CHILD DEVELOPMENT

Child Development, xxxx 2017, Volume 00, Number 0, Pages 1–17

Individual Differences in Prototypical Moral and Conventional Judgments and Children's Proactive and Reactive Aggression

Marc Jambon and Judith G. Smetana
University of Rochester

This article examined links between 4- and 6-year-olds' (n = 101; M_{age} = 5.12 years, SD = 0.67; 53% male) ability to distinguish moral and conventional transgressions along different criteria and teacher ratings of proactive and reactive aggression. Latent difference score modeling revealed that moral transgressions were judged more unacceptable and wrong independent of rules and authority than conventional violations, but significant variability in moral–conventional distinctions was also observed. Proactive aggression was associated with less—and reactive aggression was associated with greater—differentiation in moral and conventional concepts. Proactive aggression was not associated with deficits in moral knowledge when other common assessments of early moral understanding were employed, highlighting the importance of using theoretically informed measures of moral judgments and aggression.

Young children consider moral transgressions involving harm and unfairness to be wrong based on the negative consequences for others (Smetana, Jambon, & Ball, 2014). Nevertheless, many children engage in behaviors that harm others (Eisner & Malti, 2015). Despite clear conceptual overlap, relatively little contemporary research has examined links between children's moral judgments and aggression (Arsenio & Lemerise, 2004). Furthermore, existing studies have typically failed to find meaningful associations between aggression and children's basic understanding of moral norms, leading researchers to downplay the importance of prototypical moral judgments for children's aggressive behavior. However, past research has often relied on conceptual definitions and methods incommensurate with contemporary approaches to moral development and aggression. Integrating concepts and methods from social domain theory (Smetana et al., 2014; Turiel, 1983) with functional approaches to aggression (Crick & Dodge, 1996; Little, Henrich, Jones, & Hawley, 2003), the present study examined whether individual differences in 4- to 6-year-olds' ability to distinguish morality from convention along theoretically derived criteria were differentially associated with proactive and reactive aggression.

Moral Judgments and Aggression

A central tenet of social domain theory (Smetana et al., 2014; Turiel, 1983) is that, through social interactions, children develop an understanding of morality (i.e., norms pertaining to issues of harm, justice, and rights) as distinct from social conventions (i.e., agreed-upon expectations and rules for acceptable behavior within a given social context). To test this assertion, domain researchers have drawn on philosophical and psychological definitions of morality and convention to examine whether children use different criteria to evaluate events in the two domains (referred to as criterion judgments). This research (reviewed in Killen & Smetana, 2015; Smetana et al., 2014) has shown that by 3–4 years of age, children judge straightforward, prototypical moral transgressions (e.g., hitting an innocent victim) to be *generalizably* wrong across different social contexts, even in the absence of rules (*rule independence*) and adult prohibitions (*authority independence*). In contrast, conventional violations (e.g., breaches of etiquette) are considered

Marc Jambon is now at the Department of Psychology, University of Calgary.

We would like to thank the day-care directors and school principals for their enthusiasm and cooperation with this project, as well as the teachers and children who participated. This research is based in part on the Marc Jambon's dissertation, submitted in partial fulfillment of the requirements for the PhD at the University of Rochester. This research was partially funded through a University of Rochester Departmental Dissertation Grant.

Correspondence concerning this article should be addressed to Marc Jambon, Department of Psychology, University of Calgary, Calgary, AB, Canada, 1235D, T2N 1N4. Electronic mail may be sent to marc.jambon@ucalgary.edu.

 0009-3920/2017/xxxx-xxxx
DOI: 10.1111/cdev.12757

acceptable if the context is changed or if the rules or prohibitions against them are removed.

Despite evidence demonstrating the early emergence of moral knowledge, many children nevertheless harm others. Individual differences in aggression are evident during the preschool years, become increasingly stable with age, and are associated with a wide range of psychosocial deficits (e.g., poor peer relationships, low academic achievement, maladjustment; Eisner & Malti, 2015). Social information processing (SIP; Dodge, Coie, & Lynam, 2006) research has also shown that biases in children's and adolescents' understanding and processing of social information contribute to the development and maintenance of aggression. For instance, aggressive youth are more sensitive to external threat cues than their nonaggressive peers, often believing that others possess hostile intentions in ambiguous conflict situations (i.e., hostile attribution biases).

Relatively less research has examined whether aggression is linked to variations in moral cognitions. Studies have commonly assessed moral evaluations in response to hypothetical stories depicting harmful behaviors in different situations. A consistent finding is that moral judgments of straightforward, unprovoked harm are not associated with individual differences in aggression (Astor, 1994; Gasser & Keller, 2009; Gasser, Malti, & Gutzwiller-Helfenfinger, 2012; Malti, Gasser, & Buchmann, 2009); both aggressive and nonaggressive children consider such acts to be unacceptable. When presented with more contextualized scenarios, however, aggressive youth are more likely to consider acts of harm in response to perceived threat or provocation to be acceptable and justified (Astor, 1994; Gasser et al., 2012; Huesmann & Guerra, 1997). Similar results have been observed in children and adolescents exposed to extreme political violence (Posada & Wainryb, 2008).

Researchers from diverse perspectives have drawn on these findings to conclude that, by the preschool years, nearly all children share the core moral belief that intentionally harming innocent others is categorically wrong (Arsenio & Lemerise, 2004; Gasser et al., 2012; Nunner-Winkler, 2013). Some have gone further to claim that judgments of prototypical moral transgressions reflect a purely cognitive or informational knowledge of rules that has no relevance for children's social behavior (Hawley & Geldhof, 2012; Nunner-Winkler, 2013). Consequently, the focus has shifted away from prototypical moral judgments to understand how other factors, such as social-cognitive biases, deficits in moral motivation and emotions, and atypical patterns of reasoning about complex moral events may lead children to engage in harmful acts despite "knowing" it is wrong to do so (Arsenio & Lemerise, 2004; Eisner & Malti, 2015).

The conclusion that all children (including those who harm others) possess the same basic moral knowledge may be premature, however, given the types of measures used in previous research to assess moral judgments. Specifically, studies examining links between prototypical moral judgments and aggression have typically asked children about the general acceptability of unprovoked moral violations (e.g., "is it ok to hit someone?" Astor, 1994; Huesmann & Guerra, 1997; Malti et al., 2009) or have assessed moral criterion judgments independent of conventional evaluations (e.g., "is it ok to hit if there were no rules/if a teacher allowed it?"; Gasser et al., 2012; Malti, Gasser, & Gutzwiller-Helfenfinger, 2010).

From a domain perspective, however, these assessments do not adequately capture children's understanding of morality as a distinct form of social knowledge. Because prototypical moral transgressions entail violations of accepted norms and practices, simply asking whether it is generally wrong to harm others does not allow us to determine whether children's judgments are based on moral or nonmoral concerns (e.g., fear of punishment; Killen & Smetana, 2015; Turiel, 1983). Obtaining justifications for judgments provides important insights when studying older youth, but preschool and early school-age children often have difficulty providing elaborated and codable responses (Smetana et al., 1999). Criterion judgments therefore provide an important window into young children's conceptual understanding of morality.

Yet focusing solely on criterion judgments of moral events—rather than the ability to differentiate morality and convention—may also limit our capacity to determine whether children understand moral obligations to be *distinct* from other norms. Even toddlers clearly recognize that moral violations are prohibited and considered unacceptable by peers and adults (Dunn, 2014). Given that young children already have difficulty with counterfactual reasoning and attributing unusual beliefs to others (Conry-Murray, 2013), a child who equates morality with external rules and authority commands may err on the side of caution and "disapprove" of harm even if told that the normal restrictions or adult prohibitions have been removed. Indeed, moral development researchers have long expressed similar concerns with socially desirable responding

when studying children's sociomoral judgments (Dunn, 2014; Piaget, 1932/1965; Rutland, Cameron, Milne, & McGeorge, 2005). However, a child who conflates moral and conventional rules is likely to evaluate both types of acts in a similar manner (Blair, 1995). Thus, focusing on the ability to *distinguish* moral and conventional events along different criteria may provide researchers with an important tool for addressing this challenge.

The hypothesis that aggression may be linked to deficits in children's moral understanding, operationalized as the capacity to differentiate morality and convention, implies that children differ from one another in this ability. To date, however, individual differences and their implications for social functioning have received little systematic attention. Deficits in prototypical domain-related judgments have been associated with disruptive and antisocial behaviors at later ages (Blair, 1995; Blair, Monson, & Frederickson, 2001; Nucci & Herman, 1982). Adolescents and adults who are high in psychopathic traits—characterized by empathic deficits and antisocial behavior—evaluate prototypical moral and conventional transgressions to be wrong and deserving of punishment, yet they have considerable difficulty distinguishing these events in their criterion judgments (Blair, 1995; Blair et al., 2001; Dolan & Fullam, 2010). Importantly, these deficits are not because psychopaths consistently treat acts of harm and injustice as conventional issues; several studies have found that psychopaths judge both moral and conventional violations to be categorically wrong even in the absence of rules and prohibitions (i.e., as moral issues; Blair, 1995; Dolan & Fullam, 2010). As Blair (1995) has noted, individuals who fail to distinguish moral and nonmoral rules may simply respond in a manner they deem socially desirable or expected.

These findings, although limited to atypical samples of adolescents and adults, lend support to the assertion that focusing on the ability to conceptually distinguish between morality and convention may provide a more appropriate assessment of children's basic moral knowledge than initial moral evaluations or moral criterion judgments alone. This ability is considered a milestone in children's social-cognitive development that is believed to provide an early foundation for later moral development (Killen & Smetana, 2015; Nucci, 2009; Smetana et al., 2014; Turiel, 1983). Research examining associations between domain distinction ability and aggressive behavior in younger, less extreme populations is greatly needed.

Functional Approaches to Aggression

In addition to using appropriate assessments of moral understanding, it is also critical to consider the underlying purpose, or function, that aggression may serve (Crick & Dodge, 1996; Little et al., 2003). Although past studies have often examined moral judgments in relation to children's general tendencies toward aggressive or disruptive behavior, differentiating between reactive and proactive functions of aggression may be crucial for understanding links with moral functioning (Arsenio & Lemerise, 2004). Reactive aggression reflects a "hot blooded" response to perceived threat or frustration, driven by intense affect and emotional dysregulation. In contrast, proactive aggression entails the "cold blooded," goal-directed use of aggression aimed at achieving rewards or dominating others.

Although the two types of aggression often cooccur (Cui, Colasante, Malti, Ribeaud, & Eisner, 2016; Poulin & Bouivin, 2000), they are nevertheless uniquely correlated with different social-cognitive mechanisms, affective processes, and sociobehavioral outcomes (Card & Little, 2006; Hubbard, McAuliffe, Morrow, & Romano, 2010; Vitaro, Brendgen, & Barker, 2006). For instance, reactive aggression is associated with physiological overarousal and dysregulated emotions (particularly anger), biased SIP (e.g., hostile attribution biases), and poor social relationships (e.g., peer rejection). These findings align with the traditional "social-skills deficit" conceptualization of antisocial behavior (Eisner & Malti, 2015; Sutton, Smith, & Swettenham, 1999). In contrast, when studies control for levels of reactive aggression, proactive aggression is often uncorrelated with indicators of maladjustment or is positively associated with social competence, such as intelligence, popularity, and social skills (see Card & Little, 2006). A number of scholars have drawn on evolutionary and ethological frameworks to suggest that the effective use of aggression to successfully obtain resources, rewards, and power in social groups can be considered adaptive and socially competent (Hawley, 2014; Stoltz, Cillessen, van den Berg, & Gommans, 2016; Sutton et al., 1999).

Moral Judgments and Proactive and Reactive Aggression

Although reactive aggression may be more closely tied to maladaptive social functioning than proactive aggression, the findings differ with regard to moral functioning. Children possessing social-

cognitive biases and dysregulated anger and frustration may respond to provocations and (mis)perceived threats with force, yet there is little evidence that reactively aggressive children lack empathy or a basic appreciation for the wrongness of unprovoked harm (Arsenio & Lemerise, 2004; Arsenio, Adams, & Gold, 2009; Orobio de Castro, Verhulp, & Runions, 2012). Indeed, when others' intentions and motivations are clear, typically developing children also consider aggression performed in self-defense or to protect others as more acceptable and justified than selfishly motivated harm (Jambon & Smetana, 2014). In contrast, studies have found that children and adolescents who engage in proactive aggression are more likely than other youth to prioritize their own selfish desires at the expense of others, have greater confidence in their ability to use aggression to obtain rewards, and are less likely to feel guilt or remorse after wrongdoing (Arsenio et al., 2009; Crick & Dodge, 1996; Hubbard et al., 2010). Longitudinal research has also found proactive aggression to be uniquely associated with a heightened risk of developing psychopathic traits and violent tendencies later in life (reviewed in Hubbard et al., 2010). Given that the deliberate use of coercion and harm for self-gain represents a prototypical violation of moral norms, children exhibiting proactively aggressive tendencies may also be less advanced in their understanding of morality. However, the few studies to have empirically tested this hypothesis have resulted in mixed findings.

Drawing on an evolutionary framework, Hawley and Geldhof (2012) examined links between preschoolers' use of coercive strategies to obtain and control resources (i.e., proactive aggression) and their moral functioning. They found that proactively aggressive youth displayed low levels of guilt and were adept at manipulating others (e.g., tattling; being nice to teachers after caught transgressing), yet they were also *more advanced* than nonaggressive children in their judgments of straightforward moral transgressions. The researchers concluded that proactive aggressors' more advanced moral knowledge reflects an "enlightened self-interest" that allows them to successfully navigate the social hierarchy. The findings regarding guilt were interpreted as evidence that emotions are more important than cognitions for guiding children's social behavior (also see Hawley, 2003).

Although this interpretation is intriguing, methodological limitations call it into question. Specifically, Hawley and Geldhof (2012) ascribed advanced moral reasoning to children who judged prototypical moral transgressions such as stealing to be unacceptable based on any explanation, including nonmoral ones, that could be classified as rule or emotion based ("it is against the rules"; "it was not nice"). In contrast, children were scored as less advanced if they judged transgressions to be permissible or could not provide an explanation. Given that preschoolers rarely judge prototypical moral violations to be generally acceptable and often have difficulty providing verbal justifications, these findings must be viewed with caution.

Although judgments were not assessed, Arsenio et al. (2009) examined whether proactive and reactive aggression were differentially associated with SIP abilities, moral emotion attributions, and reasoning in a sample of low socioeconomic status (SES) adolescents. Proactive aggression was associated with higher verbal skills, greater expectations for rewards and positive emotions following aggression, and fewer references to others' welfare when explaining their emotions. In contrast, reactive aggression was linked with greater hostile attribution biases and delays in attention and verbal ability, but it was not associated with deficits in moral emotions and reasoning.

Adopting a person-centered approach, a recent longitudinal investigation (Cui et al., 2016) examined whether moral reasoning was associated with distinct developmental trajectories of proactive and reactive aggression from middle childhood to early adolescence. The researchers found that low levels of moral reasoning at 8 years of age predicted membership in a class of youth exhibiting high, stable levels of reactive and proactive aggression over time. In contrast, participants who showed elevated rates of reactive (but not proactive) aggression did not differ from nonaggressive youth in their reasoning about prototypical moral violations.

These studies provide some support for the notion that proactive (but not reactive) aggression may be associated with deficits in moral knowledge, but this hypothesis remains to be tested with more precise and theoretically grounded measures of moral understanding. That proactive aggressors are often verbally and socially skilled (Arsenio et al., 2009; Hawley, 2014) may help to account for the more advanced moral judgments observed in preschoolers (Hawley, 2003; Hawley & Geldhof, 2012). Children who lack a basic appreciation for issues of harm, justice, or rights, yet who are attuned to the wants and needs of others when it suits their desires, may be particularly adept at "appearing moral" when assessments that fail to capture their conceptual understanding are used.

Studies employing appropriate methods are needed to better understand the role of moral judgments in the development of proactive and reactive aggression in early childhood. This is critically important, given that individual differences in moral cognitions and their implications for children's broader social functioning are poorly understood. Although antisocial behavior has been extensively studied in disadvantaged and high-risk populations, researchers have increasingly emphasized the importance of considering the developmental processes associated with aggressive and coercive behaviors in more normative, typically developing populations (Arsenio & Lemerise, 2004; Hawley, 2014). Gaining greater insight into the unique correlates of proactive and reactive aggression in early childhood may also help inform intervention efforts aimed at targeting the specific deficits underlying the aggression subtypes (Eisner & Malti, 2015; Vitaro et al., 2006).

The Current Study

The primary goal of the present study was to examine whether young children's ability to distinguish prototypical moral and conventional events along theoretical criteria was differentially associated with teacher ratings of proactive and reactive aggression. We focused on 4- to 6-year-olds because this constitutes an important period for the development of moral concepts (Smetana et al., 2014) and individual differences in proactive and reactive aggression (Murray-Close & Ostrov, 2009). Consistent with past social domain research, we expected that children would distinguish moral and conventional norms in their criterion judgments. Importantly, however, we also expected to find significant variability in this ability.

Past research has shown teacher reports of proactive and reactive aggression to be highly correlated (Card & Little, 2006; Poulin & Bouivin, 2000). Although the two subtypes often co-occur, much of this overlap stems from the fact that questionnaires assessing reactive and proactive functions use similar wording and include items reflecting the same underlying behavioral form (e.g., hitting; Little et al., 2003). Despite this methodological concern, the two functions consistently load as distinct constructs and, when controlling for their shared variance, demonstrate unique patterns of association with various behavioral and social-cognitive factors (Arsenio et al., 2009; Card & Little, 2006; Hubbard et al., 2010; Poulin & Bouivin, 2000). We therefore expected that, after controlling for reactive aggression, children rated higher in proactive aggression would make less of a distinction between morality and convention. However, we did not expect domain distinction ability to be associated with reactive aggression after controlling for proactive aggression.

A second important goal of the study was to examine whether domain distinction ability provided a more appropriate assessment of prototypical moral understanding compared to the two most common measures used in past research: (a) initial moral judgments and (b) moral criterion judgments excluding evaluations of conventions. Consistent with prior studies utilizing these methods (e.g., Astor, 1994; Gasser et al., 2012; Huesmann & Guerra, 1997; Malti et al., 2009), we did not expect either assessment to be significantly associated with proactive aggression.

Finally, prior research has found little evidence for gender differences in prototypical moral and conventional judgments (Smetana et al., 2014). Similarly, although boys typically exhibit higher mean levels of overt aggression than girls, links between aggression subtypes and other measures of children's social and emotional development also do not systematically differ by gender (Card & Little, 2006). We therefore tested effects for age and gender but did not expect to find differences in boys and girls or for younger and older children.

Method

Participants/Sample

The sample consisted of 101 four- to six-year-olds (53% male; M_{age} = 5.12 years, SD = 0.67) and their teachers (n = 18; 94% female; M_{age} = 35 years, SD = 15). Participants were recruited from nine schools and day-care centers serving lower to middle SES families in a moderately sized city in the northeastern United States. Data were collected between May 2014 and October 2015. The racial and ethnic composition of the sample reflected the demographics of the population from which they were drawn. Children were 65% White, 23% Black, 9% biracial, 1% Asian, and 2% unknown; 14% were of Hispanic/Latino(a) origin. Two additional child participants failed to pass the interview scale training (see below) on multiple occasions and were subsequently dropped from the study. Teachers were 75% White, 19% Black, and 6% Pacific Islander; 6% were Hispanic/Latino(a).

Social Rules Interview

Drawing on valid and reliable methods used in past social domain research (see Smetana et al., 2014), structured interviews were conducted to assess children's moral and conventional judgments. Interview scripts and all visual stimuli are included in Appendix S1.

Story Stimuli

Children were presented with four hypothetical vignettes depicting everyday moral and conventional transgressions (two each). Hypothetical transgressors (matched to the participant's gender) were depicted as having clearly defined intentions, and all events were described as occurring in a classroom setting.

Each vignette presentation consisted of four story pictures: rule introduction, actor goal, transgression, and outcome. The *rule introduction* picture showed the specific transgression and included a red "X" over the image. This was to ensure that all children understood that the behaviors (particularly the conventional events) were considered rule violations. Children were told that the act was prohibited (e.g., "In this class, kids are not supposed to hit."). The *actor goal* picture described character(s) and their motives (e.g., wanting a toy that another child is playing with). The *transgression* picture showed the character committing the transgression to achieve the goal (e.g., hitting the child to get the toy). The *outcome* picture showed the goal achieved and the resulting consequences (e.g., the victim was depicted as hurt, and the actor was shown holding the toy).

Moral stories included physical harm (*hit*: a child hitting another child to obtain a desired toy) and property theft (*steal*: a child taking food from another child after finishing their own snack). The conventional transgressions included violations of classroom organization (*backpack*: a child putting his or her backpack on the ground instead of placing it in a cubby to play with a desired toy) and school custom (*hand*: a child getting up to get another snack from the teacher's desk without raising their hand and asking permission first).

Initial Judgments

After each transgression, children's initial judgments of the *act's acceptability* were assessed dichotomously: "Is it ok or not ok for [transgressor] to ___?" This question was accompanied by two images, one of a green "thumbs up" reflecting a judgment that the act was acceptable (*ok*, scored as 0) and a red "thumbs down" reflecting a judgment that the act was not acceptable (*not ok*, scored as 1).

Initial *severity* ratings were then assessed by asking, "Do you think it was not bad for [transgressor] to ___, a little bad to do, or very bad to do?" The answer choices were represented by a "badness meter," depicted using three rectangles. *Not bad* was represented by an empty rectangle, *little bad* was represented by a rectangle half filled with orange, and *very bad* was depicted by a rectangle completely colored in. Responses were scored on a 3-point scale ranging from 0 (*not bad*) to 2 (*very bad*).

Criterion Judgments

Following initial judgments, additional picture slides depicting *authority* or *rule contingency* were presented. The pictures were designed to aid children's understanding of hypothetical, counterfactual scenarios where normal rules and expectations do not apply. Presenting children with detailed scenarios depicting rule and authority independence, rather than simply asking if it would be wrong if there were no rules/prohibitions, is consistent with the original procedures used by domain researchers to assess criterion judgments (Weston & Turiel, 1980). Children were presented with only one criterion per story because pilot testing indicated that assessing acceptability and severity judgments of both authority and rule contingency for each story, in addition to initial judgments, made the interview excessively long and tiresome. The criterion children received for a particular story was counterbalanced within domain. Thus, children evaluated one story within each domain involving authority contingency and one depicting rule contingency, resulting in a total of four different story–criterion combinations.

Authority contingency was illustrated by presenting an image of a female teacher shrugging her shoulders, with her palms turned up to indicate nonchalance. Above the teacher's head was a thought bubble depicting the target transgression, indicating that she was explicitly thinking about the act. This was accompanied by the verbal explanation:

> What if it was a different class with a different teacher. *This* teacher (point to teacher) does not care if kids ___ in her class. See, she doesn't

mind at all. So these questions are about what [transgressor] did if (s)he had *this* teacher who didn't care/mind and let kids ___.

For rule contingency, the rule introduction picture containing the red "X" from the first story slide was reintroduced at the top of the screen. Directly below this was the same picture, now without the red "X," positioned next to an image of a school. Children were told, "Some schools have rules against ___, and say that kids aren't supposed to ___" while the experimenter pointed to the top picture. Pointing to the bottom picture, the experimenter explained, "but in *this* school, there is *no* rule against ___. Kids are allowed to ___ if they want to. So these questions are about what [transgressor] did if (s)he went to *this* school where children are allowed to ___."

Acceptability and severity judgments (as described above) were then reassessed. Each judgment question was preceded by the statement, "If the teacher doesn't care/mind if kids ___ and lets them do it . . ." (for authority contingency), or "If there was no rule against ___and they let them do it . . ." (for rule contingency). For all judgments, higher scores indicated more "moral" evaluations (i.e., more unacceptable/wrong).

Aggression Measures

Reactive Aggression

Teachers reported on children's reactive aggression using six items from the physical and relational subscales of the Preschool Proactive and Reactive Aggression–Teacher Report (PPRA–T; Ostrov & Crick, 2007). Items were scored on a 5-point scale ranging from (1) *never or almost never true* to (5) *always or almost always true* of the child. The subscales demonstrated good internal consistency (reactive-physical $\alpha = .97$, reactive-relational $\alpha = .84$) and have been shown to reliably predict observational assessments of reactive aggression (Murray-Close & Ostrov, 2009; Ostrov & Crick, 2007).

Proactive Aggression

Teachers also completed proactive subscales from the PPRA–T. However, consistent with a recent study using the same measure (Poland, Monks, & Tsermentseli, 2016), preliminary analyses revealed that the proactive and reactive scales were highly correlated ($r = .93$) and did not load as separate factors. This likely stems from the high degree of similarity in item wording. For example, items reflecting reactive aggression include "if other children anger this child, s/he will often hit, kick, or punch them" (reactive-physical) and "when this child is upset with others s/he will often ignore or stop talking to them" (reactive-relational), whereas proactive items include "this child often hit, kicks, or punches to get what s/he wants" (proactive-physical) and "to get what this child wants, s/he often will ignore or stop talking to others" (proactive-relational). Given that the PPRA did not differentiate reactive and proactive functions, the proactive items were dropped from the study.

Proactive aggression was therefore assessed using teacher ratings of six items reflecting coercive resource control (Hawley, 2003; Hawley & Geldhof, 2012). Despite differences in terminology, coercive control and proactive aggression reflect the same underlying construct (i.e., instrumental aggression aimed at obtaining rewards). Furthermore, small differences in content and wording in the coercive control scale allowed us to model the proactive and reactive functions as separate factors. Items were scored on a 7-point scale ranging from (1) *hardly true* to (7) *mostly true* of the child, with higher scores reflecting a greater use of aggressive, coercive means of controlling others to accomplish goals (e.g., this child ". . .gets what s/he wants by bullying others"; ". . .tricks others to get what s/he wants"). The scale demonstrated good internal consistency ($\alpha = .93$) and was highly correlated with proactive aggression ratings from the PPRA–T ($r = .82$), but was less highly correlated with reactive aggression ($r = .78$) compared to the PPRA–T proactive scale.

Procedures

Children between 4 and 6 years of age attending local day cares and schools and their teachers were invited to participate in the study. Letters describing the study, along with parent permission forms, were given to all eligible families through their centers/schools. Participating children were interviewed individually in a separate area of their school or day care. Most teachers (14 of 18) completed online surveys regarding children's behavior; the rest completed paper surveys.

Child interviews were conducted on an iPad using the online survey software *SurveyGizmo*. This allowed for an interactive experience where children were able to point and choose different answer responses (described above). To train participants to use the device, children were first shown

each scale (e.g., "This is our 'badness' meter. You can use this to tell me how bad you think something is"), and the experimenter demonstrated its use. Children were then shown the scales again and asked (in a random order) to select pictures corresponding to different judgments. The interviews commenced only after children clearly understood how to use the scales.

All responses were automatically recorded online. Judgment questions were presented in the fixed order described above, whereas the software presented the stories and judgment response options in a randomized order. The specific story–criterion combination was counterbalanced within gender. Each interview took approximately 20–25 min, and participants were offered breaks as needed. Children were given a sticker after each story and received a colorful pencil at the end of the interview. Teachers received a small payment ranging from $10 to $30 depending on the number of surveys completed.

Analysis Plan

Unless otherwise noted, structural equation modeling with latent variables in M*plus* 7.3 (Muthén & Muthén, 1998-2012) was utilized to test our hypotheses. Adapting standard procedures for analyzing change scores, domain distinction ability was operationalized as a latent difference score (LDS; McArdle, 2009). This latent distinction construct represented the difference between criterion judgments of events in the two domains, with higher scores reflecting evaluations that moral transgressions were less acceptable and more wrong independent of rules and authority than conventional violations (see Appendix S2 for a complete description of the procedures and example M*plus* syntax).

LDS modeling provided a number of benefits over approaches commonly used in moral development research (e.g., analysis of variance, regression). We were able to explicitly test the hypothesis that variations in the ability to distinguish morality from social convention reflect meaningful individual differences rather than resulting from measurement error. It also allowed us to directly examine links between domain distinctions and aggression rather than requiring complex tests of interaction effects. Consistent with past research (Arsenio et al., 2009; Poulin & Bouivin, 2000), separate models were tested to compare the *unique* associations between judgments and proactive and reactive aggression while controlling for the nonfocal function.

Results

Preliminary Analyses of Moral and Conventional Judgments

Preliminary inspection of the data revealed that nearly all children initially evaluated both moral and conventional transgressions as unacceptable and bad. Scores for individual stories within domains were similar for both authority and rule contingency evaluations. Criterion judgments were therefore aggregated across stories within domain (see Table 1; complete descriptive statistics by story and criteria are presented in Appendix S3).

One-way analyses of variance examining whether criterion judgments differed as a function of the story–criterion combination revealed no significant effects (ps = .25–.97), suggesting that the specific combination of story and criteria (e.g., authority contingency of hitting vs. rule contingency of hitting) did not affect judgments. Additionally, authority and rule contingency judgments were significantly correlated for acceptability (moral $r = .54$; convention $r = .50$, ps < .001) and severity ratings (moral $r = .44$; conventional $r = .36$, ps < .001) and were therefore combined within domain. For ease of interpretability, criterion acceptability judgments were summed, whereas criterion severity judgments were averaged. Judgment ratings therefore ranged from 0 to 2, with higher scores reflecting acts were viewed as less acceptable/more wrong. These aggregated criterion judgments of acceptability and severity were used as manifest indicators in all subsequent analyses. Descriptive statistics and manifest correlations for all study variables can be found in Table 2.

Table 1
Descriptive Statistics for Moral and Conventional Criterion Judgments (Collapsed Across Story)

	Moral		Conventional	
Judgment	*M*	*SD*	*M*	*SD*
Authority contingency				
Acceptability[a]	0.55	0.50	0.41	0.50
Severity[b]	1.09	0.95	0.73	0.86
Rule contingency				
Acceptability[a]	0.56	0.50	0.48	0.50
Severity[b]	1.14	0.85	0.79	0.86

Note. [a]Scores range from 0 (*ok*) to 1 (*not ok*). [b]Scores range from 0 (*not bad*) to 2 (*very bad*).

Table 2
Descriptive Statistics and Bivariate Correlations for All Study Variables

	M	*SD*	Range	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age	5.11	0.68	4–6.80	1	−0.07	0.24*	0.12	0.07	0.11	0.08	−0.10	−0.01	0.02	−0.01	−0.02	0.03
2. Gender (male)	0.53	0.50	0–1		1	−0.01	0.07	−0.17†	−0.13	0.03	−0.13	0.17†	0.00	0.10	0.04	−0.11
3. Race (White)	0.68	0.42	0–1			1	0.13	−0.06	0.02	0.27*	−0.04	−0.05	0.07	0.13	0.02	0.05
4. Initial moral accept	1.93	0.32	0–2				1	0.43**	0.00	0.01	0.18†	0.13	−0.01	−0.16	−0.13	0.02
5. Initial conv accept	1.92	0.35	0–2					1	0.16	0.04	−0.04	0.12	−0.14	−0.13	−0.13	0.02
6. Initial moral sev	1.83	0.31	1–2						1	0.13	−0.14	−0.11	−0.07	−0.06	−0.12	0.09
7. Initial conv sev	1.52	0.55	0–2							1	0.04	0.05	0.25*	0.29**	−0.10	−0.13
8. Criterion moral accept	1.10	0.88	0–2								1	0.66**	0.68**	0.50**	−0.09	−0.14
9. Criterion conv accept	0.91	0.86	0–2									1	0.52**	0.56**	−0.13	−0.06
10. Criterion moral sev	1.12	0.76	0–2										1	0.65**	−0.07	−0.11
11. Criterion conv sev	0.76	0.71	0–2											1	−0.08	−0.14
12. Reactive aggression	2.25	1.00	1–5												1	0.78**
13. Proactive aggression	2.08	1.42	1–7													1

Note. All estimates based on observed manifest variables. Accept = acceptability judgments; sev = severity judgments; conv = conventional. Gender: 0 = female, 1 = male. Race: 0 = non-White, 1 = White. *$p < .05$. **$p < .01$. †$p < .10$.

Missing Data

There was relatively little missing data (range = 0%–7.8%), either in child interviews or teacher reports. Six children left their day cares before interviews could be conducted, and five additional children did not have teacher reports (three due to teachers not completing surveys, and two because teachers left the center before data could be collected and did not provide contact information). Little's missing completely at random (MCAR) test was nonsignificant, $\chi^2(59, N = 101) = 48.71$, $p = .828$. Missing data were therefore accounted for based on the MCAR assumption using the full information maximum likelihood (FIML) estimator with robust standard errors.

Links Between Domain Distinctions and Aggression

Domain Differentiation as a Latent Construct

A confirmatory factor analysis with two latent factors was conducted on moral and conventional criterion judgments. Each factor consisted of two indicators (criterion acceptability and severity). This initial model provided a relatively poor fit to the data, $\chi^2(3) = 19.26$, $p < .001$, root mean square error of approximation (RMSEA) = 0.239, comparative fit index (CFI) = 0.862, standardized root mean square residual (SRMR) = 0.053. Modification indices suggested that moral and conventional acceptability judgments shared considerable error variance. Because these evaluations were measured using different stimuli as the severity ratings, the error terms were allowed to covary. This modified two-factor model fit the data well, $\chi^2(2) = 1.25$, $p = .58$, RMSEA = 0.00, CFI = 1.00, SRMR = 0.031, and provided a significantly better fit to the data than a one-factor model, $\Delta\chi^2(1) = 10.76$, $p = .001$. An LDS model was then estimated to capture children's ability to differentiate morality and convention (see Appendix S2).

Parameter estimates for the LDS model are shown in Figure 1a. Consistent with previous domain research, the latent mean of the distinction factor was significantly different from zero (Cohen's $d = .94$, $p < .001$), confirming that overall, children significantly distinguished between morality and convention in their criterion judgments. However, the variance of the distinction factor was also significantly different from zero ($p = .006$), providing support for the hypothesis that children would differ from one another in this ability. Conventional judgments were not significantly correlated with

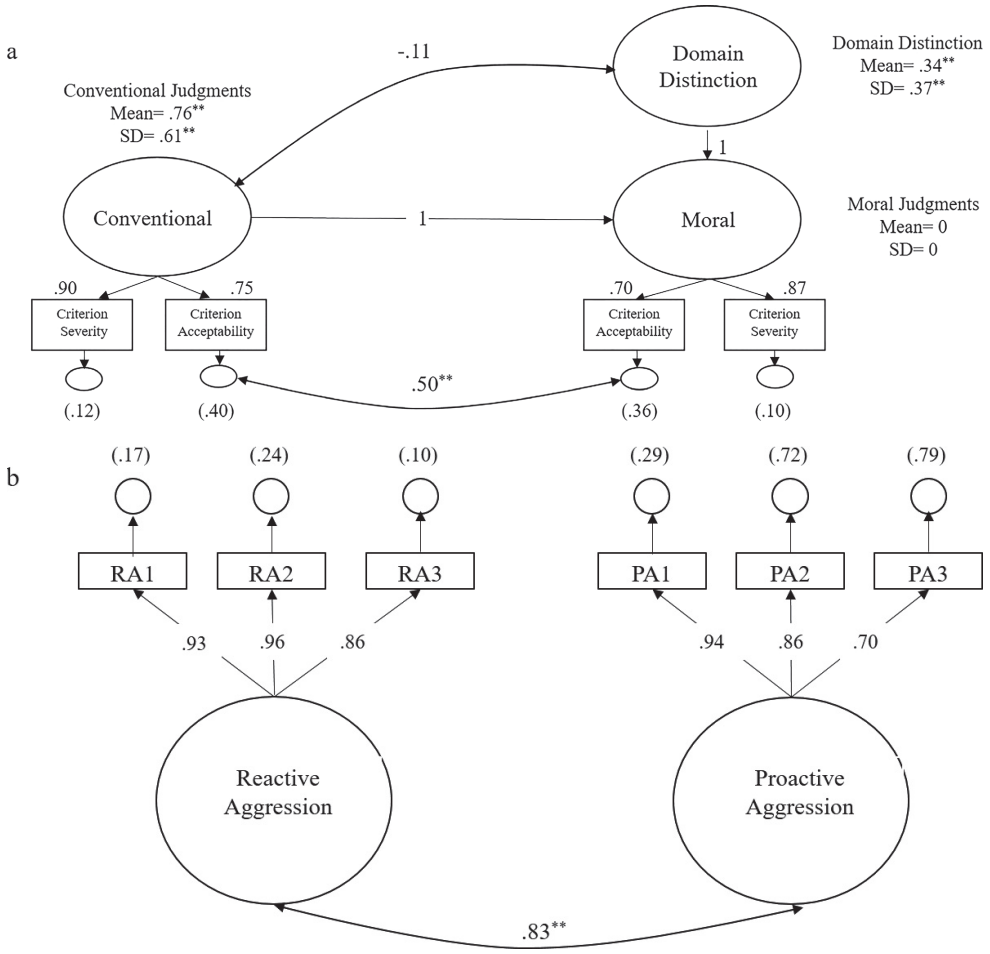

Figure 1. Measurement models for (a) latent moral/conventional difference scores and (b) reactive and proactive aggression. Factor loadings and covariance estimates are standardized. Residual variance estimates are unstandardized. **$p < .01$.

domain distinction scores ($p = .48$). The inclusion of demographic covariates revealed that conventional judgments and domain distinction ability did not significantly differ according to gender (conventional: $r = .02$, $p = .55$; distinction: $r = .17$, $p = .26$), race (conventional: $r = .10$, $p = .37$; distinction: $r = -.01$, $p = .94$), or age (conventional: $r = .00$, $p = .98$; distinction: $r = -.03$, $p = .79$).

Aggression Measurement Model

Given the modest sample size and the large number of items, parceling was used to create aggregated manifest indicators for proactive and reactive aggression (Little, Rhemtulla, Gibson, & Shoemann, 2013). Because the focus of the present study was on the different functions, items with similar wording from the physical and relational subscales of reactive aggression were combined to create three manifest parcels. Similarly, theory and modification indices were used to create three manifest parcels reflecting proactive aggression (see Appendix S4 for a description of items and parceling procedures).

As shown in Figure 1b, the resulting two-factor model evidenced adequate fit: $\chi^2(8) = 12.68$, $p = .12$, RMSEA = 0.078, CFI = 0.98, SRMR = 0.03, and provided a significantly better fit than a one-factor model, $\Delta\chi^2(1) = 8.45$, $p = .004$. The inclusion of demographic covariates revealed that neither function differed by gender (proactive: $r = -.07$, $p = .43$; reactive: $r = -.03$, $p = .70$), race (proactive: $r = -.09$, $p = .88$; reactive: $r = .04$, $p =. 80$), or age (proactive: $r = -.07$, $p = .54$; reactive: $r = -.03$, $p = .78$). These variables were subsequently dropped from the analyses.

Domain Distinctions and Aggression (Models 1 and 2)

Parameter estimates for all structural models can be found in Table 3. To test the hypothesis that domain distinctions would be uniquely associated

with proactive aggression, this factor was regressed onto domain distinctions, conventional judgments, and reactive aggression (see Figure 2a). Covariances were estimated between all predictors. Conventional judgments were not associated with proactive aggression ($\beta = -.01$) and the path was constrained to zero.

As expected, after controlling for reactive aggression, domain distinctions were negatively associated with proactive aggression ($p = .014$); children rated as more proactively aggressive were less able to differentiate moral and conventional transgressions in their criterion judgments. We then tested a second model predicting reactive aggression scores from proactive aggression and domain distinctions (see Figure 2b). Unexpectedly, after controlling for proactive aggression, domain distinction ability was positively associated with reactive aggression ($p = .046$); children rated as more reactively aggressive were better able to differentiate moral and conventional transgressions in their criterion judgments.

Follow-up analyses were conducted to examine whether associations between domain distinctions and aggression differed by age or gender. The latent variable interaction for age was tested using the XWITH command in M*plus*, whereas multigroup analyses tested for gender differences. Effects did not differ by age or gender (ps > .30).

Comparing Domain Distinction Ability to Other Common Judgment Assessments

To test the second aim of the study—whether domain differentiation provided a more appropriate assessment of children's prototypical moral understanding than assessments commonly employed in past research—two additional models examining links between proactive aggression and initial moral judgments and moral criterion judgments were estimated (see Table 3).

Initial Moral Judgments (Model 3)

Only four children initially judged either of the moral transgressions to be acceptable. Given this lack of variability, initial moral severity ratings (averaged across the hitting and stealing stories) were used. Proactive aggression was regressed on reactive aggression and observed initial moral severity scores, $\chi^2(12) = 17.03$, $p = .15$, RMSEA = 0.06, CFI = 0.99, SRMR = 0.03. Interestingly, initial moral severity ratings were positively associated with proactive aggression ($p = .001$); although most children judged moral transgressions to be wrong, proactively aggressive children rated them to be more wrong than less aggressive youth.

Moral Criterion Judgments (Model 4)

Proactive aggression was then regressed on reactive aggression and a latent variable representing moral criterion judgments, $\chi^2(18) = 18.18$, $p = .15$, RMSEA = 0.01, CFI = 1.00, SRMR = 0.03. Moral criterion judgments alone were not significantly associated with proactive aggression ($p = .19$).

Discussion

Drawing from a community sample of diverse, lower to upper middle SES 4- to 6-year-olds, we

Table 3
Parameter Estimates for Structural Models

	Unstandardized				Standardized		
	Coefficient (*b*)	Standard error	Lower $CI_{.95}$	Upper $CI_{.95}$	Coefficient (β)	Lower $CI_{.95}$	Upper $CI_{.95}$
Model 1							
Domain dist → Proactive	−0.916	0.401	−1.580	−0.260	−.220	−.368	−.073
Reactive → Proactive	1.196	0.139	0.967	1.425	.835	.738	.933
Model 2							
Domain dist → Reactive	0.532	0.276	0.077	0.987	.184	.032	.337
Proactive → Reactive	0.613	0.052	0.527	0.699	.879	.790	.984
Model 3							
Initial moral sev → Proactive	0.930	0.256	0.509	1.35	.186	.089	.283
Model 4							
Moral crit → Proactive	−0.221	0.090	−0.239	0.058	−.096	−.215	.024

Note. Dist = distinction; proactive = proactive aggression; reactive = reactive aggression; sev = severity; crit = criterion judgments.

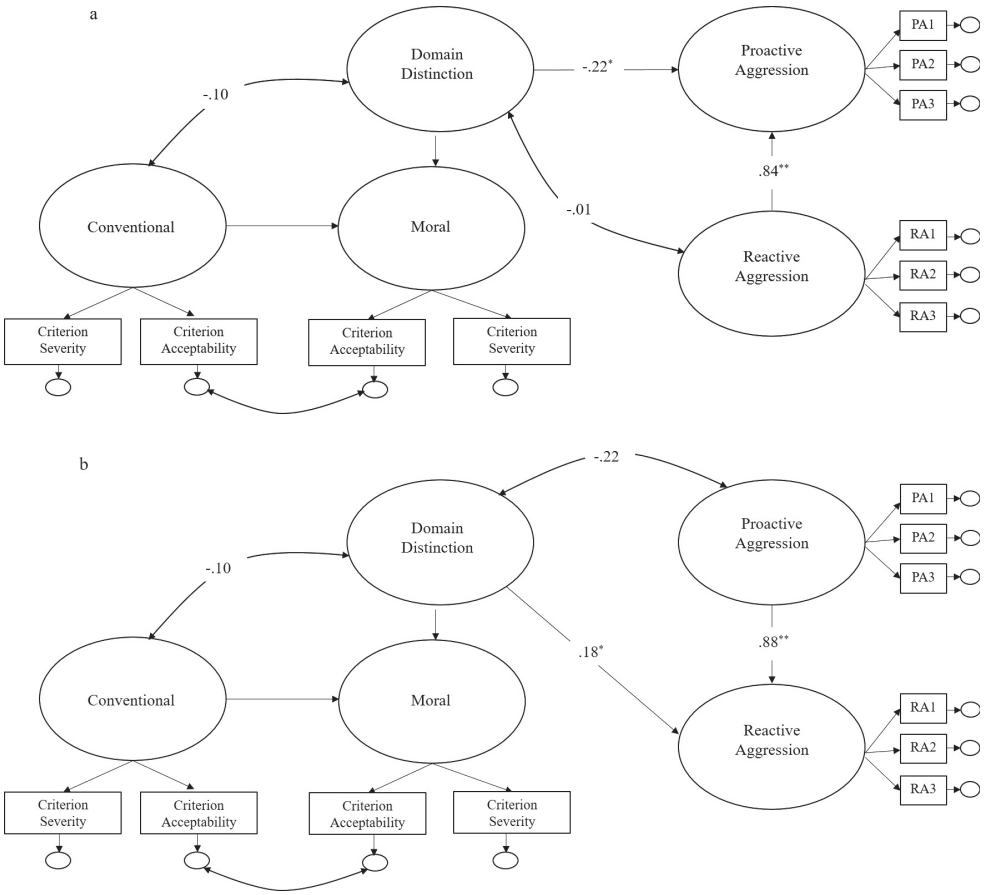

Figure 2. Structural models predicting (a) proactive aggression and (b) reactive aggression from domain distinctions. All parameter estimates are standardized. The covariance paths between the conventional factor and (a) reactive aggression ($r = -.12$, $p = .13$) and (b) proactive aggression ($r = -.09$, $p = .13$) were included in the models but are not depicted in the figure. *$p < .05$. **$p < .01$.

employed theoretically derived and validated measures to test the hypothesis that variations in young children's ability to distinguish prototypical moral and conventional events along theoretical criteria would be differentially associated with teacher ratings of proactive and reactive aggression. Proactive aggression was uniquely associated with deficits in children's ability to differentiate moral and conventional events, whereas, unexpectedly, reactive aggression was linked with more advanced moral judgments. However, proactive aggression was not associated with deficits in moral cognitions when other commonly used measures of moral knowledge were employed.

Individual Differences in Moral and Conventional Concepts

The understanding that morality represents a distinct domain of social knowledge is considered a normative, universal feature of young children's social-cognitive development (Killen & Smetana, 2015). Due to the limited cognitive and linguistic abilities of preschool and early school-aged children, criterion judgments represent the best available method of assessing this conceptual distinction in early childhood. Consistent with a large body of past social domain research, we found that when comparing mean levels, children distinguished moral and conventional transgressions in their judgments. However, we also found significant interindividual variation in this ability. Although little attention has focused on individual differences in early moral judgments, this finding is nevertheless fully consistent with a constructivist view of development.

Children are predisposed toward care and concern for others from the earliest of ages (Eisenberg, Spinrad, & Morris, 2014), and observational research has shown that naturally occurring moral and conventional events elicit different types of reactions from peers and adults (Smetana et al.,

2014). Resulting from these experiences, children develop different ways of thinking about morality as compared to other social norms (Turiel, 1983). However, children also differ markedly from one another in their capacity for empathy (Eisenberg et al., 2014) and the ability to read and pick up on social cues (Dodge et al., 2006), which may have important implications for how they interpret and process morally relevant information. The use of LDS modeling to capture variations in children's domain distinction ability could provide future researchers with a powerful and flexible tool to examine the developmental processes associated with individual differences in early moral understanding.

Moral Judgments and Proactive and Reactive Aggression

From a moral developmental perspective, intentionally harming others for self-gain constitutes an underdeveloped moral orientation. Consistent with this assertion, we found that proactive aggression was associated with a reduced capacity to differentiate moral and conventional norms, an ability that typically emerges during the preschool years and provides an early foundation for later moral development (Killen & Smetana, 2015; Nucci, 2009; Smetana et al., 2014; Turiel, 1983). Nevertheless, past studies have found that proactively aggressive children are often powerful and popular within their peer groups (Hawley, 2014; Stoltz et al., 2016), and may not exhibit the same signs of maladjustment observed in other aggressive youth (Poulin & Bouivin, 2000; Sutton et al., 1999). These findings have led theorists to construe some types of aggression, when flexibly employed to accomplish goals and obtain resources, as reflecting social competence and adaptation (Hawley, 2014). Although this perspective has garnered strong empirical support and provides an important alternative to the traditional deficit-based view of aggression (Sutton et al., 1999), it fails to adequately account for the negative impact such behaviors have on others. Given that proactive aggression may be uniquely associated with serious violence, delinquency, and psychopathic traits later in life (Hubbard et al., 2010), our findings underscore the need to consider moral functioning as a critical yet distinct aspect of early social competence (Arsenio & Lemerise, 2004; Gasser & Keller, 2009).

These results have important theoretical implications for how researchers conceptualize links between morality and aggression. Moral development researchers have shifted away from studying prototypical moral judgments to understand why some children deliberately harm others despite "knowing" it is wrong to do. Although other factors (e.g., empathic deficits, atypical patterns of reasoning regarding complex moral dilemmas) are certainly important to consider (Arsenio et al., 2009; Eisner & Malti, 2015), our findings call into question the assumption (Arsenio & Lemerise, 2004; Nunner-Winkler, 2013) that proactive aggressors possess a similar understanding of morality as less aggressive children. This also provides support for the assertion that prototypical moral judgments do reflect sincere beliefs regarding obligations to refrain from harm (Smetana et al., 2014) rather than a purely cognitive knowledge of moral rules (Hawley & Geldhof, 2012; Nunner-Winkler, 2013).

Because early detection and intervention is key to preventing problematic behavioral trajectories from emerging (Vitaro et al., 2006), these results may also help to inform efforts to prevent or reduce proactive aggression in young children. To date, researchers have stressed the need for intervention strategies that specifically target empathic deficits unique to proactive aggression (Arsenio & Lemerise, 2004; Hubbard et al., 2010). Our findings suggest that focusing on children's thinking and reasoning about issues of harm, justice, and rights may also be important, particularly during the preschool and early school-age years when moral concepts and individual differences in aggression first emerge. For instance, parents and educators could be trained to identify and respond to different types of misbehaviors with domain-appropriate messages that explicitly highlight the moral or conventional nature of the transgressions (Nucci, 2009). Additionally, controlled interventions utilizing concepts and methods from contemporary moral development research could provide invaluable information regarding the extent to which moral judgments and reasoning play a causal role in children's social behavior (Eisner & Malti, 2015; Vitaro et al., 2006).

In contrast to the findings for proactive aggression, we found that children rated higher in reactive aggression were better able to distinguish moral and conventional events (after controlling for proactive aggression). Although this finding was unexpected, it may stem from the types of social interactions and social-cognitive biases that are indicative of reactively aggressive youth. Resulting from past experiences (e.g., victimization, harsh family environments), reactive aggressors are particularly likely to believe that others possess hostile

intentions and retaliate with force (Dodge et al., 2006). Given that moral experiences (e.g., with aggression, unfairness) are believed to facilitate the emergence of moral concepts (Smetana et al., 2014; Turiel, 1983), reactive aggressors may be uniquely attuned to the implications and wrongness of unprovoked harm. Indeed, reactively aggressive youth often consider aggression in response to provocation and threat as a necessary moral imperative aimed at what they perceive as an injustice (Orobio de Castro et al., 2012). Given the relatively small effect size and unexpected nature of the finding, however, future replication is needed before drawing any firm conclusions.

Nevertheless, greater research is needed to understand how everyday experiences of conflict, victimization, and aggression color children's conceptions of right and wrong. To date, moral development and aggression researchers have primarily relied on hypothetical stories to document children's social and moral reasoning (Dodge et al., 2006; Killen & Smetana, 2015). It is less clear how children and adolescents understand their own harmful behaviors after they have occurred or whether adult and peer responses influence this interpretive process. Because engaging in actions that negatively affect others is an inevitable part of social life, the extent to which individuals self-reflect on their wrongdoing is a critical yet understudied aspect of moral development (see Wainryb & Recchia, 2014).

Assessing Prototypical Moral Judgments

Our findings underscore the importance of considering moral evaluations in a theoretically consistent manner. In line with with past research, few children stated that moral transgressions were generally acceptable, and moral criterion judgments were not significantly associated with proactive aggression. Consistent with Hawley's findings (Hawley, 2003; Hawley & Geldhof, 2012), however, we found that proactively aggressive children initially rated moral transgressions as *more wrong* than less aggressive youth. This may reflect the nature of proactive aggression and the social desirability concerns inherent in asking children about socially unacceptable behaviors and beliefs (Dunn, 2014; Piaget, 1932/1965; Rutland et al., 2005).

Proactively aggressive children are often intelligent and socially skilled. Thus, they may know how to appear "moral" and maintain that acts of harm would be wrong even if external sanctions or rules were removed. Yet they appear to overgeneralize this strategy when evaluating conventional events as well. This interpretation is consistent with research showing that adolescents and adults high in psychopathy may at times treat all rule violations as morally wrong (Blair, 1995; Dolan & Fullam, 2010). Operationalizing young children's rudimentary moral understanding as the ability to differentiate moral and conventional norms may therefore provide a useful method for countering social desirability concerns.

Limitations and Future Directions

Although the present study addresses important gaps in our understanding of moral judgments and aggression, several limitations should be mentioned. First, aggression was assessed using teacher ratings. Although teacher reports have been shown to reliably predict observational measures of aggression subtypes (Murray-Close & Ostrov, 2009; Ostrov & Crick, 2007), future research should include other informants. Additionally, coercive control was used to operationalize proactive aggression because the PPRA did not adequately distinguish between the two functions. Given that coercive control and proactive aggression reflect the same construct (aggression aimed at obtaining rewards), are measured using similar items (harming others "to get what they want"), and coercive control—but not the PPRA proactive items—loaded separately from reactive aggression, we believe this decision was justified. Nevertheless, research that employs both experimental and observational assessments is warranted.

Given the ubiquity of questionnaire assessments of children's social behavior in developmental research, however, more valid and reliable questionnaires aimed at differentiating between aggressive subtypes are needed (Hubbard et al., 2010). Although procedures have been developed to limit the confounding of aggressive forms and functions in behavioral ratings (Little et al., 2003), the calculating, callous, and selfish nature of proactive aggression is largely absent in existing measures asking whether children behave aggressively to "get what they want." A child who angrily lashes out against a threatening peer has certainly accomplished what they set out to do, yet this does not constitute an act of proactive or instrumental aggression as it is typically conceptualized (Crick & Dodge, 1996; Hubbard et al., 2010; Little et al., 2003; Murray-Close & Ostrov, 2009). To better understand links to moral functioning, future

studies could therefore benefit from shifting the focus to the characteristics unique to proactive aggression. For instance, proactive aggression has been associated with the presence of callous–emotional (CU) traits entailing low levels of empathy and guilt and a tendency to use others for personal gain (Frick & White, 2008), and children and adolescents exhibiting CU traits and proactive aggression show similar social-cognitive, affective, and behavioral profiles (see Frick & White, 2008). Future research could therefore examine whether deficits in moral judgments and reasoning are more strongly associated with aggression in children exhibiting high levels of CU traits (i.e., proactive aggressors) compared to children low on this dimension. This approach could avoid some of the limitations of existing questionnaires while shedding light on how different social-cognitive and affective mechanisms may contribute to the same observable behaviors.

Consistent with previous social domain research, we focused on hypothetical moral vignettes depicting straightforward instances of harm that closely resemble proactive aggression. As such, reactive aggression may have been linked with less advanced judgments if situations entailing reactive aggression (e.g., provocation or retaliation) had been included. Although these types of complex situations have been the subject of much previous research (Arsenio et al., 2009; Astor, 1994; Gasser et al., 2012), they do not constitute prototypical moral transgressions and were therefore not the focus of the present study. However, future research should examine whether proactive and reactive aggression are differentially associated with evaluations of hypothetical situations that are matched in kind to ensure that these findings are robust.

We also relied on a relatively small subset of moral and conventional vignettes and focused on a limited age range of participants. Although the stories and age range were appropriate for the current investigation, links between judgments and aggression would be expected to differ depending on the types of stories and developmental period examined. For example, judgments of more covert moral events (e.g., social exclusion) may be more strongly associated with proactive aggression that is relational or social in nature. Taking into account the different forms of aggression may be particularly important at later ages. Physical aggression, common in young children, normatively declines with age and is associated with a host of problematic outcomes later in life (Eisner & Malti, 2015). As such, proactive aggression entailing serious physical violence in older youth is likely indicative of broader psychosocial deficits rather than a lack of moral values. Understanding developmental shifts in the relation between moral judgments and different forms and functions of aggression is an important avenue for future study.

Finally, greater research is needed to understand the developmental antecedents and consequences of variations in children's understanding of right and wrong. To date, research on individual differences in moral development has primarily focused on affective (guilt, empathy) and behavioral constructs (prosociality, compliance; see Killen & Smetana, 2015). Yet intrapersonal (e.g., temperament), interpersonal (e.g., parenting practices), and contextual factors (e.g., neighborhood disadvantage) all influence how children make sense of the world. Furthermore, although deficits in prototypical moral understanding have been documented in adolescents and adults exhibiting extreme antisocial behavior (Blair, 1995; Blair et al., 2001; Nucci & Herman, 1982), we do not know whether impairments in early social and moral judgments have implications for later moral functioning in typically developing populations. Examining the developmental processes associated with individual differences in moral cognitions would provide a more integrated and nuanced understanding of children's moral and social development.

References

Arsenio, W., Adams, E., & Gold, J. (2009). Social information processing, moral reasoning, and emotion attributions: Relations with adolescents' reactive and proactive aggression. *Child Development*, *80*, 1739–1755. doi:10.1111/j.1467-8624.2009.01365.x

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

Astor, R. (1994). Children's moral reasoning about family and peer violence: The role of provocation and retribution. *Child Development*, *65*, 1054–1067. doi:10.2307/1131304

Blair, R. (1995). A cognitive developmental approach to morality: Investigating the psychopath. *Cognition*, *57*, 1–29. doi:10.1016/0010-0277(95)00676-p

Blair, R., Monson, J., & Frederickson, N. (2001). Moral reasoning and conduct problems in children with emotional and behavioral difficulties. *Personality and Individual Differences*, *31*, 799–811. doi:10.1016/s0191-8869(00)00181-1

Card, N., & Little, T. (2006). Proactive and reactive aggression in childhood and adolescence: A meta-analysis of differential relations with psychosocial adjustment. *International Journal of Behavioral Development*, *30*, 466–480. doi:10.1177/0165025406071904

Conry-Murray, C. (2013). Young children's understanding of beliefs about moral and conventional rule violations. *Merrill-Palmer Quarterly*, *59*, 489–510. doi:10.13110/merrpalmquar.59.4.0489

Crick, N., & Dodge, K. (1996). Social-information processing mechanisms in reactive and proactive aggression. *Child Development*, *67*, 993–1002. doi:10.1037/0033-2909.115.1.74

Cui, L., Colasante, T., Malti, T., Ribeaud, D., & Eisner, M. P. (2016). Dual trajectories of reactive and proactive aggression from mid-childhood to early adolescence: Relations to sensation seeking, risk taking, and moral reasoning. *Journal of Abnormal Child Psychology*, *44*, 663–675. doi:10.1007/s10802-015-0079-7

Dodge, K., Coie, J., & Lynam, D. (2006). Aggression and antisocial behavior in youth. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Social, emotional, and personality development* (Vol. 3, 6th ed., pp. 719–788). New York, NY: Wiley.

Dolan, M., & Fullam, R. (2010). Moral/conventional transgression distinction and psychopathy in conduct disordered adolescent offenders. *Personality and Individual Differences*, *49*, 995–1000. doi:10.1016/j.paid.2010.08.011

Dunn, J. (2014). Moral development in early childhood and social interaction in the family. In M. Killen, & J. Smetana (Eds.), *Handbook of moral development*, 2nd ed. (pp. 135–159). New York, NY: Psychology Press.

Eisenberg, N., Spinrad, T., & Morris, A. (2014). Empathy-related responding in children. In M. Killen, & J. Smetana (Eds.), *Handbook of moral development*, 2nd ed. (pp. 184–207). New York, NY: Psychology Press.

Eisner, M., & Malti, T. (2015). Aggressive and violent behavior. In R. Lerner (Editor-in-Chief) & M. Lamb (Vol. Ed.), *Handbook of child psychology and developmental science: Socioemotional processes* (Vol. 3, 7th ed., pp. 794–841). New York, NY: Wiley-Blackwell.

Frick, P., & White, S. (2008). Research review: The importance of callous-unemotional traits for developmental models of aggressive and antisocial behavior. *Journal of Child Psychology and Psychiatry*, *49*, 359–375. doi:10.1111/j.1469-7610.2007.01862.x

Gasser, L., & Keller, M. (2009). Are the competent morally good? Perspective taking and moral motivation of children involved in bullying. *Social Development*, *18*, 798–816. doi:10.1111/j.1467-9507.2008.00516.x

Gasser, L., Malti, T., & Gutzwiller-Helfenfinger, E. (2012). Aggressive and non-aggressive children's moral judgments and moral emotion attributions in situations involving retaliation and unprovoked aggression. *Journal of Genetic Psychology*, *173*, 417–439. doi:10.1080/00221325.2011.614650

Hawley, P. (2003). Strategies of control, aggression, and morality in preschoolers: An evolutionary perspective. *Journal of Experimental Child Psychology*, *85*, 213–235. doi:10.1016/s0022-0965(03)00073-0

Hawley, P. (2014). The duality of human nature: Coercion and prosociality in youths' hierarchy ascension and social success. *Current Directions in Psychological Science*, *23*, 433–438. doi:10.1177/0963721414548417

Hawley, P., & Geldhof, J. (2012). Preschoolers' social dominance, moral cognition, and moral behavior: An evolutionary perspective. *Journal of Experimental Child Psychology*, *112*, 18–35. doi:10.1016/j.jecp.2011.10.004

Hubbard, J., McAuliffe, M., Morrow, M., & Romano, L. (2010). Reactive and proactive aggression in childhood and adolescence: Precursors, outcomes, processes, experiences, and measurement. *Journal of Personality*, *78*, 95–118. doi:10.1111/j.1467-6494.2009.00610.x

Huesmann, L., & Guerra, N. (1997). Children's normative beliefs about aggression and aggressive behavior. *Journal of Personality and Social Psychology*, *72*, 408–419. doi:10.1037//0022-3514.72.2.408

Jambon, M., & Smetana, J. (2014). Moral complexity in middle childhood: Children's evaluations of necessary harm. *Developmental Psychology*, *50*, 22–33. doi:10.1037/a0032992

Killen, M., & Smetana, J. (2015). Origins and development of morality. In R. Lerner (Editor-in-Chief) & M. Lamb (Vol. Ed.), *Handbook of child psychology and developmental science: Socioemotional processes* (Vol. 3, 7th ed., pp. 701–749). New York, NY: Wiley-Blackwell.

Little, T., Henrich, C., Jones, S., & Hawley, P. (2003). Disentangling the "whys" from the "whats" of aggressive behaviour. *International Journal of Behavioral Development*, *27*, 122–133. doi:10.1080/01650250244000128

Little, T., Rhemtulla, M., Gibson, K., & Shoemann, A. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods*, *18*, 285–300. doi:10.1037/a0033266

Malti, T., Gasser, L., & Buchmann, M. (2009). Aggressive and prosocial children's emotion attributions and moral reasoning. *Aggressive Behavior*, *35*, 90–102. doi:10.1002/ab.20289

Malti, T., Gasser, L., & Gutzwiller-Helfenfinger, E. (2010). Children's interpretive understanding, moral judgments, and emotion attributions: Relations to social behavior. *The British Journal of Developmental Psychology*, *28*, 275–292. doi:10.1348/026151009x403838

McArdle, J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, *60*, 577–605. doi:10.1146/annurev.psych.60.110707.163612

Murray-Close, D., & Ostrov, J. (2009). A longitudinal study of forms and functions of aggressive behavior in early childhood. *Child Development*, *80*, 828–842. doi:10.1111/j.1467-8624.2009.01300.x

Muthén, L., & Muthén, B. (1998-2012). *Mplus user's guide (version 7)*. Los Angeles, CA: Author.

Nucci, L. (2009). *Nice is not enough: Facilitating moral development*. Upper Sadle River, NJ: Pearson.

Nucci, L., & Herman, S. (1982). Behavioral disordered children's conceptions of moral, conventional, and

personal issues. *Journal of Abnormal Child Psychology*, *10*, 411–426. doi:10.1007/bf00912330

Nunner-Winkler, G. (2013). Moral motivation and the happy victimizer phenomenon. In K. Heinrichs, F. Oser, & T. Lovat (Eds.), *Handbook of moral motivation: Theories, models, applications* (pp. 267–288). Rotterdam, the Netherlands: Sense.

Orobio de Castro, B., Verhulp, E., & Runions, K. (2012). Rage and revenge: Highly aggressive boys' explanations for their responses to ambiguous provocation. *European Journal of Developmental Psychology*, *9*, 331–350. doi:10.1080/17405629.2012.680304

Ostrov, J., & Crick, N. (2007). Forms and functions of aggression during early childhood: A short-term longitudinal study. *School Psychology Review*, *36*, 22–43. Retrieved from http://http://naspjournals.org/loi/spsr

Piaget, J. (1965). *The moral judgment of the child*. New York, NY: Free Press. (Original work published 1932)

Poland, S., Monks, C., & Tsermentseli, S. (2016). Cool and hot executive function as predictors of aggression in early childhood: Differentiating between the function and form of aggression. *The British Journal of Developmental Psychology*, *34*, 181–197. doi:10.1111/bjdp.12122

Posada, R., & Wainryb, C. (2008). Moral development in a violent society: Colombian children's judgments in the context of survival and revenge. *Child Development*, *79*, 882–898. doi:10.1111/j.1467-8624.2008.01165.x

Poulin, F., & Bouivin, M. (2000). Reactive and instrumental aggression: Evidence of a two-factor model. *Psychological Assessment*, *12*, 115–122. doi:10.1037/1040-3590.12.2.115

Rutland, A., Cameron, L., Milne, A., & McGeorge, P. (2005). Social norms and self-presentation: Children's implicit and explicit intergroup attitudes. *Child Development*, *76*, 451–466. doi:10.111/j.1467-8624.2005.00856.x

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

Smetana, J., Toth, S., Cicchetti, D., Bruce, J., Kane, P., & Daddis, C. (1999). Maltreated and nonmaltreated preschoolers' conceptions of hypothetical and actual moral transgressions. *Developmental Psychology*, *35*, 269–281. doi:10.1037/0012-1649.35.1.269

Stoltz, S., Cillessen, A., van den Berg, Y., & Gommans, R. (2016). Popularity differentially predicts reactive and proactive aggression in early adolescence. *Aggressive Behavior*, *42*, 29–40. doi:10.1002/ab.21603

Sutton, J., Smith, P., & Swettenham, J. (1999). Bullying and "theory of mind": A critique of the "social skills deficit" view of antisocial behavior. *Social Development*, *8*, 117–127. doi:10.1111/1467-9507.00083

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

Vitaro, F., Brendgen, M., & Barker, E. (2006). Subtypes of aggressive behaviors: A developmental perspective. *International Journal of Behavioral Development*, *30*, 12–19. doi:10.1177/0165025406059968

Wainryb, C., & Recchia, H. (Eds.) (2014). *Talking about right and wrong: Parent–child conversations as contexts for moral development*. New York, NY: Cambridge University Press.

Weston, D., & Turiel, E. (1980). Act-rule relations: Children's concepts of social rules. *Developmental Psychology*, *16*, 417–424. doi:10.1037/0012-1649.16.5.417

Supporting Information

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

Appendix S1. Social Rules Interview Stimuli and Script

Appendix S2. Estimating a Latent Domain Distinction Factor

Appendix S3. Full Descriptive Statistics for Moral and Conventional Judgments, by Story and Criteria

Appendix S4. Parceling of Proactive and Reactive Aggression Items