical instances of physical aggression, despite potential differences in the extent to which the transgressions caused harm (i.e., getting wet vs. physical pain). Similarly, although Rule et al. (1974) controlled for the type of transgression, the consequences for the story victims were downplayed (e.g., "Jim trips, but does not fall"), making it ambiguous whether harm occurred. Additionally, judgments were restricted to ratings of punishment or actor naughtiness. However, more recent research has demonstrated that even when younger and older children make what appear to be similar judgments of right or wrong, the underlying reasons for their evaluations may differ in important ways (Nucci & Turiel, 2009). Given these limitations, the lack of age differences in these studies must be interpreted with caution.

Parallels may be drawn between necessary harm and recent research on judgments of prosocial lie-telling. Although all children negatively evaluate antisocial deception, elementary school-

This article was published Online First May 6, 2013.

Marc Jambon and Judith G. Smetana, Department of Clinical and Social Sciences in Psychology, University of Rochester.

We thank the local daycare directors and staff for their enthusiasm about the project as well as the children who participated in this research. We also thank Emily Rus and Janine Steffan for their assistance in data collection and coding.

Correspondence concerning this article should be addressed to Marc Jambon, Department of Clinical and Social Sciences in Psychology, Melliora Hall, RC 270266, University of Rochester, Rochester, NY 14627. E-mail: mjambon@psych.rochester.edu

age children begin to hold more favorable views of lies aimed at preventing harm or discomfort to others (compared with antisocial lies). Evaluations of prosocial lie-telling become more forgiving, and “blunt truth-telling” becomes less forgiving across childhood (Popliger, Talwar, & Crossman, 2011; Xu, Bao, Fu, Talwar, & Lee, 2010). This is accompanied by older children’s increased recognition of the consequences such actions have for others (Heyman, Sweet, & Lee, 2009) and a greater ability to coordinate concerns about honesty with the desire to avoid harm (Xu et al., 2010). Thus, the ability to flexibly coordinate moral beliefs about lying with concerns about preventing harm continues to develop beyond early childhood. However, the fact that recipients of “white lies” do not experience negative repercussions makes it unclear whether these judgments also apply to situations involving actual harm. Therefore, it is an open question whether children’s judgments of necessary harm continue to develop at later ages. Additionally, it is unclear what features children attend to in making their judgments and whether different contextual or developmental factors influence the salience of these concerns at different ages. Social domain theory (Helwig, 1995; Smetana, 2006; Turiel, 2006) provides a useful framework for examining these questions.

Studies from this perspective have explicitly focused on how children of different ages attend to and coordinate different concerns in judging complex situations (Helwig, 1995; Helwig, Hildebrandt, & Turiel, 1995; Nucci & Turiel, 2009; Wainryb et al., 2005). Although individuals hold uniformly negative views of straightforward, prototypical moral transgressions, evaluations of multifaceted situations become more mixed and less categorical with age. This is due in part to increases in the cognitive capacity to balance, coordinate, and integrate multiple elements of a situation. As noted by Nucci and Turiel (2009), the increased recognition that many everyday situations entail a conflict between different moral (and nonmoral) concerns results in greater flexibility in applying moral concepts in specific social contexts.

Intersections Between Psychological Knowledge and Children’s Moral Evaluations

In addition to general age-related increases in the ability to balance and coordinate concerns, specific advances in children’s psychological understanding of others may contribute to their moral development. Perspective-taking abilities, including the capacity to coordinate one’s own perspective with those of another, have long been theorized to play an important role in the emergence of more mature moral reasoning (e.g., Colby & Kohlberg, 1987; Selman, 1980). More recently, much research has examined how young children’s developing understanding of mental states, such as beliefs and intentions (i.e., their theory of mind), informs moral judgments (Killen, Mulvey, Richardson, Jampol, & Woodward, 2011; Smetana, Jambon, Conry-Murray, & Sturge-Apple, 2012; Wainryb & Brehl, 2006). For instance, Wainryb and Ford (1998) found that 5- and 7-year-olds, but not 3-year-olds, understood that others may hold different beliefs from their own (e.g., a teacher thinks girls need more food than boys). As a result, they were more likely than younger children to excuse unfair practices based on those beliefs (e.g., unequal distribution of food based on gender). Similarly, Killen et al. (2011) found that children who lacked an understanding that others may hold mistaken or false

beliefs—an ability typically in place by 4–5 years of age (Wellman, Cross, & Watson, 2001)—incorrectly attributed negative motives to a character who unwittingly transgressed against others. Consequently, they judged these accidental transgressors as more deserving of punishment than did children who grasped false beliefs. This suggests that a rudimentary understanding of others’ minds allows children to make more informed moral evaluations during early childhood.

Despite the relevance of mental state understanding for moral judgments, little research has explored these connections at later ages. This is surprising given that a large body of developmental research has identified middle to late childhood as an important period of growth in children’s social perspective taking and psychological knowledge (Hoffman, 2000; Selman, 1980; Wainryb & Brehl, 2006). During this time, children come to appreciate that mental states are constructed and maintained by active psychological agents rather than stemming from an objective external reality. For instance, although 4- to 5-year-olds understand that others may hold different or false beliefs about the world, they often still do not see them as legitimate, judging others with different beliefs to be wrong or mistaken (Carpendale & Chandler, 1996). With the gradual shift to a more active understanding of mental life and the ability to coordinate one’s psychological perspective with others, children come to appreciate that individuals may hold differing, yet equally legitimate interpretations of the same event. This is accompanied by increases in the frequency and complexity of mental state information children use in evaluating their experiences (Wainryb et al., 2005) and a more advanced ability to disentangle intentions from outcomes (Nobes, Panagiotaki, & Pawson, 2009; Selman, 1980). Additionally, around 7 or 8 years of age, children begin to appreciate the possibility of multiple, conflicting desires and emotions and coordinate this mental state knowledge with social and moral norms (Lagattuta, 2005; Shaw & Wainryb, 2006; Wainryb & Brehl, 2006).

These age-related changes have important implications for the development of more complex and flexible moral judgments in middle childhood. Although preschoolers are sensitive to others’ thoughts, it is not until somewhat later that their moral evaluations begin to consistently and systematically reflect a concern for others’ beliefs and motives when judging morally relevant practices (Selman, 1980; Wainryb & Brehl, 2006; Wainryb et al., 2005). For example, Sokol, Chandler, and Jones (2004) found that 5- to 7-year-olds who lacked an interpretive understanding of the mind were less likely to attribute responsibility to a transgressor who tried but failed to cause harm, despite their full awareness of the actor’s malicious intent. Others have found similar discrepancies between early school-age children’s basic knowledge of mental states and their ability to integrate that information into their moral evaluations (Killen et al., 2011; Nobes et al., 2009; Shaw & Wainryb, 2006; Wainryb & Ford, 1998).

Psychological knowledge may also be important for understanding more abstract moral concepts, such as psychological harm (e.g., name-calling, teasing; Helwig, 1995). Studies indicate that children’s thinking about psychological harm may develop later than physical harm. Prior to the age of 7 or 8, children show deficits in their ability to attend to the moral features of psychologically harmful acts, such as the victim’s perspective and an actor’s intent (Helwig et al., 1995; Helwig, Zelazo, & Wilson, 2001). As a result, younger children are often more accepting of

actions leading to emotional rather than physical distress. More research is needed, however, to understand how specific developments in children's psychological knowledge of others beyond the preschool years are implicated in the emergence of more complex, flexible moral evaluations.

The Current Study

We sought to address this gap by examining age-related changes in children's understanding that causing harm to others may be morally justifiable in certain circumstances. Using semistructured interviews, we examined 5- to 11-year-olds' moral and psychological conceptions of complex, hypothetical situations in which an *actor* intentionally harms a *friend* to stop the friend from potentially causing serious bodily harm. We varied the type of harm used by the actor (physical, psychological) and the identity of the target (actor/third party/friend).

The major goal of the study was to examine the extent to which judgments of necessary harm show age-related changes in complexity across the later childhood years. On the basis of past research suggesting that increases in the ability to coordinate multiple concerns and to consider the psychological perspective of others influences moral judgments (Helwig, 1995; Nucci & Turiel, 2009; Selman, 1980; Wainryb et al., 2005), we hypothesized that necessary harm would be seen as more acceptable and less deserving of punishment with age. Additionally, we expected that older children would focus more on the actor's positive motives and preventive actions, and would evidence a greater tendency to coordinate different competing concerns than younger children. We did not anticipate age differences in judgments of prototypical harm, as all children were expected to negatively evaluate these events. We view the capacity to coordinate multiple concerns as a general social-cognitive ability that increases across childhood and adolescence. Other theorists (e.g., Selman, 1980) have also used the concept of coordination to refer to children's and adolescents' emerging ability to consider others' perspectives in relation to one's own. This latter use of coordination shares considerable overlap with our conceptualization of psychological knowledge.

Given that 4- to 5-year-olds can correctly attribute beliefs and intentions to others (Carpendale & Chandler, 1996; Wainryb & Brehl, 2006; Wellman et al., 2001), we expected that most students would understand that the protagonist in each story was trying to prevent harm. Nevertheless, younger children were expected to have a more limited understanding of how these protagonists would interpret and evaluate their own actions. Specifically, we expected older children to rate actors engaging in necessary harm as having more positive and accepting views of their actions than younger children. This appreciation for an *actor's interpretation* of an event entails more than simple knowledge of an agent's intentions (e.g., protect themselves or others), but requires the ability to take their perspective and appreciate how the actors themselves would subjectively appraise their own harmful behavior (Selman, 1980).

We expected that children who rated actors as having more positive interpretations of necessary harm would offer more forgiving moral evaluations, regardless of age and simple knowledge of harmful intent. Given prototypical actors' selfish motivation, it was unclear what interpretations children would ascribe to prototypical transgressors. Regardless, all children were expected to

base their evaluations of straightforward harm on the negative consequences of the act (Helwig et al., 2001; Smetana, 2006). Actor interpretation ratings were therefore not expected to be associated with moral judgments of prototypical harm.

Previous research suggests that younger children have a more constrained understanding of psychological harm than older children, often leading to more permissive views of emotional compared with physical distress (Helwig, 1995; Helwig et al., 1995, 2001). However, the goal of previous research was to explore how children weigh intentions and outcomes in determining whether a moral transgression has occurred. Thus, in past studies, scenarios requiring children to infer the mental states of transgressors and victims were used. In contrast, the present study depicted scenarios in which the motives and consequences were unambiguous. It was therefore unclear whether older and younger children would differ in their evaluations of physical and psychological harm. We examined harm type but did not test specific hypotheses.

Method

Participants/Sample

The total sample consisted of 76 children ranging from 5.17 to 11.42 years of age ($M = 7.79$, $SD = 1.85$). Age was normally distributed, with skewness of .326 ($SE = .276$) and kurtosis of -1.061 ($SE = .545$), and was therefore treated as a continuous variable in all analyses (see the Analysis Plan section below). However, we describe the sample here in terms of a median split by age. The younger group consisted of 38 children (20 boys) ranging in age from 5.17 to 7.75 years ($M = 6.20$, $SD = .70$). The older group consisted of 38 children (23 boys) ranging in age from 7.92 to 11.42 ($M = 9.38$, $SD = 1.01$). The total sample was 72% European American, 13% Asian, 9% African American, 4% Hispanic, and 2% "other" (primarily biracial). Participants were recruited from seven after-school care and summer camp programs serving lower-middle to upper-middle socioeconomic families in a moderately sized city in the Northeastern United States.

Measures

Necessary harm vignettes. The stimuli consisted of 12 8×11 -in. colored drawings depicting four hypothetical stories (three *necessary harm* stories and one *prototypical harm* control story). Each necessary harm story depicted an actor harming a friend to stop him or her from performing an act that would likely result in serious bodily injury. The first picture in each story introduced the two characters and revealed the friend's intention to engage in the injurious act. The second picture showed the actor unsuccessfully imploring his or her friend not to engage in the act. It was clearly stated that the actor did not have time to seek help before the friend's action would be completed, necessitating an immediate response if serious harm were to be avoided. The final frame depicted the actor (with a neutral facial expression) deliberately hurting his or her friend to stop the event from occurring. Each story ended with the friend running home crying without having caused injury. The prototypical control story followed a similar narrative structure, but instead of preventing harm, the actor was described as hurting his or her friend in order to get a turn on a swing.

The type of harm used by the actor was treated as a between-subject factor, with half the children ($n = 38$: 20 younger children, 18 older children) hearing stories (prototypical and necessary harm) in which actors caused *psychological harm* to the friend (e.g., calling her a bad name and hurting her feelings), whereas the other half heard stories involving *physical harm* (e.g., hitting friend). A between-subject design for this factor was chosen to ensure that perceptions of one harm type would not influence judgments of the other type of harm.

The target of the friend's actions (i.e., the person the actor was trying to protect) was treated as a within-subject factor; the friend was depicted as potentially causing harm to him- or herself (*friend protect*), an innocent third party (*innocent protect*), or the actor (*self-defense*). Because the friend in the prototypical story was not trying to cause harm, there was no "target" in this condition. All characters were matched to the participants' age and gender. See Table 1 for the major components of each story, and the supplemental material for the full stories.

Moral evaluations. After each story, children's moral evaluations were assessed using interview procedures adapted from previous research (Helwig et al., 2001; Smetana, 2006). For each story, children were asked the following questions in a fixed order: "Was it ok or not ok for [actor] to [hurt friend]?" followed by, "Was it a little bad or very bad (just ok, a little good, or very good)?" Due to limited use of the "little good" and "very good" options, responses were collapsed across those response categories and scored on a 3-point scale ranging from 0 (*just ok/little good/very good*) to 2 (*very bad*), assessing *act acceptability*. *Justifications* for their act evaluation were assessed by asking children to explain their judgments ("Why?"). Next, children were asked, "Should [actor] get into trouble?" and if yes, "a little bit or a lot?" Answers were scored on a 3-point *deserved punishment* scale ranging from 0 (*no trouble*) to 2 (*a lot of trouble*).

On the basis of theory, pilot testing, and coding 20% of the protocols, justifications were coded using categories developed for the current study. The final justification coding system, presented in Table 2, included nine categories. Categories that were used were assigned a score of 1; if the category was not used, they were assigned a score of 0. There was no limit on the number of justifications allowed. Prior to coding justifications, a trained research assistant scored an additional 20% of randomly selected protocols for reliability. Interrater agreement was excellent ($\kappa = .94$).

To assess whether concerns with an actor's positive intention to prevent serious harm were seen as in conflict with moral concerns about actively hurting others, justifications were also coded for the

presence or absence of a *coordination of concerns*. Responses attempting to balance the negative aspects of the actor's actions with the positive consequences of preventing harm (e.g., "It's not good to hit people, but sometimes it's ok if she's trying to stop her from shaking the tree") were coded as 1, whereas responses that did not contain multiple, conflicting concerns were coded as 0. Interrater agreement on the coding of coordination responses, also obtained on 20% of the protocols, was $\kappa = .82$.

Psychological knowledge. Following the moral judgment questions for each story, children were asked further questions to get at their knowledge of the actor's psychological experience. First, a straightforward understanding of the actor's intent was assessed using both open and forced-choice responses. Children were asked: "What do you think the actor was really trying to do?" followed by, "Do you think he/she was trying to hurt [their friend]?" Open-ended responses were coded for whether children attributed harmful (e.g., "get in a fight"; "boss her friend around") or nonharmful intentions (e.g., "just trying to get the swing"; "protect the girl in the tree") to the actor. A number of responses coded as "nonharmful" reflected attributions of self-serving motives, predominately for prototypical actors (e.g., "He wanted his turn on the swing"). However, the purpose of this question was to determine whether children attributed intentions other than causing negative consequences for the friend (psychological or physical) to actors. Interrater agreement for the coding of open-ended responses, based on 20% of the protocols, was perfect ($\kappa = 1.00$). Children who attributed harmful intentions to an actor in *either* their open-ended responses or by answering "yes" to the forced-choice question were assigned a score of 1, whereas children were given a score of 0 if they attributed nonharmful intentions to the actor and responded "no" to the forced-choice item. Given that children were scored as attributing harmful intentions to transgressors if they indicated this on either of the two assessments, the *harmful intent belief* item represents a conservative test of children's ability to correctly understand whether actors intended to cause harm.

Children's psychological knowledge of the actor's interpretation was assessed using an item from Killen et al. (2011) that asked: "Did [the actor] think he/she was doing something that was alright or not alright to do?" followed by "Did he/she think it was a little bad or very bad (just alright, a little good, or very good?)". Responses were scored on a 5-point *actor interpretation* scale ranging from 0 (*very bad*) to 4 (*very good*). Higher scores reflected the judgment that an actor would hold a more positive and accepting view of his or her own actions.

Table 1
Main Story Components

Condition	Friend's intended action	Harm prevented	Target of friend's harm
Prototypical	Swinging on a swing	^a	^a
Friend protect	Inexperienced climber ascending a ladder to a roof to retrieve a ball	Falling off of a roof	Friend's own self
Innocent protect	Shaking an inexperienced climber out of a tall tree	Falling out of a tall tree	Third party
Self-defense	Pushing someone underwater in the deep end of a pool	Getting dunked underwater	Actor

^a These components are not applicable in the prototypical condition.

Table 2
Justification Categories

Coding category	Description and examples
Actor transgression	References to the actor's harmful actions and its consequences (e.g., "She made her friend cry" or "It was mean for him to do that").
Friend transgression	References to the harm the friend would/could have caused (e.g., "He was trying to dunk him" or "Because that kid was going to fall out of the tree").
Actor positive act/intent	Specific references to the actor's intentions or their attempts to prevent harm or protect themselves or others (e.g., "He saved himself" or "She didn't want her friend to fall off of the roof").
Minimizing harm	References that explicitly minimized the actor's harmful actions or the consequences of their actions (e.g., "All he did was yell at him"; "It's not like she hit her"; "Hurting his friends' feelings isn't that big of a deal in the end).
Unnecessary harm	Responses stressing that the actor's actions were unnecessary and/or suggestions for other courses of action the actor could or should have taken (e.g., "He [the friend] wouldn't have hurt anyone but himself"; "She wasn't doing anything to deserve that").
Authority/punishment	Appeals to the existence of rules, punishment, or authority (e.g., "He'll get in trouble").
Relationship concerns	Statements that specifically implicate the friendship/relationship as the reason for an evaluation (e.g., "They won't be friends anymore").
Undifferentiated	Ambiguous or unelaborated responses (e.g., "I don't know" or "It's bad").
Other/uncodable	Responses that are not relevant, are incoherent, or do not fit into the above categories (e.g., "He's a silly face").

We hypothesized that higher scores would reflect more advanced psychological knowledge. To provide some support for this hypothesis, we conducted a pilot study using the same interview on a sample of 30 college students (47% female; $M_{\text{age}} = 20.27$ years, $SD = 1.14$). We found that 90%–93% of adult participants had actor interpretation scores of 2 (i.e., actors would evaluate their own actions to be "just alright") or higher across the necessary harm stories ($M_s = 3.48$ – 3.88 , $SD_s = .88$ – $.94$). These findings lend support to the assertion that individuals with a more mature psychological conception of others will rate the well-intended actors in the necessary harm stories as having more positive interpretations of their behavior.

Procedure

Permission to participate in the study was obtained from parents of children enrolled in kindergarten through fifth grade attending local after-school care centers and summer camp programs. Each child was read all four vignettes, with the type of harm used by the actor (psychological or physical) varying between participants. The three necessary harm stories were presented together, with the order of presentation counterbalanced. The prototypical story was also counterbalanced and was presented as either the first or last story. After each vignette and before conducting the assessments, children were required to correctly answer three memory checks to ensure comprehension of the story. Specifically, children were asked to identify the actor (e.g., "Can you point to Zoey?"), confirm their understanding of the actor's transgression (e.g., "What did Zoey do to her friend in the story?"), and acknowledge that the friend was successfully prevented from performing their action (e.g., "Did Zoey ever get dunked by her friend?"). Incorrect responses were corrected by recounting the relevant details of the story, and they were asked the questions again. The majority of children (68%) successfully passed all memory checks on the first trial. Approximately 28% (13 younger and eight older children) incorrectly answered one or two items, whereas 4% (three younger children) incorrectly answered more than two memory questions (range = 3–7). However, all children successfully passed all memory checks after the salient aspects of the story were repeated,

and children were retested. The moral judgment questions were presented first, followed by the psychological knowledge items. All questions were presented in the fixed order described above. Each interview took approximately 20 min to complete. Children's responses to all items were recorded by a research assistant.

Analysis Plan

Analyses were first conducted with each vignette (prototypical, friend protect, innocent protect, self-defense) treated as a four-level within-subject *target* factor. These four levels reflected the manipulation of the target of the friend's harm (see Table 1). However, these analyses revealed few consistent or significant differences between judgments in the three necessary harm vignettes. To simplify analyses, judgments were averaged across the three vignettes to create a composite *necessary harm* score for each dependent variable. Therefore, *story effects* refer to differences in judgments between the prototypical and necessary harm conditions. Additionally, approximately 5% (3/54) of the effects or interactions involving gender were significant, with no discernible pattern, and no significant effects of prototypical story order (i.e., whether it was presented first or last) were found; therefore, gender and prototypical story order were not considered further.

Unless otherwise noted, an analysis of covariance (ANCOVA) approach was used to examine the data. Researchers have argued against breaking continuous variables into discrete groups (e.g., younger and older age groups) to test moderation effects (Whisman & McClelland, 2005). ANCOVA was originally developed for experimental designs to provide a more powerful test of the effect of an independent variable on a dependent variable by eliminating or "controlling" for variations in participants' scores on a covariate. This is accomplished statistically by estimating the effects of the independent variable when scores on the covariate are 0. In this way, ANCOVA is comparable to multiple regression analyses when the independent variable of interest (i.e., age) is continuous and normally distributed (for a full discussion, see Rutherford, 2001). Thus, ANCOVA provided a powerful and appropriate analytic tool for examining our data.

Specifically, we examined judgments using a 2×2 mixed-model ANCOVA, with *harm type* (psychological, physical) as the between-subject condition, *story* (prototypical, necessary harm) as the within-subject condition, and *age* (centered at the mean, as recommended [Aiken & West, 1991; Whisman & McClelland, 2005]) as the predictor. All possible two- and three-way interaction terms involving harm type, story, and age were included in each of the analyses.

Significant Age \times Story interactions were probed by first assessing the effect of age within each story condition by estimating the unstandardized slope (b) and effect size (η_p^2) of age on judgments separately for prototypical and necessary harm. The b values represent the average amount of predicted change in judgments associated with a 1-year increase in age. Next, simple slopes analyses (Aiken & West, 1991) were conducted by rerunning the ANCOVAs separately, with age centered at low and high values (rather than the mean). This allowed us to compare younger and older children's ability to distinguish between prototypical and necessary harm by estimating these story effects at different age values, rather than collapsing children into different age groups and testing observed means. We chose to use 6 and 9 years of age to represent "low" and "high" values based on prior research showing shifts in children's moral judgments and psychological knowledge during this time (Smetana, 2006; Wainryb & Brehl, 2006), and because it roughly approximates the average age for younger and older children in our sample if we were to dichotomize age at the median. Significant Age \times Harm Type interactions were probed in a similar manner.

Results

Moral Judgments of Act Acceptability and Deserved Punishment

Overall descriptive statistics for the study variables are presented in Table 3. As expected, the separate mixed-model ANCOVAs revealed significant Age \times Story interactions for both judgments of act acceptability and deserved punishment, $F_s(1, 72) = 28.79, 15.17, p_s \leq .001, \eta_p^2_s = .29, .17$. The examination of the age effect within prototypical versus necessary harm stories revealed that with increasing age, children rated necessary harm to be less wrong ($b = -.17, \eta_p^2 = .11$) and less deserving of punishment ($b = -.15, \eta_p^2 = .15$), whereas the effects of age on judgments of prototypical harm were nonsignificant ($b_s = .05, .00; \eta_p^2_s = .03, .00; p_s = ns$, for acceptability and deserved punishment, respectively). As illustrated in Figure 1a and 1b, comparisons of the story effects at younger and older ages showed that older children made greater distinctions between well-intended and selfish harm than younger children in evaluations of acceptability: younger, $t(74) = 2.37, \eta_p^2 = .07, p = .02$; older, $t(74) = 10.12, \eta_p^2 = .58, p < .001$, and deserved punishment: younger, $t(74) = 2.10, \eta_p^2 = .06, p = .03$; older, $t(74) = 7.76, \eta_p^2 = .45, p < .001$. No significant effects or interactions involving harm type were found.

Acceptability Justifications

The proportion of children endorsing each category is presented in Table 3. The *actor transgression*, *friend transgression*, and *actor's positive act/intent* response categories accounted for over

Table 3
Means (and Standard Deviations) for Judgments, Acceptability Justifications, and Coordination Responses, by Story

Variable	Prototypical harm		Necessary harm	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Judgments ^a				
Act acceptability	1.71	(.48)	1.08	(.72)
Deserved punishment	1.42	(.72)	0.95	(.72)
Harmful intent beliefs	0.20	(.41)	0.13	(.27)
Actor interpretations	1.26	(1.27)	1.65	(1.16)
Acceptability justifications ^b				
Actor transgression		.83		.57
Friend transgression		.00		.24
Actor positive act/intent		.03		.29
Minimizing harm		.04		.04
Unnecessary harm		.13		.05
Authority/punishment		.07		.02
Relationships		.04		.04
Unelaborated		.04		.04
Other/uncodable		.03		.01
Coordination of concerns ^b		.00		.16

^a Act acceptability and deserved punishment judgments range from 0 to 2, with higher scores reflecting less permissible acts and greater deservingness of punishment. Harmful intent belief judgments range from 0 to 1, with higher scores reflecting greater intention to hurt or cause harm. Actor interpretation judgments range from 0 to 4, with higher scores reflecting actors' more positive views of their own actions. ^b Scores represent percentages.

85% of all justifications offered; analyses were therefore limited to these three categories. Additionally, the friend transgression and actor positive act/intent justifications were only endorsed with any frequency in the necessary harm condition. Thus, the analyses for those justification categories did not include story as a factor. The results are presented below for each category in turn.

Actor transgression. A significant Age \times Story interaction, $F(1, 72) = 24.45, p \leq .001, \eta_p^2 = .25$, revealed that references to the actor's transgression showed age-related increases in the prototypical condition ($b = .05; p = .03; \eta_p^2 = .06$), whereas endorsement of this category decreased with age in the necessary harm condition ($b = -.08; p \leq .001; \eta_p^2 = .15$). As illustrated in Figure 2, tests of the story effects at younger and older ages showed that whereas use of this category did not significantly differ by story for younger children, $t(74) = 0.32, p = .75$, older children were more likely to mention the actor's transgression in the prototypical condition compared with the necessary harm condition, $t(74) = 7.15, p < .001, \eta_p^2 = .42$. No significant effects or interactions involving harm type were found.

Actor positive act/intent. A significant Age \times Harm Type interaction, $F(1, 72) = 11.42, p \leq .001, \eta_p^2 = .14$, revealed that age was strongly associated with references to the positive or preventive nature of the actor's intentions and behavior when they engaged in psychological harm ($b = .16, p \leq .001, \eta_p^2 = .52$), whereas the age effect was not significant in the physical harm condition ($b = .04, p = .15, \eta_p^2 = .06$). As can be seen in Figure 3a, younger children rarely made use of this category, with no differences between the two harm type conditions, $t(74) = 0.78, p = .44, \eta_p^2 = .01$. In contrast, older children made use of this justification for both harm types, although they were significantly

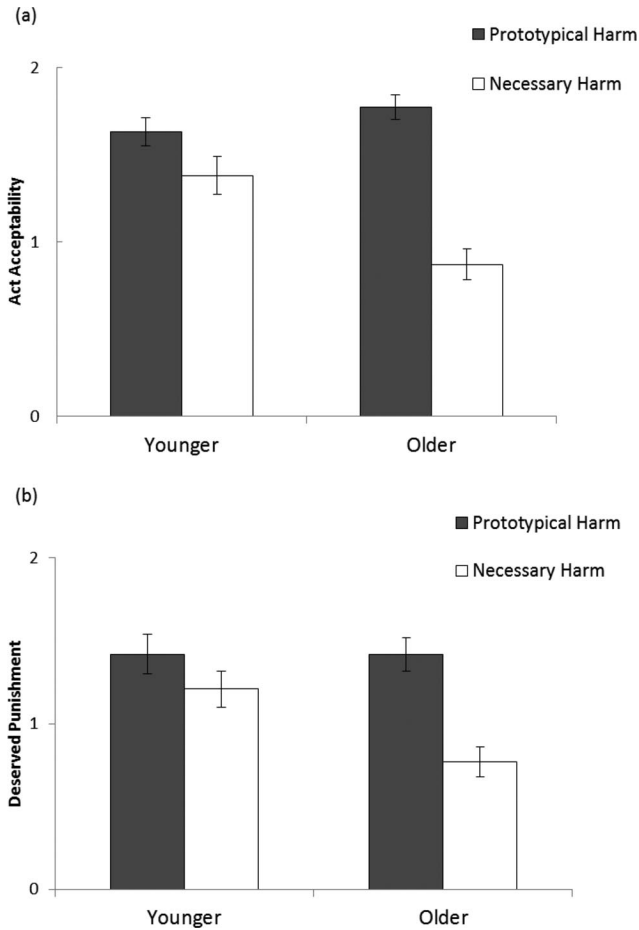


Figure 1. Estimated mean ratings for (a) act acceptability and (b) deserved punishment for younger and older children in the prototypical and necessary harm conditions. Error bars are standard errors.

more likely to do so in the psychological harm condition compared with the physical harm condition, $t(74) = 3.71, p < .001, \eta_p^2 = .16$.

Friend transgression. No significant age or harm type main effects or interactions were found for this justification category.

Coordination of Concerns

Children only coordinated concerns in the necessary harm condition, so the story factor (prototypical vs. necessary harm) was not included in the model. The analysis yielded a significant Age \times Harm Type interaction, $F(1, 72) = 10.74, p = .002, \eta_p^2 = .13$. Although age was positively associated with coordinated concerns in the psychological harm condition ($b = .08, p \leq .001, \eta_p^2 = .26$), age was not significantly associated with coordinated judgments of physical harm ($b = -.02, p = .32, \eta_p^2 = .03$). Interestingly, an examination of the effect of harm type at different ages (shown in Figure 3b) revealed that although older children were more likely to coordinate concerns in response to psychological compared with physical harm, $t(74) = 2.21, p = .03, \eta_p^2 = .06$, younger children did so more when actors engaged in physical compared with psychological harm, $t(74) = 1.94, p = .05, \eta_p^2 = .05$.

Psychological Knowledge

Descriptive statistics for harmful intent beliefs and actor interpretation ratings are presented in Table 3. As expected, only a minority of children attributed negative intentions to the well-intended actors. Interestingly, only about one fifth of children attributed harmful intentions to prototypical transgressors. Overall, children rated actors as holding fairly negative interpretations of their own actions.

Significant main effects of story revealed that actors in the necessary harm condition were attributed fewer harmful intentions, $F(1, 72) = 4.02, p = .05, \eta_p^2 = .05$, and were rated as having more positive interpretations of their actions than transgressors in the prototypical condition, $F(1, 72) = 4.95, p = .03, \eta_p^2 = .06$. Additionally, significant main effects of age for both questions revealed that with age, children were less likely to attribute harmful intentions to all actors, $F(1, 72) = 8.34, p = .005, b = -.05, \eta_p^2 = .10$, and instead rated them as having more positive interpretations of their own actions, $F(1, 72) = 6.13, p = .01, b = .18, \eta_p^2 = .08$. No significant effects or interactions involving harm type were found.

Associations Between Children's Moral Judgments and Psychological Knowledge

Preliminary analyses revealed that acceptability and punishment ratings were highly correlated for necessary harm ($r = .76$) and moderately correlated for prototypical harm ($r = .46$). Therefore, these two variables were combined and reverse coded to create a 3-point composite *forgiveness* variable (with higher scores indicating that a transgression was seen as more acceptable and the actor was less deserving of punishment) for necessary ($M = .99, SD = .68$) and prototypical harm ($M = .43, SD = .51$).

Bivariate correlations between moral judgments and psychological knowledge items by story condition are presented in Table 4. In both story conditions, actor interpretation ratings were negatively correlated with harmful intent beliefs. As hypothesized, children who were more forgiving of necessary harm were less likely to attribute harmful intentions to the actors and instead rated

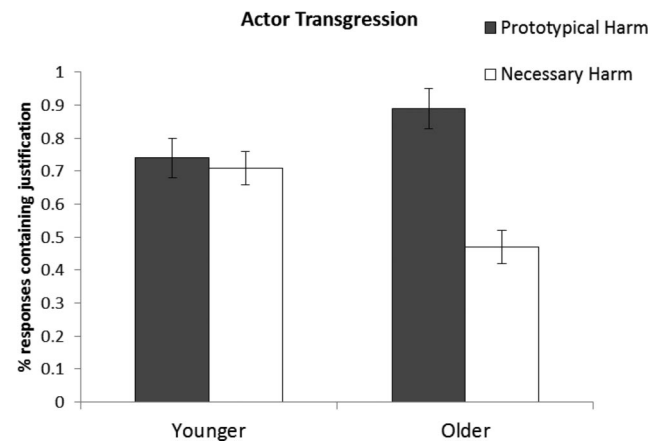


Figure 2. Estimated proportion of younger and older children referencing the actor's transgression in their justifications in the prototypical and necessary harm conditions. Error bars are standard errors.

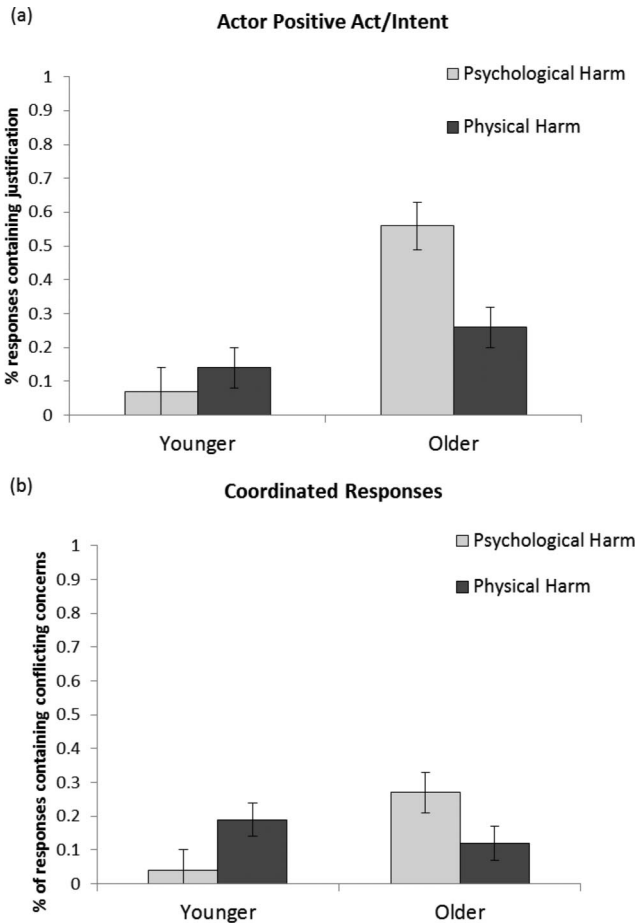


Figure 3. Estimated proportion of younger and older children in the psychological and physical harm conditions (a) referencing the actor's positive act/intent and (b) offering coordinated responses containing conflicting concerns in their justifications. Error bars are standard errors.

them as having more positive interpretations of their actions. As expected, harmful intent beliefs were negatively correlated with references to the actor's positive act/intent and coordinated responses in the necessary harm condition, whereas actor self-interpretations were positively correlated with use of these justification responses. Actor interpretation ratings were also

associated with fewer justifications pertaining to the actor's transgression.

Despite the consistent associations between actor interpretation ratings and more forgiving and flexible moral judgments of necessary harm, as expected, actor interpretation ratings were not significantly correlated with moral judgments in the prototypical condition. However, children who believed prototypical transgressors had harmful intentions were less forgiving of their actions.

Next, we tested the hypothesis that children's appreciation for the actor's psychological experience would be associated with more forgiving judgments, focusing less on the actor's harm and more on the prevention of harm. Controlling for age and simple intent knowledge, we ran five separate hierarchical regression analyses on forgiveness judgments and justification responses (i.e., actor transgressions, friend transgressions, actor positive act/intent, and coordination responses).

To control for the effects examined in the previous analyses, age, harm type (with psychological harm coded as -1 and physical harm coded as 1), harmful intent beliefs, and the Age \times Harm Type interaction term were entered in the first step; actor interpretation ratings were entered in the second step. The two- and three-way interactions between age, harm type, and actor interpretations were entered in a third step. No significant interactions involving actor interpretations were found for any of the models and were excluded from the analyses. All continuous variables were centered at their mean before being entered. The results of the regression analyses are presented in Table 5.

Controlling for the variables in the first step, actor interpretation ratings were positively associated with forgiveness of necessary harm and references to the actor's positive actions/intent, and negatively associated with references to the actors' transgressions. With Step 1 variables controlled, actor interpretation ratings were not significantly associated with references to the friend's transgression or with children's coordination of competing concerns.

Discussion

Many interpersonal conflicts require difficult choices between competing moral goals, where any decision may result in negative consequences (Wainryb et al., 2005). In the current study, we examined 5- to 11-year-olds' thinking about situations entailing necessary harm, where concerns about causing and preventing harm were in conflict. Consistent with recent conceptualizations of moral development (Helwig, 1995; Nucci & Turiel, 2009;

Table 4
Correlations Among Moral Judgment and Psychological Knowledge Variables, by Story

	Prototypical harm				Necessary harm						
	1	2	3	4	1	2	3	4	5	6	7
1. Harmful intent beliefs	—	-.26*	-.26*	-.21 [†]	—	-.41**	-.34*	.06	-.15	-.29*	-.25*
2. Actor interpretations		—	.00	.10		—	.62**	-.32**	.20 [†]	.45**	.28*
3. Forgiveness			—	-.23*			—	-.58**	.39**	.54**	.32**
4. Actor transgression				—				—	-.55**	-.24**	.10
5. Friend transgression									—	-.18	-.05
6. Actor positive act/intent										—	.61**
7. Coordination of concerns											—

[†] $p < .10$. * $p < .05$. ** $p < .01$.

Table 5
Hierarchical Regression Analyses Predicting Moral Evaluations of Necessary Harm From Psychological Knowledge

Variable	Forgiveness ratings				Actor transgression				Friend transgression				Actor positive act/intent				Coordinated responses			
	ΔR^2	<i>b</i>	β	<i>SE</i>	ΔR^2	<i>b</i>	β	<i>SE</i>	ΔR^2	<i>b</i>	β	<i>SE</i>	ΔR^2	<i>b</i>	β	<i>SE</i>	ΔR^2	<i>B</i>	β	<i>SE</i>
Step 1	.25**				.17**				.07				.40**				.20**			
Age		.10**	.27	.04		-.07**	-.37	.02		-.01	-.06	.02		.08**	.41	.02		.01	.09	.02
Harm type		.07	.10	.06		-.02	-.05	.04		.06	.18	.08		-.06	-.15	.03		.00	.00	.03
Harmful intent		-.07	-.03	.25		-.26	-.24	.17		-.09	-.13	.16		-.06	-.02	.14		-.07	-.14	.06
Age \times harm type		-.04	-.12	.03		-.02	-.10	.02		.03	.15	.02		-.06*	-.31	.02		-.05**	-.35	.02
Step 2	.22**				.07**				.04				.06**				.03			
Actor interpret		.32**	.54	.06		-.10**	-.30	.04		.06	.21	.04		.09**	.28	.03		.04	.19	.03
Total R^2		.47			.24				.10				.46				.23			

Note. The unstandardized slopes (*bs*), standardized regression coefficients (β s), and standard errors (*SEs*) are taken from the second step in each model. Harmful intent = harmful intent beliefs; Actor interpret = actor interpretations.

* $p < .05$. ** $p < .01$.

Smetana, 2006), we found that children's appreciation for such complex, multifaceted situations emerged gradually with age. Older children offered more forgiving evaluations and took into account the positive motivations of the well-intended actor. Importantly, although age-related increases in psychological knowledge were evident across both prototypical and necessary harm stories, how children used their understanding of others' minds to inform their moral choices depended on the concerns involved.

Age Differences in Moral Evaluations

In the current study, we extended previous research on positively motivated aggression (Darley et al., 1978; Darley & Shultz, 1990; Rule et al., 1975, 1974) by showing that both judgments and justifications of well-intended harm continue to develop beyond the preschool and early school years. Although younger children made rudimentary distinctions between prototypical and necessary harm in their judgments of acceptability and deserved punishment, with age participants offered increasingly more forgiving evaluations of necessary harm. Older children were less likely to focus on the well-intended actor's transgression and more likely to consider the positive or preventive nature of their actions. These results are similar to the findings obtained in studies of prosocial lie-telling (Heyman et al., 2009; Popliger et al., 2011; Xu et al., 2010).

However, older children were not simply more forgiving of harm in general. Regardless of age, all children judged prototypical harm to be very wrong and highly deserving of punishment. Additionally, older children did not typically view necessary harm as a positive or desirable act. They rarely judged these scenarios to be "good," and references to the actor's transgression were the most frequently used justification. It was also not the case that younger children were simply less aware of the potentially negative consequences of the friend's actions. No age differences were found in children's references to the friend's transgression (e.g., "She [friend] was going to shake that kid out of the tree"). This suggests that age differences in moral evaluations cannot be attributed to differences in children's basic understanding of the need to promote well-being or prevent harm, which develops during the toddler and early preschool years (Smetana & Braeges, 1990; Smetana, Rote, et al., 2012). Rather, variations in moral judgments appear to be related to differences in children's psy-

chological knowledge of others and in their emerging ability to simultaneously coordinate multiple, competing concerns.

Age Differences in Psychological Knowledge

Despite mean-level differences between the two story conditions, children rated both prototypical and well-intended actors as having more positive views of their own actions with age. Studying 3- to 7-year-olds, Killen et al. (2011) found similar age-related increases in the expectation that both accidental and malicious transgressors would have more accepting views of their actions. In making sense of their experiences, older children's perception of others as psychological agents acting on the basis of their personal goals may lead to the realization that even individuals engaging in seemingly straightforward moral transgressions may not see their own behavior in such a negative light. This is consistent with the findings from narrative studies by Gutzwiller-Helfenfinger, Gasser, and Malti (2010) and Wainryb et al. (2005) showing that older children and adolescents seldom construe their everyday moral experiences as straightforward or unconflicted. Similarly, social psychological research has shown that adults rarely judge others to hold beliefs that threaten their psychological well-being or conflict with their chosen actions (Ross & Ward, 1996). This was also supported in our pilot study, in which over 80% of college students stated that prototypical actors would hold neutral or positive views about their harmful behavior.

One implication is that the prototypical transgressions researchers typically use to assess preschoolers' moral judgments of intentional harm (Smetana, 2006) may be construed very differently by older individuals. Because harmful intent beliefs were ascribed to children if they answered in the affirmative on either the forced-choice or open-ended questions, we believe this represented a conservative test of simple intention knowledge. Nevertheless, 80% of children did not state that the prototypical actor's goal was to cause harm. In their study of 5- to 16-year-olds' narratives of past moral conflicts, Wainryb et al. (2005) found that 5-year-olds made uniformly negative moral judgments and focused primarily on the concrete harm that occurred (e.g., "I hit him; he cried"). In contrast, older participants' judgments were more forgiving and mitigated as they systematically attended to both the behavioral and psychological dimensions of their interpersonal conflicts.

Children may come to see even prototypical instances of harm in everyday contexts as somewhat ambiguous (see also Gutzwiller-Helfenfinger et al., 2010), highlighting the importance of considering interpretations and psychological attributions in studies of moral judgment. Beyond the preschool years, this likely involves going beyond assessments of discrete mental states such as beliefs or intentions (Carpendale & Chandler, 1996; Selman, 1980; Sokol et al., 2004).

Psychological Knowledge, Harm Type, and Moral Evaluations

Although global shifts in psychological knowledge may underlie changes in how children interpret both prototypical and well-intended transgressors' behavior, how children ultimately use this information in forming their own moral evaluations depends on the types of concerns involved. Children who rated well-intended actors as having more positive interpretations of their behavior were themselves more forgiving of necessary harm. They were also more likely to focus on the protective and prosocial aspects of the situation and less likely to focus on the actor's transgression. However, more advanced psychological knowledge does not guarantee more flexible moral judgments. An important finding was that actor self-evaluations were unrelated to children's evaluations of prototypical harm. Previous research has shown that children who are fully aware of the reasons for causing harm are nevertheless disapproving of those actions when performed for malevolent or immoral purposes (Wainryb & Brehl, 2006). As Wainryb and Brehl (2006) have noted, the ability to appreciate another's perspective does not imply that individuals will view that perspective as acceptable.

Unlike transgressions involving physical harm, psychological harm is inherently symbolic, requiring the victim to interpret a transgression to experience its consequences. Previous studies have found that children younger than 7–8 years of age show a less advanced understanding of psychological harm, resulting in more permissive views compared with physical aggression (Helwig et al., 1995, 2001). Although acceptability and deserved punishment judgments did not differ by harm type in the current study, strong differences emerged in justifications and coordination responses. Our finding that older children referenced the actor's positive goals and offered responses containing multiple, conflicting concerns was primarily for psychological harm. At first glance, these results appear to conflict with previous studies of psychological harm. Because harm type was treated as a between-subjects factor in our study, differences in evaluations were unlikely to stem from direct comparisons of the two types of harm.

With age, children are increasingly able to contrast an action with other possible behaviors an actor could or should have taken (Beck, Robinson, Carrol, & Apperly, 2006). In the current scenarios, older children may have interpreted the use of psychological harm as a deliberate choice to use a less serious form of aggression to accomplish the goal of preventing harm. This may have served to highlight their positive intentions, making them more readily salient. Importantly, however, beliefs about the actor's interpretations did not significantly differ by harm type, although the effect approached significance ($p = .06$). Null findings may have stemmed from a lack of power. Nevertheless, robust age effects for actor positive act/intent justifications and coordinated responses

were obtained for psychological harm, whereas this effect was nonsignificant for physical harm.

In contrast, few differences according to harm type emerged for younger children because they predominately focused on the actor's transgression and rarely justified their responses with appeals to positive intentions or attempts to prevent harm. This is consistent with studies showing that younger children are more heavily influenced by consequences rather than intentions when making moral judgments (Helwig et al., 1995; Killen et al., 2011; Nobes et al., 2009; Wainryb et al., 2005). Given that the observable consequences were similar for both physical and psychological harm (e.g., the friend did not cause harm and went home crying), our findings for younger children are not surprising.

One exception was that younger children coordinated concerns more often when evaluating physical than psychological harm. It may be that their more limited understanding of psychological harm (Helwig et al., 2001) precluded them from seeing how hurting someone's feelings—as opposed to physically stopping them—could legitimately prevent another's behavior, regardless of the actor's intent. This constrained view, coupled with younger children's difficulty in distinguishing between real and apparent emotions (Wainryb & Brehl, 2006), may have led them to view psychological harm as a less effective, yet equally harmful strategy for preventing their friend's actions. However, given the relative dearth of younger children's coordinated justifications, these effects need to be replicated before firm conclusions can be drawn.

The obtained findings for harm type suggest that moral judgment development cannot be reduced to global increases in psychological knowledge. Likewise, an understanding of another's psychological experience cannot be reduced to increased cognitive sophistication, as actor interpretations were no longer associated with coordinated responses after controlling for age. Rather, these abilities should be seen as tools that allow for the emergence of novel, more complex patterns of thought, and do not in themselves determine how children will evaluate a given situation. In our view, studies of moral development must consider the meanings and interpretations individuals give to their experiences. Although our results are in line with research showing that children come to offer more forgiving judgments of multifaceted situations with age (Smetana, 2006; Wainryb et al., 2005), this does not mean that moral development should be seen as a linear progression toward a shared moral point of view (Nucci & Turiel, 2009). Indeed, seemingly similar ratings of acceptability and punishment for psychological and physical harm stemmed from very different patterns of reasoning.

Limitations and Future Directions

Although we made a number of novel contributions in the current study, there are also some limitations. First, the sample size and relatively broad age range potentially limit the generalizability of our findings. This may have influenced the power necessary to detect small effects or higher order interactions, and may have prevented us from pinpointing specific ages at which certain types of judgment may emerge. However, treating age as a continuous predictor provided a more powerful test for our hypotheses than dichotomizing children into younger and older groups.

In the current study, the potential consequences of the friend's behaviors (e.g., falling off a roof or a tree, being dunked under-

water) were purposefully designed to be physically harmful acts that were more serious than the actor's attempts to prevent them. Additionally, the actor's motives were clear, and children were repeatedly reminded that the actor had failed to prevent harm peacefully before resorting to aggression. However, real-life situations of this type are inherently more ambiguous. Future research should explore how older children's increasingly more forgiving judgments may be moderated by the seriousness of the harm being prevented. Additionally, our control story was limited to a single vignette, as judgments regarding prototypical harm have been well studied in past research. Greater variation in judgments may have been found had we presented different types of selfish harm.

We chose to combine judgments across the three necessary harm stories because few significant differences emerged, although the types of concerns (e.g., protecting self vs. others) differed in important ways. Although we found strong age differences in the ability to distinguish necessary from prototypical harm, finer distinctions among different types of necessary harm may emerge later in development. For example, as autonomy and personal choice become salient concerns during adolescence (Nucci & Turiel, 2009), youth become less accepting of restrictions on another's self-injurious behavior (Flanagan, Stout, & Gallay, 2008). This was also supported in our pilot study; college students were less forgiving of preventing self-injury than other types of necessary harm.

Although psychological knowledge in our study focused on abilities often considered aspects of children's theory of mind, standard theory-of-mind tasks primarily assess abilities that develop prior to middle childhood (but see Carpendale & Chandler, 1996). Thus, we used a measure of actor interpretation drawn from research with older children (Killen et al., 2011) that produced robust associations with moral judgments. Nevertheless, there is some ambiguity as to which aspect of children's psychological knowledge it measures. Although Killen et al. (2011) used it to assess intention knowledge, the extent to which individuals interpret their own behavior as acceptable is influenced by a variety of factors other than their intent. Despite being well-meaning in their actions, individuals often hold ambivalent views toward their own interpersonal conflicts (Gutzwiller-Helfenfinger et al., 2010; Wainryb et al., 2005). Although our findings provide support for the utility of assessing beliefs about others' interpretations of their own behavior, more research is needed to clarify how those attributions relate to other aspects of their developing psychological knowledge.

Like others (e.g., Killen et al., 2011; Wainryb & Brehl, 2006), we focused on how increases in psychological knowledge may influence age-related changes in moral judgments. However, these statements must be viewed cautiously, as the cross-sectional design precludes any substantive claims regarding the direction of the effects. Furthermore, there is evidence to suggest that associations between knowledge in these two domains are part of a bidirectional and dynamic process in early childhood (Smetana, Jambon, et al., 2012). Longitudinal research is needed to examine how these processes unfold beyond the early childhood years.

These findings build on a large body of research from the social domain perspective demonstrating that children and adolescents make increasingly more sophisticated moral evaluations with age (Helwig, 1995; Nucci & Turiel, 2009; Smetana, 2006; Smetana, Jambon, et al., 2012; Wainryb et al., 2005). The ability and

willingness to consider the thoughts, beliefs, motives, and intentions of others can influence how individuals evaluate right and wrong. More research is needed to understand how this process generalizes to real-life interpersonal contexts. An avenue for future research would be to examine intrapersonal factors associated with variations in judgments. For instance, numerous studies have documented differences between reactively aggressive and nonaggressive youth in their judgments of the acceptability of retaliation and self-protective harm that stem from deficits in aggressive children's social information-processing abilities (Arsenio & Lemerise, 2004). However, we do not know whether these differences extend to situations in which aggressive acts are prosocial or aimed at protecting others. Additionally, past research has predominately focused on retaliation, which is qualitatively distinct from acting to prevent harm from occurring. It is possible that reactively aggressive youth may misconstrue well-meaning actors' intentions as negative, leading them to be less forgiving of this type of benevolent harm.

Individuals make moral decisions on the basis of their interpretations of the situation (Ross & Ward, 1996; Wainryb & Brehl, 2006). With age, children and adolescents come to hold a more sophisticated understanding of others. They become better able to coordinate and integrate psychological and moral elements, contributing to the emergence of more nuanced and complex patterns of moral understanding. More research is needed to understand how individual differences in these assumptions and attributions develop, and how they may contribute to normative variations in the ways children and adolescents navigate their complex moral worlds.

References

- Aiken, L., & West, S. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage.
- Arsenio, W., & Lemerise, E. (2004). Aggression and moral development: Integrating social information processing and moral domain models. *Child Development, 75*, 987–1002. doi:10.1111/j.1467-8624.2004.00720.x
- Beck, S., Robinson, E., Carrol, D., & Apperly, I. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development, 77*, 413–426. doi:10.1111/j.1467-8624.2006.00879.x
- Carpendale, J., & Chandler, M. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development, 67*, 1686–1706. doi:10.2307/1131725
- Colby, A., & Kohlberg, K. (Eds.). (1987). *The measurement of moral judgment* (Vols. 1–2). New York, NY: Cambridge University Press.
- Darley, J., Klosson, E., & Zanna, M. (1978). Intentions and their contexts in the moral judgments of children and adults. *Child Development, 49*, 66–74. doi:10.2307/1128594
- Darley, J., & Shultz, T. (1990). Moral rules: Their content and acquisition. *Annual Review of Psychology, 41*, 525–556. doi:10.1146/annurev.ps.41.020190.002521
- Flanagan, C. A., Stout, M., & Gallay, L. S. (2008). It's my body and none of your business: Developmental changes in adolescents' perceptions of rights concerning health. *Journal of Social Issues, 64*, 815–834. doi:10.1111/j.1540-4560.2008.00590.x
- Gutzwiller-Helfenfinger, E., Gasser, L., & Malti, T. (2010). Moral emotions and moral judgments in children's narratives: Comparing real-life and hypothetical transgressions. In B. Latzko & T. Malti (Eds.), *New directions for child and adolescent development* (pp. 11–31). San Francisco, CA: Jossey-Bass. doi:10.1002/cd.273

- Helwig, C. (1995). Social context in social cognition: Psychological harm and civil liberties. In M. Killen & D. Hart (Eds.), *Morality in everyday life: Developmental perspectives* (pp. 166–200). Cambridge, UK: Cambridge University Press.
- Helwig, C., Hildebrandt, C., & Turiel, E. (1995). Children's judgments about psychological harm in social context. *Child Development, 66*, 1680–1693. doi:10.2307/1131903
- Helwig, C., Zelazo, P., & Wilson, M. (2001). Children's judgments of psychological harm in normal and noncanonical situations. *Child Development, 72*, 66–81. doi:10.1111/1467-8624.00266
- Heyman, G., Sweet, M., & Lee, K. (2009). Children's reasoning about lie-telling and truth-telling in politeness contexts. *Social Development, 18*, 728–746. doi:10.1111/j.1467-9507.2008.00495.x
- Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. New York, NY: Cambridge University Press. doi:10.1017/CBO9780511805851
- Killen, M., Mulvey, K. L., Richardson, C., Jampol, N., & Woodward, A. (2011). The accidental transgressor: Morally-relevant theory of mind. *Cognition, 119*, 197–215. doi:10.1016/j.cognition.2011.01.006
- Lagattuta, K. (2005). When you shouldn't do what you want to do: Young children's understanding of desires, rules, and emotions. *Child Development, 76*, 713–733. doi:10.1111/j.1467-8624.2005.00873.x
- 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*, 382–397. doi:10.1016/j.jecp.2009.08.001
- Nucci, L., & Turiel, E. (2009). Capturing the complexity of moral development and education. *Mind, Brain, and Education, 3*, 151–159. doi:10.1111/j.1751-228X.2009.01065.x
- Popliger, M., Talwar, V., & Crossman, A. (2011). Predictors of children's prosocial lie-telling: Motivation, socialization variables, and moral understanding. *Journal of Experimental Child Psychology, 110*, 373–392. doi:10.1016/j.jecp.2011.05.003
- Ross, L., & Ward, A. (1996). Naïve realism in everyday life: Implications for social conflict and disagreement. In E. Reed, E. Turiel, & T. Brown (Eds.), *Values and knowledge* (pp. 103–136). Mahwah, NJ: Erlbaum.
- Rule, B., Dyck, R., McAra, M., & Nesdale, A. (1975). Judgments of aggression serving personal versus prosocial purposes. *Social Behavior and Personality, 3*, 55–63. doi:10.2224/sbp.1975.3.1.55
- Rule, B., Nesdale, A., & McAra, M. (1974). Children's reactions to the information about the intentions underlying an aggressive act. *Child Development, 45*, 794–798. doi:10.2307/1127848
- Rutherford, A. (2001). *Introducing ANOVA and ANCOVA: A GLM approach*. London, England: Sage.
- Selman, R. (1980). *The growth of interpersonal understanding: Clinical and developmental analyses*. New York, NY: Academic Press.
- Shaw, L., & Wainryb, C. (2006). When victims don't cry: Children's understandings of victimization, compliance, and subversion. *Child Development, 77*, 1050–1062. doi:10.1111/j.1467-8624.2006.00918.x
- Smetana, J. (2006). Social-cognitive domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (pp. 119–153). Mahwah, NJ: Erlbaum.
- Smetana, J., & Braeges, J. L. (1990). The development of toddlers' moral and conventional judgments. *Merrill-Palmer Quarterly, 36*, 329–346. doi:10.1037/0012-1649.25.4.499
- Smetana, J., Jambon, M., Conry-Murray, C., & Sturge-Apple, M. (2012). Reciprocal associations between young children's developing moral judgments and theory of mind. *Developmental Psychology, 48*, 1144–1155. doi:10.1037/a0025891
- Smetana, J., Rote, W., Jambon, M., Tasopoulos-Chan, M., Villalobos, M., & Comer, J. (2012). Developmental changes and individual differences in young children's moral judgments. *Child Development, 83*, 683–696. doi:10.1111/j.1467-8624.2011.01714
- Sokol, B., Chandler, M., & Jones, C. (2004). From mechanical to autonomous agency: The relationship between children's moral judgments and their developing theories of mind. In J. Baird & B. Sokol (Eds.), *Connections between theory of mind and sociomoral development* (pp. 19–36). San Francisco, CA: Jossey-Bass. doi:10.1002/cd.95
- Turiel, E. (2006). The development of morality. In N. Eisenberg (Ed) & W. Damon (Series Ed.), *Handbook of child psychology, 6th Edition, Volume 3: Social, emotional, and personality development* (pp. 789–857). New York, NY: Wiley.
- Wainryb, C., & Brehl, B. (2006). I thought she knew that would hurt my feelings: Developing psychological knowledge and moral thinking. In R. Kail (Ed), *Advances in child development and behavior* (pp. 131–171). New York, NY: Elsevier. doi:10.1016/S0065-2407(06)80006-6
- Wainryb, C., Brehl, B., & Matwin, S. (2005). Being hurt and hurting others: Children's narrative accounts and moral judgments of their own interpersonal conflicts. *Monographs of the Society for Research in Child Development, 70*(3, Serial No. 281). doi:10.1111/j.1540-5834.2005.00349
- Wainryb, C., & Ford, S. (1998). Young children's evaluations of acts based on beliefs different from their own. *Merrill-Palmer Quarterly, 44*, 484–503.
- Wellman, H., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*, 655–684. doi:10.1111/1467-8624.00304
- Whisman, M., & McClelland, G. (2005). Designing, testing, and interpreting interactions and moderator effects in family research. *Journal of Family Psychology, 19*, 111–120. doi:10.1037/0893-3200.19.1.111
- Xu, F., Bao, X., Fu, G., Talwar, V., & Lee, K. (2010). Lying and truth-telling in children: From concept to action. *Child Development, 81*, 581–596. doi:10.1111/j.1467-8624.2009.01417.x

Received August 12, 2012

Revision received January 14, 2013

Accepted March 26, 2013 ■