



ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Cognitive Development

journal homepage: www.elsevier.com/locate/cogdev

When it's okay to hit: How Turkish and U.S. preschoolers and adults make judgments about permissible and impermissible acts of force

Charles P. Baxley Jr ^{a,*}, Muhammet Ali Karaduman ^b, Isabelle Gross ^a, Emma Ford ^a, Audun Dahl ^a

^a University of California, Santa Cruz, USA

^b Artvin Coruh University, Turkey

ARTICLE INFO

Keywords:

Moral judgments
Aggression
Emotion attribution
Preschool children

ABSTRACT

To navigate their complex social worlds, children need to make judgments about when, if ever, it is okay to use force against others. By adulthood, most come to condemn violence in most situations yet accept or even encourage force in other situations, such as self-defense or contact sports. This research examined key factors expected to guide how preschoolers and adults make judgments about permissible and impermissible force. These factors included the antecedent events—preceding the force act—and the emotion attributed to the victim. In four studies with Turkish and U.S. preschoolers and adults, an interviewer presented participants with vignettes in which one child used force against another. The antecedent event significantly affected children's and, to an even greater extent, adults' judgments about force. Participants were also more likely to judge force as permissible when they attributed positive or neutral emotions to the victim. Some cultural differences also emerged. The findings shed light on how children begin to draw moral distinctions between permissible and impermissible force.

The moral complexities of violence are easy to overlook and difficult to learn. Whereas many violent acts are universally condemned, other acts of interpersonal force are widely accepted (Baxley & Dahl, *in press*; Fiske & Rai, 2014). By adulthood, people commonly approve of hitting in self-defense and even celebrate forceful tackles in contact sports. Preschool-age children face their own conundrums regarding acts of force. Children may elicit anger from peers or adults when they harm an innocent victim, yet elicit laughter when they push a peer during rough-and-tumble play (Dahl, 2016; Dahl, Sherlock, Campos & Theunissen, 2014; Smith & Boulton, 1990; Tannock, 2008). Distinctions between right and wrong acts of force help children navigate their social worlds. Children who draw these distinctions adaptively can skillfully condemn, use, and applaud force in everyday interactions (Baker & Liu, 2021; Hawley, 2003; Jambon & Smetana, 2018). In contrast, children who struggle with these distinctions can hit too hard in the wrong situation—alienating peers—or fail to defend themselves against attacks from others.

The preschool years are a crucial time in the development of children's ability to judge some acts of force acts as okay, or *permissible*, and other force acts as wrong, or *impermissible*. From three to five years of age, children gradually use fewer acts of force and form more nuanced moral judgments (Dahl & Freda, 2017; Hay, 2005). Simultaneously, preschoolers encounter new social events and relationships, prodding children to learn new social rules and expectations (Rubin et al., 2006; Siegal & Storey, 1985). To make

* Correspondence to: Department of Psychology, University of California, 1156 High Street, Santa Cruz, CA 95064, United States.
E-mail address: cbaxley@ucsc.edu (C.P. Baxley).

<https://doi.org/10.1016/j.cogdev.2021.101152>

Received 12 August 2021; Received in revised form 21 December 2021; Accepted 28 December 2021

Available online 8 January 2022

0885-2014/© 2022 Elsevier Inc. All rights reserved.

judgments about permissible and impermissible acts of force, children eventually need to consider factors such as the antecedents of the act (e.g., a provocation) and the effects of the act on the victim.

The present research examined factors that lead preschoolers to judge whether everyday acts of force are permissible or impermissible. Preschoolers from Turkey and the United States evaluated hypothetical vignettes involving acts of force among children. For comparison, we also sampled adults from both countries. Our main focus was on the role of *antecedent events* and *emotion attributions* in preschoolers' and adults' judgments about acts of force.

1. The development of moral orientations toward force

During the preschool years, children encounter and evaluate everyday acts of force (Dahl & Freda, 2017; Dahl & Turiel, 2019; Hay, 2005; Rubin et al., 2006; Smetana et al., 2018). Peaking around the third year of life, rates of physical aggression slowly decrease in the subsequent years (Hay, 2005; Tremblay et al., 2018). Still, most children will encounter and use more interpersonal force between three to five years of age than in any later period in their life. These acts of force in early childhood come in many forms, for instance unprovoked acts of hitting, forceful retrieval of toys, pushing in self-defense, or wrestling during rough-and-tumble play (Dahl, 2016; Hay, 2005; Tannock, 2008; Vitaro & Brendgen, 2005). As preschoolers navigate an increasingly complex social environment, children develop nuanced judgments about permissible and impermissible acts of force.¹

By three years of age, most children judge that hitting or pushing others is generally wrong (Dahl & Turiel, 2019; Nucci & Weber, 1995). In most studies of preschoolers' views about interpersonal force, interviewers have presented children with unprovoked acts of force: events in which one child hit or pushed another without apparent cause. Most preschoolers and older children judge that these unprovoked acts of force are wrong. They tend to judge acts of hitting as more serious than other violations, such as mess-making or violations of conventions for how to dress (Ball et al., 2017; Dahl & Kim, 2014; Smetana et al., 1984; Smetana, Rote, Jambon, Tasopoulos-Chan, Villalobos & Comer, 2012). Preschoolers make these judgments based on concerns with victims' rights and welfare and, accordingly, typically deem that hitting or kicking others is wrong even if teachers gave permission (Smetana et al., 2012; for a review, see Smetana, 2013).

At some point in childhood, children begin to judge some acts of force as permissible. By middle childhood, many children think it is okay to hit another child in self-defense, in response to a provocation, or to prevent greater harm (e.g., Dahl, Gingo, Uttich & Turiel, 2018; Jambon & Smetana, 2018; Noh, Jambon, Smetana, Lee & Killen, 2020; Nucci et al., 2017; Pnevmatikos, 2018; Wainryb et al., 2005). School-age children with more exposure to aggression tend to view physical force as more permissible, especially in response to provocations (Ardila-Rey et al., 2009; Arsenio & Lemerise, 2004; Astor, 1994; Gasser et al., 2012). By adolescence, many have come to judge it permissible to take a life in order to save others (Dahl et al., 2018; Fiske & Rai, 2014; Pnevmatikos, 2018).

For younger children, however, distinguishing permissible from impermissible force is likely challenging. Preschoolers often have difficulties balancing competing moral and other evaluative considerations (Dahl et al., 2020; Dahl & Killen, 2018; Killen et al., 2018). For instance, past research has presented children with vignettes in which a protagonist had the option of helping a child steal from another child (Dahl et al., 2020; J. G. Miller et al., 1990). The vignette thereby forced children to balance their concerns with helping against their concerns with not stealing. Whereas virtually all adults said it was wrong to help another child steal, less than half of preschoolers said it was wrong. Similarly, preschoolers struggle to incorporate multiple emotional states when they attribute emotions to others. They are, for instance, more likely than older children to attribute positive emotional states to a child who has victimized another child ("happy victimizers," e.g., Arsenio & Kramer, 1992). In one longitudinal study that followed children from around four to six years of age, children became less likely with age to attribute positive emotions to a victimizer who, for instance, shoved another child to get a cherished snack (Jambon, Colasante, & Malti, 2021). These and other findings point to the difficulties preschoolers have in balancing multiple considerations when they form judgments about permissible and impermissible actions (see also Killen et al., 2018; Nucci et al., 2017).

The present research focused on two factors that likely shape children's judgments about permissible and impermissible force: the *antecedent events* of the force (e.g., was the force unprovoked or was it a response to a prior act of force) and *emotion attributed to the victim* (e.g., positive versus negative emotion).

1.1. Antecedent events

By antecedents of force, we mean the circumstances immediately preceding and surrounding the forceful act (Baxley & Dahl, in press; Dahl et al., 2018; Nucci et al., 2017; Smetana & Ball, 2018). As we discussed above, adults often deem force permissible in many contexts, for instance when force is used to save another life, in self-defense, or as part of contact sports (Dahl et al., 2018; Fiske & Rai, 2014; Kimble, Russo, Bergman & Galindo, 2010; Pnevmatikos, 2018). While these contexts are familiar to adults, however, they are far removed from the everyday acts of force about which children form judgments (Dahl & Turiel, 2019; Davidson et al., 1983).

Here, we will focus on four antecedent contexts of force: *Unprovoked*, *property conflict*, *self-defense*, and *physical play*. These contexts are common in the everyday lives of preschoolers, yet they have—with the exception of unprovoked contexts—received limited attention in prior research on early moral development.

¹ Our research examines situational and psychological factors that lead individuals to judge acts of force as permissible. It does *not* examine which acts truly are permissible, in a normative or evaluative sense, nor do we assume that everyone in a community agrees on which acts are permissible. (For discussion, see e.g., Dahl & Waltzer, 2018).

1.1.1. Unprovoked force

Perhaps the simplest form of force is *unprovoked* force, wherein one person uses abrupt force against another person's body for no apparent reason (Dahl, 2016). Unprovoked acts, as noted, have been the focus of most prior research on young children's judgments about acts of force (e.g., Dahl & Kim, 2014; Nucci & Weber, 1995; Smetana et al., 2012). This work has shown that nearly all preschoolers—like older children and adults—judge it generally wrong to hit, push, or otherwise use excessive force against others. Judgments about unprovoked force, evident by three years of age, emerge after years of everyday experiences with force. By one estimate, about half of one-year-olds' daily acts of force are unprovoked (Dahl, 2016). Although judgments about unprovoked force have been studied extensively, they serve as a useful comparison for other contexts in which the general prohibition against force is pitted against other considerations.

1.1.2. Property conflict

In early childhood, many acts of force happen during conflicts about the control of objects (Brownlee & Bakeman, 1981; Caplan, Vespo, Pedersen & Hay, 1991; Hay, 2006; Hay, Hurst, Waters & Chadwick, 2011). When a two-year-old takes a toy from another child, the latter child will often strike in return, be it to defend property or to take revenge (Baxley & Dahl, in press). Property conflicts constitute a key context in which preschoolers must decide whether force is permissible. Many societies grant individuals the right to use violence to defend, or even acquire, property (Fiske & Rai, 2014). In some U.S. states, people have a right to use violence against someone who tries to rob them, even if they do not perceive themselves to be in physical danger (Ward, 2014). The present research examined whether children and adults deemed it permissible for a child to use force against another child when the latter sought to take a toy from the former.

1.1.3. Self-defense

Self-defense is viewed as a valid reason to use force in most communities (Fiske & Rai, 2014; Pnevmatikos, 2018; Ward, 2014). Still, prior research has yielded mixed findings on whether young children judge it permissible to use force in self-defense. Martin and Ross (1996) asked 3- to 4-year-olds about how much punishment a child should receive for using force against another child. They found no significant difference in how much punishment participants allocated between a vignette involving unprovoked force and a vignette in which a protagonist responded forcefully to a prior act of force. In other research with preschoolers and older children, however, participants have generally been more accepting, or at least less negative, in their evaluation of force used in self-defense than of unprovoked force (Ardila-Rey et al., 2009; Nucci et al., 2017; Smetana, Daddis, Toth, Cicchetti, Bruce & Kane, 1999, 2003). One goal of the present research was to examine preschoolers' and adults' judgments about events in which a child used force to defend themselves against a physical attack.

1.1.4. Physical play

Rough-and-tumble play, or play that involves physical contact, is common in interactions among young children (Martin & Ross, 1996; Pellegrini & Smith, 1998; Smith & Boulton, 1990; Tannock, 2008). As they grow older, many also partake in contact sports that involve intense, and sometimes painful, physical contact (Fiske & Rai, 2014; Mintah et al., 1999). To participate successfully in forceful play, children need to draw subtle distinctions between permissible and impermissible acts of force. Acts that are permissible inside a game—for instance forcefully bumping your shoulder into another child—may be severely prohibited or even dangerous outside of that context. Another goal of the present research was to examine whether preschoolers and adults were more accepting of force during physical play than of unprovoked force.

1.2. Emotion attributions

Insofar as children and adults condemn acts of force, they do so largely based on concerns with the victims' welfare and rights (Dahl et al., 2018; Dahl & Turiel, 2019; Smetana, 2013; Turiel, 2015; Wainryb et al., 2005). That is, they judge that hitting others is wrong because it has negative consequences for the victim. Hence, children and adults tend to condemn even unfamiliar acts if they believe that a victim reacted negatively to the act (Helwig et al., 2001; Smetana, 1985). Indeed, Smetana and colleagues (2003) found significant relations between evaluations of transgressions and emotion attributions. In the present research, we expected that participants would be more likely to judge force acts as permissible when they attributed neutral or positive reactions to the victim.

1.3. The present research

In the present research, we presented preschoolers and adults with hypothetical vignettes in which one child used force against another. We experimentally varied the antecedent events of force act, and we hypothesized that participants would judge force as most permissible in *forceful play* and *self-defense* contexts followed by *property conflict* contexts. We expected very few participants to judge force as permissible in *unprovoked* contexts. We also hypothesized that participants would deem force more permissible when they attributed neutral or positive emotions to the victim.

We included adult participants both to validate our stimuli and to identify a state toward which children were developing. Given young children's difficulties in incorporating multiple considerations into their judgments, discussed above, we expected that adults would distinguish more sharply than preschoolers among permissible and impermissible acts of force. Specifically, we expected that adults and children would be similarly unlikely to judge force as acceptable in the unprovoked context but that adults would be even more likely than children to judge force as acceptable in the play and self-defense contexts.

Before presenting the studies, we will make three comments about our sampling of events and participants, to which we will return in the General Discussion. First, the present study focused on acts of force of intermediate intensity that could occur in any of the four contexts. Extreme acts of force, such as lethal force, would presumably be judged impermissible in most contexts, unless one's own or another life was threatened (Dahl et al., 2018; Pnevmatikos, 2018; Ward, 2014). However, such extreme acts are rare in the lives of young children.

Second, we recruited participants from communities in two countries: Turkey and the United States. By recruiting from these two locations, we aimed to assess the generalizability of our findings rather than to examine cultural differences in children's and adults' judgments about force. Although cross-cultural comparisons were not the main goal of this research, we briefly analyze and discuss such comparisons in the last part of the paper. Unlike the United States, Turkey has not been the site of much psychological research on moral development. The work that has been conducted in Turkey nonetheless suggests overarching similarities between Turkish children and children elsewhere, for instance in the sense that children across the world come to think that unprovoked harm against others is wrong (Kuyel & Glover, 2010; Nisan & Kohlberg, 1982; Turiel, 2015; Turiel et al., 1978). Thus, we expected that both Turkish and U.S. samples would draw similar distinctions between permissible force (e.g., in play contexts) and impermissible force (e.g., in unprovoked contexts). Developmentally, we expected that adults would differentiate more sharply among the contexts of force than would preschoolers in both countries.

Third, all studies aimed for a sample size of at least 40. Pilot data led us to expect a 20% difference between the antecedent events with the highest and lowest acceptance. To estimate power, we simulated data and fitted Generalized Linear Models for each simulated dataset (see details on analyses below). These simulations indicated that a sample size of 40 yielded estimated power above .90 for detecting an effect of antecedent event on judgments, which was the principal goal of this research. As we note below, however, data collection with Turkish preschoolers (Study 1A) was interrupted due to the COVID-19 pandemic. In contrast, sample sizes for Studies 1B, 2A, and 2B all exceeded 40 and thus yielded adequate estimated power for addressing our main research questions.

2. Study 1A: Turkish preschoolers

2.1. Method

2.1.1. Participants

Although the target sample size for Study 1A was 40, data collection was cut short due to the COVID-19 pandemic. Participants were recruited through a preschool in Northeastern Turkey. The final sample consisted of 19 children (10 female, 9 male, mean age = 5.1 years). As power analyses indicated that this yielded an estimated power of about .70, we decided that the findings were nevertheless worth reporting, especially since we had the opportunity to compare findings across the four studies (see below). Children were native Turkish speakers. Seven children were excluded from the sample because of parental interference or because the child was unable or unwilling to complete the procedures. Parents provided written permission prior to data collection.

2.1.2. Materials

The original materials were developed in English. The Turkish adaptation was created by using translation, back-translation, and expert panel discussion sessions to provide cross-cultural and conceptual equivalence. Three bilingual researchers and two translators took part in the adaptation process. To examine the effects of antecedent events on judgments about force, we created vignettes that represented each of the four antecedent conditions (*unprovoked*, *property conflict*, *self-defense*, and *physical play*) and four different act types (*bumping* another child, *pulling* another child to the ground, *hitting* another child with a stuffed animal, and *throwing* a rubber ball at another child). The contexts and act types were combined to create a set of eight vignettes for Studies 1A and 1B such that each antecedent event was presented with two act types. Half of the vignettes had a female protagonist, and the other half had a male protagonist. To facilitate children's comprehension, the interviewer presented two pictures with each vignette. The first picture showed the scene right before the protagonist's act of force and the second picture showed the scene in which the protagonist used force.

For instance, for the vignette for the *physical play* condition and throwing action with a female protagonist, the interviewer would say the following: "Eren is at her preschool. Eren and Ates are playing with toys at the playground. Eren and Ates are playing a game together, where they try to throw a ball and hit each other with the ball. During the game, Eren throws a rubber ball Ates' chest" (see SOM for complete texts). The victim's appearance and name left their gender ambiguous. In the second picture, the victim's face was turned away so that the participant could not see their facial expression so as not to influence participants' emotion attribution. In addition, we used a five-point visual rating scale for assessing emotion attributions. The scale showed five faces ranging from very negative to very positive.

2.2. Procedure

Interviews were conducted in the family home in Turkish by the second author, who is a native speaker. The interviews lasted around 25 min. Each participant saw eight vignettes in random order. After each vignette, the interviewer asked the following questions to all participants: Was it okay for [protagonist] to [use force in context] (*permissibility*)? Was it not bad, a little bad, or really, really bad that. (*severity rating*)? Can you point to the face that [victim] was feeling right after. (*victim emotion attribution*)? Can you point to the face that [protagonist] was feeling right after. (*protagonist emotion attribution*)? The interview was audio recorded for coding purposes.

2.3. Data analysis

Non-dichotomous variables were dichotomized before analysis in the same way for all four studies. Since the distribution of the severity ratings was severely non-normal, we analyzed whether participants gave a *severity* rating of “not bad” (vs. “a little bad” or “really really bad”).² For *emotion attributions*, we analyzed whether participants attributed negative (versus neutral or positive) emotions to the protagonist and the victim. Data were analyzed using Generalized Linear Mixed Models with logistic link function and binomial error distribution (Hox, 2010). Models included fixed effects for antecedent event, act type, participant gender, and participant age (in years), as well as random intercepts for participants. As preliminary analyses revealed no effects of transgressor gender for any of the studies (¹A–²B), this predictor was not included in the final models. Hypotheses were tested using likelihood ratio tests (D) and Wald tests for regression coefficients (W). For all four studies, the alpha level was .05.

2.4. Results

2.4.1. Permissibility judgments

Children's judgments about force varied significantly by antecedent event, $D(3) = 14.09, p = .003$ (Table 1). Participants were more accepting of force in the *self-defense* context (24%) than in the *play* (8%), *property* (3%), and *unprovoked* (3%) conditions, $Ws(1) > 4.08, ps < 0.044$. No other cells differed significantly, $Ws(1) < 1.15, ps > 0.28$. Permissibility judgments did not vary significantly by type of force, $D(3) = 5.88, p = .12$, gender, $D(1) = 3.00, p = .083$, or age, $D(1) = 1.54, p = .21$.

2.4.2. Severity rating

No Turkish preschoolers responded that the force action was “not bad.”

2.4.3. Victim emotion attribution

In 93% of cases, children attributed negative emotions to the victim. Participants' attributions of negative emotions to victims varied significantly by antecedent event, $D(3) = 17.23, p < .001$. Although numerically different, however, none of the contexts differed significantly in pairwise comparisons (*Unprovoked*: 100%, *Property*: 97%, *Self-defense*: 87%, *Play*: 87%), $Ws(1) < 2.88, ps > 0.08$. There were no significant effects of force type, $D(1) = 5.57, p = .13$, age, $D(1) = 2.40, p = .12$, or gender, $D(1) = 0.83, p = .36$.

Next, we fitted a separate model predicting permissibility judgments from negative emotion attributions to victims, controlling for antecedent event, action type, gender, and age. As predicted, children were likely to say the action was okay when they attributed neutral or positive emotions to the victim (45% okay, 5/11 trials) than when they attributed negative emotions to the victim (6%, 9/141 trials), $D(1) = 6.44, p = .011$.

2.4.4. Protagonist emotion attribution

Participants attributed negative emotions to the protagonist in 68% of trials overall. Protagonist emotion attributions varied by antecedent event, $D(3) = 10.65, p = .014$. Participants were less likely to attribute negative emotions to the protagonist in the *play* context (53%) than in the *unprovoked* (74%) and *property* (79%) contexts, $Ws(1) > 5.74, ps < 0.017$. No other pairwise comparisons were significant, $Ws(1) < 3.37, ps > 0.06$. (For completeness, we note that participants attributed negative emotions to the protagonist 68% of the time in the *self-defense* context.) There were no significant effects of force type, $D(3) = 2.44, p = .49$, age, $D(1) = 2.24, p = .13$, or gender, $D(1) = 1.19, p = .28$.

Attributions of negative emotions to the protagonist did not significantly predict permissibility judgments, $D(1) = 0.37, p = .54$, when controlling for antecedent event, force type, gender, and age.

3. Study 1B: Turkish adults

3.1. Method

3.1.1. Participants

We recruited 53 adults (35 female, 18 male, mean age = 21.6 years) from a university in Northeastern Turkey. Students were recruited via fliers on campus, and they received bonus course credit for their participation. All participants were native Turkish speakers. Written consent was obtained before data collection started.

² For all four studies, the distribution of the ordinal severity ratings was significantly non-normal, as assessed by the Shapiro-Wilk test, $ps < 0.001$, and had either severe skew (< -1 , Studies 1A, 2A, and 2B) or kurtosis (< -1 , Study 1B). The dichotomization of quantitative variables can be warranted in the analysis of dependent variables whose distribution is severely non-normal (MacCallum, Zhang, Preacher & Rucker, 2002). This practice can particularly be appropriate here, since Generalized Linear Mixed Models using Maximum Likelihood estimation are less robust to violations of normality than Ordinary Least Squares Regression and Analysis of Variance (Hox, 2010). We also note that, since the original variable—severity rating—only had three levels, the conversion to a two-level variable amounted to little loss of information. We thank an anonymous reviewer for prompting us to explain our analytical strategy.

Table 1
Proportion of okay judgments by antecedent event and sample population.

Antecedent event	Study 1: Turkey		Study 2: United States	
	Preschoolers (1A)	Adults (1B)	Preschoolers (2A)	Adults (2B)
Unprovoked	0.03 _a	0.04 _a	0.05 _a	0.04 _a
Property conflict	0.03 _a	0.09 _a	0.16 _b	0.12 _b
Self-defense	0.24 _b	0.35 _b	0.08 _a	0.51 _c
Physical play	0.08 _a	0.66 _c	0.16 _b	0.60 _c

Note. Cells show the proportion of participants who said the act was okay. For each column, cells with different subscripts differ significantly, $p < 0.05$.

3.1.2. Materials and procedures

The materials, procedures, and data analysis for Study 1B were as for Study 1A, with the following exceptions. First, the interviews took place in a room on the university campus rather than in participant homes. Second, the pictures illustrating the vignettes and the visual rating scale were deemed unnecessary for adult interviews and were therefore not used. Third, the analyses of adult data did not include participant age as a predictor.

3.2. Results

3.2.1. Permissibility judgments

Turkish adults indicated whether the acts of force were okay or not okay. There was a significant effect of antecedent events on permissibility judgments, $D(3) = 150.83$, $p < .001$. Participants were more likely to say that force was okay in the *play* events (69%) than the *self-defense* (34%), *property* (7%), and *unprovoked* (5%) events, $Ws(1) > 27.74$, $ps < 0.001$. In turn, the force act was more often deemed okay in the *self-defense* context than in the *property* and *unprovoked* events, $Ws(1) > 25.68$, $ps < 0.001$. The *property* and *unprovoked* events did not differ significantly, $W(1) = 0.39$, $p = .53$. There was no significant effect of force type, $D(3) = 4.77$, $p = .19$. Male participants were more likely to say the force was okay (40%) than were female participants (23%), $D(1) = 11.84$, $p < .001$.

3.2.2. Severity ratings

Participants' tendency to say that the protagonist's action was "not bad" varied, $D(3) = 117.37$, $p < .001$. The *play* events (55%) were more likely to be rated as "not bad" than the *self-defense* (9%), *unprovoked* (7%), and *property* (0%) events, $Ws(1) > 48.19$, $ps < 0.001$. The *self-defense* and *unprovoked* events also differed from the *property* events, $Ws(1) > 4.39$, $ps < 0.036$, but did not differ from each other, $W(1) = 0.61$, $p = .44$. "Not bad" ratings also differed by force type, $D(3) = 30.43$, $p < .001$. "Not bad" ratings were more common for the *pulling* (36%) and *throwing* (25%) events than for *hitting* (5%) and *bumping* (5%) events, $Ws(1) > 15.11$, $ps < 0.001$. No other pairwise comparisons were significant, $Ws(1) < 2.90$, $ps > 0.088$. Males were more likely to rate acts of force as "not bad" (26%) than were females (14%), $D(1) = 9.07$, $p = .003$.

3.2.3. Victim emotion attribution

Adults attributed negative emotions to victims in 83% of trials. Participants' attribution of negative emotion varied significantly by antecedent event, $D(3) = 39.70$, $p < .001$. Adults were less likely to attribute negative emotions to the victim in *play* events (63%) than in the *self-defense* (87%), *unprovoked* (91%), or *property* (92%) contexts, $Ws(1) > 15.22$, $ps < 0.001$. No other comparisons were significant, $Ws(1) < 1.86$, $ps > 0.17$. Negative emotion attributions also varied by force type, $D(3) = 11.85$, $p = .008$. Negative emotion attributions were less common in the *throwing* (75%), *pulling* (78%), and *hitting* (84%) contexts than in the *bumping* (95%) contexts, $Ws(1) > 6.67$, $ps < 0.010$. No other comparisons were significant, $Ws(1) < 2.44$, $ps > 0.11$. Lastly, women were more likely to attribute negative emotions to the victim (87%) than were men (76%), $D(1) = 6.33$, $p = .012$.

Separate models revealed a significant relation between negative emotion attributions to victims and permissibility judgments: Participants were more likely to judge the act as okay when they attributed neutral or positive emotions to the victim (69%, 49/71 trials) than when they attributed negative emotions to the victim (20%, 72/353 trials), $D(1) = 25.63$, $p < .001$, controlling for antecedent event, action type, gender, and age.

3.2.4. Protagonist emotion attribution

On average, adults attributed negative emotions to protagonists in 56% of cases. Emotion attributions varied significantly by antecedent event, $D(3) = 62.04$, $p < .001$, as participants were far less likely to attribute negative emotions to the protagonist in the *play* context (25%) than in the *unprovoked* (58%), *property* (69%), and *self-defense* (70%) contexts, $Ws(1) > 28.14$, $ps < 0.001$. The *unprovoked* event also differed significantly from the *property* and *self-defense* events, $Ws(1) > 3.92$, $ps < 0.048$, whereas the later two did not differ significantly, $W(1) = 0.27$, $p = 0.05$. Negative emotion attributions to protagonist's also varied significantly by force type, $D(3) = 17.66$, $p < 0.001$. Participants were the least likely to attribute negative emotions in the *pulling* (40%) and *throwing* (48%), both of which differed significantly from the *hitting* (68%) and *bumping* (69%) vignettes, $Ws(1) > 13.57$, $ps < 0.001$. No other pairwise comparisons were significant, $Ws(1) < 1.92$, $ps > 0.16$. There was no significant effect of participant gender, $D(1) = 0.422$, $p = 0.52$.

Separate analyses revealed that participants were far more likely to judge the protagonist's action as permissible when they attributed neutral or positive emotions to the protagonist (49% okay, 93/188 trials) than when they did attribute negative emotions to

the protagonist (12%, 28/236 trials), $D(1) = 38.17, p < .001$, again controlling for antecedent event, action type, gender, and age.

3.3. Study 1 discussion

The findings from these first studies with Turkish samples of preschoolers and adults provided initial support for two of our hypotheses. First, participants' judgments about force varied by the antecedent events of the force act. Acts of force in the *self-defense* events and, for adults, the *play* events were judged more acceptable than acts of force in *property* and *unprovoked* events. Somewhat surprisingly, children rarely accepted force in the *play* events. Second, children and adults were more likely to judge force acts as okay when they attributed neutral or positive emotions to the victim. This aligns with prior work, which has shown that varying assumptions about how a force act affects a victim's welfare can lead to contextual variation in judgments about force (Wainryb, 1991).

Before discussing these findings in further detail, we will consider whether the findings were replicated in samples of U.S. preschoolers and adults in Studies 2A and 2B. Study 2A contained a full sample of preschoolers, as opposed to the reduced sample of preschoolers in Study 1A that resulted from COVID-19 restrictions on data collection. In addition, for the U.S. samples in Studies 2A and 2B, we analyzed participants' explanations for their reasoning about why force was or was not okay (Dahl et al., 2018; Dahl & Turiel, 2019; Jambon & Smetana, 2014). The justification data would offer further evidence on whether judgments about permissible and impermissible force were rooted in concerns with others' welfare, authority commands, or other considerations.

3.4. Authority permission and alternative actions

Studies 2A and 2B also assessed two further factors that may shape evaluations of force acts in the everyday lives of preschoolers: authority permission and possibilities for alternative action. Although judgments about force are typically less affected by authorities than are judgments about violations of social conventions or religious rules, authorities can still affect judgments about force (Davidson et al., 1983; Rhodes & Chalik, 2013; Srinivasan et al., 2019). Authorities may be particularly influential for young children's views about potential moral violations in multifaceted contexts (Braine, Pomerantz, Lorber & Krantz, 1991; Smetana, Yoo, Nguyen & Ball, 2021; Tisak, 1986). For instance, Smetana and colleagues (2021) found that 4- to 10-year-olds were more accepting of parental commands to retaliate against harm than of parental commands to steal from others. Thus, we expected that children, more than adults, would be more accepting of forceful actions permitted by teachers than forceful actions without such permission.

The final addition in Studies 2A and 2B was the examination of participants' views about alternative actions. At least under some circumstances, judgments that a protagonist ought to have acted differently—and not used force—implies that an alternative course of action was available to the protagonist (see Buckwalter & Turri, 2015; Kurthy et al., 2017). Thus, we expected that participants would be more likely to deem the act of force as permissible when they could not provide any alternative courses of action, such as seeking help from teachers or responding verbally rather than physically to a provocation (Arsenio & Lemerise, 2004).

4. Study 2A: U.S. preschoolers

4.1. Method

4.1.1. Participants

We recruited 47 3- to 5-year-olds (24 female, 23 male, mean age = 4.5 years) from preschools in a coastal region in the Western United States. Data from nine additional participants were removed prior to analysis because of experimenter error or because the children did not complete the interview. Parents provided written permission prior to the interview. Parental reports indicated that 85% of children were white, 4% were Latinx, 6% were Asian/Asian American, and 5% were of mixed or other ethnic/racial backgrounds.

Table 2
Coding Scheme for Justifications.

Code	Definition	Example
Consent	Indications that children consented to the activity.	"They both agreed to play the game"
Force evaluation	Mere evaluative labeling of the act of force	"It is never okay to hit somebody"
Game rules	References to the existence of a game, without mentioning consent	"It was okay because they were playing a game"
Non-aggressive strategies	Statement that protagonist could have taken on a non-violent action instead	"She should have used words"
Provocation	Reference to the prior provocation.	"She is doing to the other child what the other child did to her"
Self-Defense	Statement about the self-defense of the protagonist.	"He was defending himself"
Unprovoked	Reference to the victim's innocence or lack of a provocation.	"He was just sitting there and the other child was aggressive"
Welfare	Statement about how the force act would affect the welfare of the victim (non-protagonist).	"Because they could have hurt the other child"
Other	Statements not fitting into above categories.	"He couldn't control himself"

4.1.2. Materials and procedure

The procedures and materials in Study 2A were as in Study 1A, except that Study 2A used eight different order conditions to counterbalance, through a Latin Square design, the combinations of force, contexts, and protagonist genders. All interviews were conducted in English by trained research assistants in a separate space in the child's preschool. In addition to the questions listed in Study 1A, the interviewer asked, after the permissibility questions, why it was (not) okay to use force (*justification*) and, after the emotion attribution questions, whether the act would have been okay if all the teachers said the act was okay (*teacher permission*) and what, if anything, the protagonist could have done instead of using force (*alternative action*).

4.2. Data coding and analysis

Justification responses were coded into categories derived from prior research and preliminary review of a subset of the data: *Consent*, *Force evaluation*, *Game rules*, *Item retrieval*, *Non-aggressive strategies*, *Provocation*, *Self-defense*, *Unprovoked*, *Welfare*, and *Other* (Table 2). For responses to the questions about alternative actions, coders assessed whether children proposed any alternative courses of action (e.g., "talk to him" or "get help from a teacher"). To assess interrater agreement, two independent coders double-coded 20% of the data. For justifications, average κ_{Cohen} was .92. For alternative actions, κ_{Cohen} was .78.

Data were analyzed as in Study 1A. Justification and alternative action codes were converted to dichotomous variables representing whether a participant had a given code for a given vignette. Only justification categories used more than 10% of the time were analyzed.

4.3. Results

4.3.1. Permissibility judgments

Preschoolers' permissibility judgments varied by the antecedent event, $D(3) = 9.79, p = .020$ (Table 1). Participants were more likely to say that the protagonist's action was okay in the *property* (16%) and *play* (16%) events than in the *unprovoked* (8%) and *self-defense* (5%) events, $W_s(1) \geq 4.44, ps < 0.035$. No other comparisons among antecedent events were significant, $ps > 0.47$. Permissibility judgments also varied by type of force, $D(3) = 16.58, p < 0.001$. Participants judged *hitting* (18%) as permissible more often than *throwing* (9%) or *pulling* (3%), $W_s(1) > 5.08, ps < 0.025$, and they judged *bumping* (15%) as permissible more often than *pulling*, $W(1) = 6.58, p = 0.010$. No other comparisons were significant, $W_s(1) < 2.28, ps > 0.13$. Younger children were more likely than older children to provide positive evaluations, as evidenced by a significant effect of child age, $D(1) = 4.51, p = 0.034$ (3-year-olds: 21% okay, 4-year-olds: 7%, 5-year-olds: 7%). In contrast, participant gender was not a significant predictor of permissibility judgments, $D(1) = 2.04, p = 0.15$.

4.3.2. Justifications for permissibility judgments

The most common justifications for why the forceful act was okay were references to the victim's *welfare* (44% of okay judgments, e.g., "it doesn't hurt them"), existing *rules* (28%), and *consent* (16%, Table 3). None of these justification types varied significantly by antecedent event, act type, participant age, or gender, $ps > 0.06$. The most common justifications for why the forceful act was wrong were references to the victim's *welfare* (71% of cases, e.g., "He'd get hurt") and *other* (16%, e.g., "because I don't do that to my brother"). None of the justification types varied significantly by antecedent event, act type, participant age, or gender, $ps > 0.15$.

4.3.3. Severity rating

In 7% of cases, participants rated the protagonist's action as "not bad." The propensity to say that the action was "not bad" did not depend significantly on antecedent event, $D(3) = 6.07, p = 0.11$, type of force, $D(3) = 2.69, p = 0.44$, age, $D(1) = 0.12, p = 0.73$, or gender, $D(1) = 2.72, p = 0.10$.

Table 3

Justifications for permissibility judgments.

Code	Study 2A: U.S. Preschoolers		Study 2B: U.S. Adults	
	Not okay	Okay	Not okay	Okay
Consent	0.02	0.16	0.02	0.21
Force evaluation	0.07	0.04	0.29	0.04
Game rules	0.01	0.28	0.07	0.33
Non-aggressive strategies	0.05	0.00	0.25	0.03
Provocation	0.01	0.04	0.12	0.05
Self-Defense	0.00	0.00	0.03	0.36
Unprovoked	0.02	0.00	0.23	0.02
Welfare	0.71	0.44	0.23	0.19
Other	0.16	0.16	0.08	0.06

Note. Cells show the proportion of cases in which participants provided a given justification for a given judgment. For instance, the upper-left cell indicates that Study 2A participants provided *Consent* justifications in 2% of the cases when they judged the force as permissible.

4.3.4. Victim emotion attribution

Children attributed negative emotions to the victim in 71% of cases. Older children were more likely than younger children to attribute negative emotions to the victim, $D(1) = 13.97$, $p < 0.001$ (3-year-olds: 54% negative, 4-year-olds: 73%, 5-year-olds: 87%). Attribution of negative emotions to victims did not vary significantly by antecedent event, $D(1) = 2.00$, $p = 0.57$, force type, $D(3) = 0.60$, $p = 0.90$, or gender, $D(1) = 1.69$, $p = 0.19$.

Next, we fitted a separate model testing whether children were less to judge the action as okay when they attributed negative emotions to the victim. As expected, children were more likely to say that acts of force were okay when they attributed neutral or positive emotions (19% okay, 21/108 trials) than when they attributed negative emotions to the victim (8% okay, 21/265 trials), $D(1) = 7.63$, $p = .006$, controlling for antecedent event, force type, age, and gender.

4.3.5. Protagonist emotion attribution

Overall, children attributed negative emotions to the protagonist in 43% of cases. Negative emotion attributions to the protagonist did not vary significantly by antecedent event, $D(3) = 2.04$, $p = 0.56$, force type, $D(3) = 1.69$, $p = 0.64$, age, $D(1) = 1.43$, $p = 0.23$, or gender, $D(1) = 3.45$, $p = 0.06$.

Emotion attributions to the protagonist did not significantly predict permissibility judgments, $D(1) = 3.60$, $p = .06$, when controlling for antecedent event, force type, age, and gender.

4.3.6. Authority permission

Children were more likely to judge that the protagonist's action was okay when the teachers had given permission (36%) than in their initial judgments (11%), $D(1) = 123.61$, $p < 0.001$. There was no significant interaction between authority permission and antecedent event, $D(3) = 4.30$, $p = 0.23$. As in the analyses of the initial permissibility judgments, the effects of antecedent event, $D(3) = 15.64$, $p = 0.001$, force type, $D(3) = 14.56$, $p = 0.002$, and age, $D(1) = 8.28$, $p = 0.004$, remained significant, whereas the effect of participant gender was again not significant, $D(1) = 1.07$, $p = 0.30$.

4.3.7. Alternative actions

When asked whether the protagonist could have done anything differently, 52% proposed an alternative action (e.g., "Talk to her"). There were no significant effects of antecedent event, $D(3) = 1.36$, $p = 0.75$, force type, $D(3) = 3.34$, $p = 0.34$, age, $D(1) = 0.70$, $p = 0.40$, or gender, $D(1) = 1.80$, $p = 0.18$.

Next, we predicted participants' permissibility judgments from their responses about alternative actions. As expected, children who provided no alternatives were more likely to say that the protagonist's action was okay (17% okay) than participants who suggested an alternative course of action (6%), $D(1) = 13.06$, $p < .001$, when controlling for antecedent event, force type, age, and gender.

5. Study 2B: U.S. adults

5.1. Method

5.1.1. Participants

The U.S. adult sample consisted of 40 undergraduate students (33 female, 7 male, mean age = 21.1 years). Participants were recruited from a research participant pool at a large public university in the Western United States and received course credit for their participation. One participant was removed because they did not finish the interview.

5.1.2. Materials and procedure

Procedures in Study 2B were as in Study 2A with the following exceptions: Adults were interviewed about 16 vignettes, representing all possible combinations of the four antecedent events and the four force types. The order of presentation was counterbalanced using a Latin square design to create four order conditions. The adult interviews did not include the pictures that were used to facilitate child comprehension of the vignettes. When adults were asked to rate the severity of the protagonist's violation, they were asked to point to a pictorial scale representing five faces, from very negative to very positive. A final difference was that adults' emotion attributions were assessed using open-ended questions ("How do you think ... felt after ...?").

5.2. Data coding and analysis

Data were coded and analyzed as in Study 2A, except that coders classified adults' emotion attributions as either negative or non-negative (neutral or positive). As in the other studies, severity ratings were converted into a dichotomous variable indicating whether participants gave a negative rating (one of the two negative faces) or a non-negative rating. As in Study 2A, two coders double-coded 20% of the data to allow for assessment of interrater agreement. Average agreements were as follows: Justifications: $\kappa_{\text{Cohen}} = 0.88$. Emotion attribution: $\kappa_{\text{Cohen}} = 0.92$. Alternative actions: $\kappa_{\text{Cohen}} = 0.78$.

5.3. Results

5.3.1. Permissibility judgments

Permissibility judgments varied by antecedent event, $D(3) = 234.73$, $p < 0.001$ (Table 1). Participants were more likely to say that

force was okay in the *play* (60%) and *self-defense* (51%) events than the *property* (12%) and *unprovoked* (4%) events, $Ws(1) > 58.15$, $ps < 0.001$. The *property* and *unprovoked* events also differed significantly, $Ws(1) = 8.03$, $p = 0.005$, whereas the *play* and *self-defense* events did not, $D(1) = 3.41$, $p = 0.06$. Male participants judged the force as okay more often (45%) than females (29%), $D(1) = 4.47$, $p = 0.035$. There was no significant effect of force type on permissibility judgments, $D(3) = 7.52$, $p = 0.06$.

5.3.2. Justifications for permissibility judgments

Justifications for why the force act was permissible referenced *Self-defense* (36%), *Rules* (33%), *Consent* (21%), and others' *Welfare* (19%). *Self-defense* references were most common in the *self-defense* event (86%, vs. 23% for *property* and 0% for *game* and *unprovoked*). All references to *Rules* occurred in the *play* event (69%, vs. 0% for other contexts), $D(3) = 139.86$, $p < 0.001$. References to *consent* were most common in the *play* (39%) and *unprovoked* events (33%, vs. *property*: 6%, *self-defense*: 0%), $D(3) = 53.57$, $p < 0.001$. *Welfare* references varied significantly by force type (*Hitting*: 33%, *throwing*: 21%, *pulling*: 13%, *bumping*: 8%), $D(3) = 10.28$, $p = 0.016$. No other effects were significant, $ps > 0.05$.

Justifications for why the force act was wrong referenced *force evaluations* (29%), *non-aggressive alternatives* (25%), the *unprovoked* nature of the act (23%), the *welfare* of others (23%), and the presence of *provocations* (12%). *Force evaluations* varied significantly by force type, $D(3) = 8.62$, $p = 0.035$ (*Hitting*: 36%, *pulling*: 33%, *throwing*: 25%, *bumping*: 21%) and were more common among females (32%) than among males (10%), $D(1) = 7.14$, $p = 0.008$. *Non-aggressive alternative* references were most common in the *self-defense* (52%) and *property* events (vs. 8% for *play* and 6% for *unprovoked* contexts), $D(3) = 103.38$, $p < 0.001$. Not surprisingly, statements about how the act was *unprovoked* was the most common in the *unprovoked* contexts (59%, vs. 3–4% for the other three events), $D(3) = 50.61$, $p < 0.001$. Lastly, references to others' *welfare* were most common in the *game* event (34%, vs. 26% for *unprovoked*, 21% for *property*, and 13% for *self-defense*), $D(3) = 12.55$, $p = 0.006$, and also varied by force type, $D(3) = 19.25$, $p < 0.001$ (*Bumping*: 29%, *Pulling*: 26%, *Throwing*: 25%, *Hitting*: 11%). No other effects were significant, $ps > 0.05$.

5.3.3. Severity ratings

Participants' non-negative ratings, akin to the "not bad" ratings in the other studies, varied significantly by antecedent event, $D(3) = 81.55$, $p < .001$. Participants' provided significantly more non-negative ratings for the *self-defense* (20%) and *play* (16%) events than for the *property* (5%) and *unprovoked* (0%) events, $Ws(1) > 12.45$, $ps < 0.001$. No other comparisons were significant, $Ws(1) < 1.86$, $ps > 0.17$. There were no significant effects of force type, $D(3) = 0.80$, $p = .85$, or gender, $D(1) = 0.17$, $p = .68$.

5.3.4. Victim emotion attribution

Adults attributed negative emotions to victims in 82% of cases. Adults' attribution of negative emotions to the victim varied by antecedent event, $D(3) = 33.57$, $p < .001$. Participants attributed negative emotions to the victim less often in the *play* events (69%) than in the *self-defense* (82%), *unprovoked* (87%), and *property* (90%) events, $Ws(1) > 8.69$, $ps < 0.004$. In addition, the *self-defense* and *property* events differed significantly, $W(1) = 5.38$, $p = .020$. No other pairwise comparisons were significant, $Ws(1) < 2.14$, $ps > 0.14$. There were no significant effects force type, $D(3) = 7.31$, $p = .06$, or participant gender, $D(1) = 0.41$, $p = .52$.

As hypothesized, participants were more likely to say that acts of force were okay when they attributed positive or neutral emotions to the victim (56% okay, 64/115 trials) than when they attributed negative emotions (27%, 129/385 trials), $D(1) = 7.46$, $p = .006$, when controlling for antecedent event, force type, and gender.

5.3.5. Protagonist emotion attribution

For protagonists, adults attributed negative emotions in 42% of cases. Protagonist negative emotion attributions varied by antecedent event, $D(3) = 153.93$, $p < .001$, being more common in *self-defense* (60%) and *property* (56%) contexts than in *unprovoked* (40%) events, which in turned differed from *play* events (11%), $Ws(1) > 13.04$, $ps < 0.001$. *Self-defense* and *property* events did not differ significantly, $W(1) = 0.71$, $p = .40$. There were no significant effects of force type, $D(3) = 1.01$, $p = .80$, or gender, $D(1) = 0.55$, $p = .46$.

Participants were more likely to judge the protagonist's action to be okay when they attributed no negative emotions to the protagonist (40% okay, 147/371 trials) than when they did attribute negative emotions to the protagonist (21%, 56/268 trials), $D(1) = 15.76$, $p < .001$, again controlling for antecedent event, force type, and gender.

5.3.6. Authority permission

Authority permission did not significantly affect adults' judgments about force, $D(1) = 0.66$, $p = .42$, nor was there any interaction between context and adult permission, $D(3) = 5.75$, $p = .12$. As in the analyses of the initial permissibility judgments, there were still significant effects of antecedent event, $D(3) = 413.23$, $p < .001$, force type, $D(1) = 11.00$, $p = .012$, and gender, $D(1) = 5.26$, $p = .022$.

5.3.7. Alternative actions

Participants' propensity to suggest non-aggressive alternative actions varied by antecedent event, $D(3) = 26.77$, $p < .001$. Participants were less likely to state alternative actions in the *play* event (83%) than in the *self-defense* (92%), *unprovoked* (95%), or *property* (96%) events, $Ws(1) > 8.31$, $ps < 0.004$. No other pairwise comparisons were significant, $Ws(1) < 2.16$, $ps > 0.14$, and there were no significant effects of force type, $D(3) = 0.23$, $p = .97$, or gender, $D(1) = 0.31$, $p = .58$.

Next, we fitted a separate model testing whether adults were more likely to judge the action as okay when they suggested a non-aggressive alternative action. As hypothesized, participants who provided no non-aggressive alternatives more often accepted the protagonist's action (65% okay) than participants who did propose an alternative action (29%), $D(1) = 5.23$, $p = .022$, when controlling for antecedent event, force type, and gender.

5.4. Study 2 discussion

Studies 2A and 2B provided further support for two of our main hypotheses. In our U.S. samples, both preschoolers and adults incorporated antecedent events into their judgments about force. As expected, participants were especially likely to think that force was okay in the *play* events, as well as the *property* conflict (especially preschoolers) and *self-defense* events (especially adults). Second, both children and adults were more likely to judge the act of force permissible if they attributed positive or neutral emotions, rather than negative emotions, to the victim.

Studies 2A and 2B also went beyond Studies 1A and 1B. Participants' justifications provided further evidence about how children and adults make judgments about permissible and impermissible force. In both child and adult samples, participants often reasoned in terms of others' welfare, existing rules, and consent when they explained why a given act of force was or was not okay. These justifications mapped onto the features we had manipulated, such as whether the victim had willingly entered into a forceful game, further pointing to the centrality of the features studied here for developing judgments about permissible and impermissible force.

Two further additions of Studies 2A and 2B were the inclusion of questions about teacher permission and alternative actions. Although preschoolers became more accepting of force when they imagined that teachers had given permission, adults did not. When prompted about alternative actions, most children and adults suggested non-aggressive courses of actions that the protagonist could have taken instead of using force. As predicted, participants were less accepting of acts of force when they believe that the aggressor could have taken other courses of action. For instance, during a property conflict, if they believed that a protagonist could have used verbal means to retrieve their toy instead of hitting, children and adults became more likely to judge that hitting was wrong. We will discuss the implications of each of these points in the General Discussion.

6. Pooled analyses

Before turning to the General Discussion, we report analyses that pool data on permissibility judgments across the four studies. (We thank the editor and reviewers of a prior submission for suggesting these analyses.) Since the research was not designed with these analyses in mind, and the studies differed along multiple dimensions, these analyses must be interpreted with caution, especially in light of the limited sample size of Study 1A. Our main question was whether the effect of antecedent event and force type depended on age group (preschool vs. adult) and country (United States vs. Turkey).

These analyses revealed a significant three-way interaction between age group, context, and country, $D(3) = 14.75, p = .002$. That is, the interaction between age group and antecedent events varied significantly between Turkey and the United States. Although the interaction between age group and antecedent event was significant for both countries, $Ds(3) > 22.34, ps < 0.001$, the interaction was more pronounced in the U.S. data than in the Turkish data. In the Turkey samples, the difference between preschoolers' and adults' judgments was only significant for the *play* events, $W(1) = 7.63, p = .006$ (preschoolers: 8% okay, adults: 69% okay). In the U.S. samples, the difference was significant for both the *self-defense* (preschoolers: 5%, adults: 51%) and *play* (preschoolers: 16%, adults: 60%) events, $Ws(1) > 12.38, p < .001$. Pairwise comparisons between children and adults for the remaining antecedent events were not significant, $Ws(1) < 2.11, ps > 0.14$.

In contrast, there was no significant three-way interaction between force type, age group, and country, $D(3) = 3.94, p = .27$. The effect of force type did vary significantly by country, $D(3) = 9.20, p = .027$. U.S. participants were significantly more accepting of *hitting* with a toy animal (30%) than were Turkish participants (16%), $W(1) = 7.93, p = .005$, but the two countries did not differ significantly in the use of other types of force, $Ws(1) < 1.85, ps > 0.17$. There was no significant interaction between force type and age-group, $D(3) = 7.44, p = .059$.

7. General discussion

As they grow up, children develop judgments about when and how it is wrong to use force against others. The present research demonstrated that distinctions between permissible and impermissible force emerge by the preschool years. It also identified several factors that predicted judgments about force among both Turkish and U.S. preschoolers and adults.

First, the nature of the antecedent event shaped judgments about permissible and impermissible force among both preschoolers and adults. Compared to the *unprovoked* events, participants were significantly more accepting of force acts in response to property conflicts (Studies 2A, 2B), *self-defense* (Studies 1A, 1B, 2B), and physical play (Studies 1B, 2A, 2B). Similar patterns were evident for severity ratings among adults (Studies 1B, 2B), though not for children, who nearly always gave negative ratings of the protagonist's action. The judgments of preschoolers prefigure the judgments of older children, who are also more accepting of force in *self-defense* and in response to provocations (Ardila-Rey et al., 2009; Astor, 1994; Nucci et al., 2017).

The findings on children's judgments about force during physical play are particularly noteworthy since so little research has examined children's judgments about force during play. Physical play constitutes a key context of early social development that many, if not most, children encounter (Smith & Boulton, 1990; Tannock, 2008). Perceptions of the consent, or willingness, of participants to engage in rough-and-tumble play may be one key factor that lead children to view forceful acts as permissible in a play situation. Evaluations of physical play may be especially susceptible to contextual variation, since a given physical act—like pulling someone to the ground—may be acceptable in one kind of physical play but not another, depending on what the interactants have agreed on. Similarly, in contact sports for adults, tackles that are permitted in football are prohibited in basketball. Such contextual variability in norms about physical play may help explain the age and country differences observed in the present research (see below). It will be important to examine how specific, direct experiences during rough-and-tumble play and related activities shape children's social

understandings and judgments.

The comparisons between children and adults evidenced considerable developmental change in judgments about force. Although the effect of antecedent event was significant for all four studies, adults distinguished more sharply between the antecedent events than did children. For instance, whereas nearly all children and adults agreed that force was wrong in the *unprovoked* events, adults were far more likely than children to judge that force was okay in the *play* events. This developmental difference may be surprising since it is preschoolers—not college students—who routinely encounter the types of physical play depicted in the vignettes. Unexpectedly, in Study 2A, younger children were significantly *more* likely than older children to judge the acts of force as okay, perhaps because they were also less likely to attribute negative emotions to the victim (see below). These age differences within and beyond the preschool age highlight how judgments about force in multifaceted contexts continue to develop throughout childhood. Indeed, Nucci and colleagues (2017) found major age differences in judgments about multifaceted force events as late as the period from middle childhood and adolescence.

A plausible explanation for the difference between preschoolers and adults is that the *play* context, like the *self-defense* and *property* contexts, presented children with two competing considerations. Physical play pits the general prohibition against interpersonal force against the victim's apparent acceptance of force. Research on judgments about helping, resource distribution, emotion attribution, and other topics in socio-moral development has shown that young children often struggle to incorporate competing considerations (Arsenio & Kramer, 1992; Dahl et al., 2020; Killen et al., 2018). As children grow older, they may become better at coordinating conflicting considerations and thus more reliably determine when to make exceptions from the general prohibition against force.

Justification data in Studies 2A and 2B revealed several considerations that likely lead participants to accept acts of force. Children most often mentioned the victim's welfare when explaining their judgments about force, though they also mentioned contextual considerations such as existing rules or consent. In addition to these considerations, adults also referenced more specific antecedent events, discussing prior provocations or the need for self-defense. The justifications show how judgments about force, even among young children, draw on multiple considerations that can conflict, as when authorities require one person to harm another (Fiske & Rai, 2014; Pnevmatikos, 2018). It will be important to examine children's reasoning about force using more diverse samples from multiple communities in future research.

The findings from emotion attributions further demonstrate the centrality of victim welfare for judgments about force (Gray et al., 2012; Wainryb et al., 2005). Across the four studies, participants who attributed neutral or positive emotions to the victim were more likely to judge the action as okay. (Attribution of protagonist emotions significantly predicted judgments only for adults, not for children.) This finding reveals how differing perceptions of social consequences may lead to differing judgments: When one person perceives harm and the other does not, their judgments will tend to differ accordingly (Helwig et al., 2001; Killen, Mulvey, Richardson, Jampol & Woodward, 2011; Wainryb, 1991). Attributions of emotions to others become more accurate with age (Arsenio & Kramer, 1992; Wainryb et al., 2005): In Study 2A, older children were more likely than younger children to attribute negative emotions to the victim. Learning how their actions affect the welfare of others in everyday life is a major task in early socio-moral development (Dahl & Turiel, 2019).

Children and adults attributed fewer negative emotional states to protagonists than to the targets of force acts. Still, even preschoolers attributed negative emotions to protagonists in as many as 40–80% of trials. Although prior work has found that younger preschoolers attribute positive emotional states to "victimizers" more than older children, this work has often used scenarios in which the victimizers attained some material gain by using force (Arsenio & Kramer, 1992; Jambon et al., 2021). In contrast, protagonists in the present research did not always use force for material gain, for instance, in the unprovoked and play events. Indeed, except in Study 2A, participants were least likely to attribute negative emotions to the protagonist in the play events. This raises the possibility that the attribution of positive states to "victimizers" may depend on the context of the force, even among young children.

Studies 2A and 2B also examined whether authority permission shapes judgments about force. Preschoolers became significantly more likely to judge that force acts were okay when they imagined that teachers had given permission. As noted, preschoolers also often mentioned existing rules when explaining why the force acts were wrong. It may be surprising to find that children's judgments about force were sensitive to authority permission. Much prior research has found that judgments about force are less sensitive to authority permission than other judgments, such as judgments about conventional or religious violations (Noh et al., 2020; Smetana, 2013; Smetana et al., 2012; Srinivasan et al., 2019). Still, we view the present effects of authority permission on force judgments as compatible with prior work. First, the present study did not include any non-force violations. Hence, it remains possible that force judgments were still *less* sensitive to authority permission than judgments about non-moral violations would have been. Indeed, prior studies have found *some* effects of authority commands on children's judgments about moral violations, especially in multifaceted contexts, even if these are smaller than authority effects on judgments about non-moral violations (Braine et al., 1991; Smetana et al., 2021; Srinivasan et al., 2019; Tisak, 1986). Second, the force acts in the vignettes were likely less severe than force acts used in prior research on preschoolers' moral judgments, and thus perhaps located in more of a gray area than more harmful hits or pushes would have been. Children in Study 2A may have assumed, for instance, that if teachers gave permission, the acts must not have been very harmful.

Unlike preschoolers' judgments, adults' judgments about force in Study 2B were not significantly affected by authority permission. In their daily lives, preschoolers likely rely more on teachers' advice about how to treat others than do adults. (Although adults, like children, do grant teachers the right to alter conventions, such as dress codes: Dahl & Waltzer, 2020.) Authority commands may be especially influential when children encounter ambiguous or complex situations, which authorities can help disambiguate (Turiel & Dahl, 2017). Indeed, we surmise that adults did find the vignettes easier to process, both with respect to the perceived harm to the victim and the coordination of competing considerations, since adults differentiated more sharply than children among the antecedent events. As noted earlier, children's abilities to perceive harm and smoothly coordinate competing considerations grow dramatically

from early childhood to adolescence (e.g., Nucci et al., 2017; Wainryb et al., 2005). Heightened sensitivity to authority commands in early childhood could offer one explanation for the emergence, by middle childhood, of cultural differences about whether it is okay to hit in self-defense and other multifaceted contexts (Ardila-Rey et al., 2009; Astor, 1994).

Studies 2A and 2B also revealed that, as predicted, children and adults were less accepting of force when they believed the protagonist could have used non-forceful alternatives. This points to how violence, in some contexts, is seen as a last resort to defend oneself or one's interest rather than an inherently valuable option. As prior research has suggested, people judge violations more harshly when better alternatives are available (Buckwalter & Turri, 2015; Kurthy et al., 2017). The generation of alternative actions constitutes another area of social development that may shape children's judgments and decisions (Arsenio & Lemerise, 2004; Dahl & Turiel, 2019). Overall, adults were far more likely than preschoolers to suggest non-aggressive alternative courses of action that the protagonist could have taken.

The predictors of force acceptance do not operate in isolation but in combination. When preschoolers in Study 2A attributed a neutral or positive emotion to the victim in the *property* context, acceptance of force reached 31%. As suggested by the Social Information Processing Model of aggression, it is often the accumulation of multiple factors that lead children to accept and use force in a given situation (Arsenio & Lemerise, 2004; Dodge et al., 2006).

In the present studies, some analyses revealed significant relations between force type and judgments or ratings. Since these effects were neither predicted nor systematic, we hesitate to interpret them. Still, we have no doubt that force type generally does affect judgments. In particular, had we systematically varied the severity of the force acts in the present study—contrasting, say, physical force that yielded grave physical injury with the physical force of a high five—we would likely have found lower acceptance of the more severe force acts. Here, however, we designed the action types to be of intermediate severity so that they could plausibly occur in any of the antecedent contexts, including physical play.

On the whole, the similarities between the Turkish and U.S. samples were more pronounced than the differences: Children and adults from both samples distinguished between permissible and impermissible force and almost never thought unprovoked force was okay. Patterns of judgments of Turkish and U.S. adults were particularly similar, whereas judgments of preschoolers differed by which context elicited the most acceptance of force (*self-defense* among Turkish children and *property conflict* and *play* among U.S. children). The findings also revealed some differences between cultural differences. Most notably, the difference between children and adults was more pronounced among U.S. participants than among Turkish participants for the self-defense context (Table 1). The cultural differences in the present study must be interpreted with caution given the small sample size of Study 1A, the fact the study populations differed in multiple ways beyond age and country, and that the study was not designed to explain cross-cultural differences. Our speculations about the sources of this three-way interaction can thus only be tentative. One possibility is that, compared to U.S. participants, the Turkish children had more exposure to neighborhood violence, whereas the Turkish adults had less exposure to violence, leading to intermediate levels of disapproval in the self-defense context among Turkish children (24% okay) and adults (34%). Research from many communities has shown that individuals who grow up with more exposure to violence tend to be more accepting of using force in self-defense and against other provocations (Ardila-Rey et al., 2009; Astor, 1994; P. Miller & Sperry, 1987).

Given the limitations of the present cross-cultural comparison, and the paucity of research on early moral development in Turkey (Kuyel & Glover, 2010; Nisan & Kohlberg, 1982; Turiel et al., 1978), it will be important to seek to replicate these findings on both age and cultural differences with larger and more diverse samples. In addition, it will be valuable to collect additional information about children's social context—including adult views on children's physical play and self-defense—in order to explain cultural variability and similarity in the development of children's judgments about force.

By middle childhood, children from many, if not all, communities judge that it is sometimes okay to use force against others (e.g., Dahl et al., 2018; Jambon & Smetana, 2018; Noh et al., 2020; Nucci et al., 2017; Pnevmatikos, 2018; Wainryb et al., 2005). At the same time, as noted, children with more exposure to violence and aggression appear to become more accepting of interpersonal force in response to provocations (Ardila-Rey et al., 2009; Astor, 1994; P. Miller & Sperry, 1987). More work is needed to reveal the developmental processes that connect the rudimentary distinctions among force acts evident in preschoolers to the more sophisticated and culturally adapted judgments evident later in childhood and into adulthood. The present research looked specifically at the kinds of situations young children are likely to encounter in the everyday life of preschools and families, which differed from events used in prior research with older children and adolescents (e.g., Noh et al., 2020; Nucci et al., 2017; Pnevmatikos, 2018). It may be precisely through such everyday encounters with force that children develop their early ideas about when it is okay to hit.

To bridge the gaps in our knowledge about how preschoolers develop nuanced judgments about force, research will need to examine how children use their everyday experiences to develop such judgments (Dahl & Turiel, 2019). Many children in the present study judged that hitting or shoving was wrong even in the play context. How do these children respond when they partake in or witness physical play in which the recipient of the force laughs instead of being angry? We expect that repeated encounters like these in everyday life will be critical in shaping children's notions about when and how to use force against others. Another likely factor in the development of judgments about force is the developing cognitive abilities to balance multiple considerations at once, such as the ability to weigh the right to self-defense against the general prohibition of harmful actions (Killen et al., 2018; Nucci et al., 2017).

The present studies speak to a key yet underexamined question in early moral development: How do children in communities across the world begin to distinguish permissible from impermissible force? Although many of the most widely condemned moral violations are acts of violence, societies are also full of acts of force that many individuals come to accept or even encourage. Football players are celebrated as heroes for their ability to use force and inflict pain on others. Laws enshrine rights to self-defense. This research traced the sources of such context-sensitive judgments about force to the preschool age, yet it also highlighted that these judgments continue to develop beyond early childhood.

Author Note

We thank members of the Developmental Moral Psychology Lab at the University of California, Santa Cruz for comments on earlier versions of this manuscript.

Declarations of Interest

None

References

- Baxley, C., & Dahl, A., in press, Developing revenge in early childhood: Current evidence and future directions. In H. E. Recchia & C. Wainryb (Eds.), *Retribution and revenge across childhood and adolescence*. Cambridge University Press.
- Ardila-Rey, A., Killen, M., & Brenick, A. (2009). Moral reasoning in violent contexts: Displaced and non-displaced Colombian children's evaluations of moral transgressions, retaliation, and reconciliation. *Social Development, 18*(1), 181–209. <https://doi.org/10.1111/j.1467-9507.2008.00483.x>
- Arsenio, W. F., & Kramer, R. (1992). Victimizers and their victims: Children's conceptions of the mixed emotional consequences of moral transgressions. *Child Development, 63*(4), 915–927. <https://doi.org/10.2307/1131243>
- Arsenio, W. F., & Lemerise, E. A. (2004). Aggression and moral development: Integrating social information processing and moral domain models. *Child Development, 75*(4), 987–1002. <https://doi.org/10.1111/j.1467-8624.2004.00720.x>
- Astor, R. A. (1994). Children's moral reasoning about family and peer violence: The role of provocation and retribution. *Child Development, 65*(4), 1054–1067. <https://doi.org/10.2307/1131304>
- Baker, E. R., & Liu, Q. (2021). Moral reasoning and moral behavior: Intersections of reasoning with aggressive forms and functions in early childhood. *Early Education and Development, 32*(4), 534–552. <https://doi.org/10.1080/10409289.2020.1780561>
- Ball, C. L., Smetana, J. G., & Sturge-Apple, M. L. (2017). Following my head and my heart: Integrating preschoolers' empathy, theory of mind, and moral judgments. *Child Development, 88*(2), 597–611. <https://doi.org/10.1111/cdev.12605>
- Braine, L. G., Pomerantz, E., Lorber, D., & Krantz, D. H. (1991). Conflicts with authority: Children's feelings, actions, and justifications. *Developmental Psychology, 27*(5), 829–840. <https://doi.org/10.1037/0012-1649.27.5.829>
- Brownlee, J. R., & Bakeman, R. (1981). Hitting in toddler-peer interaction. *Child Development, 52*(3), 1076–1079. <https://doi.org/10.2307/1129115>
- Buckwalter, W., & Turri, J. (2015). Inability and obligation in moral judgment. *PLOS ONE, 10*(8), Article e0136589. <https://doi.org/10.1371/journal.pone.0136589>
- Caplan, M., Vespo, J., Pedersen, J., & Hay, D. F. (1991). Conflict and its resolution in small groups of one- and two-year-olds. *Child Development, 62*(6), 1513–1524. <https://doi.org/10.2307/1130823>
- Dahl, A. (2016). Infants' unprovoked acts of force toward others. *Developmental Science, 19*(6), 1049–1057. <https://doi.org/10.1111/desc.12342>
- Dahl, A., & Freda, G. F. (2017). How young children come to view harming others as wrong: A developmental analysis. In J. A. Sommerville, & J. Decety (Eds.), *Social Cognition: Development across the Life Span* (pp. 151–184). Routledge.
- Dahl, A., Gingo, M., Uttich, K., & Turiel, E. (2018). Moral reasoning about human welfare in adolescents and adults: Judging conflicts involving sacrificing and saving lives. *Monographs of the Society for Research in Child Development, 83*(3), 1–109. <https://doi.org/10.1111/mono.12374>
- Dahl, A., Gross, R. L., & Siefert, C. (2020). Young children's judgments and reasoning about prosocial acts: Permission, obligation, and supererogation. *Cognitive Development, 55*, Article 100908.
- Dahl, A., & Killen, M. (2018). Moral reasoning: Theory and research in developmental science. In J. T. Wixted, & S. Ghetti (Eds.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th ed., pp. 323–353). Wiley.
- Dahl, A., & Kim, L. (2014). Why is it bad to make a mess? Preschoolers' conceptions of pragmatic norms. *Cognitive Development, 32*, 12–22. <https://doi.org/10.1016/j.cogdev.2014.05.004>
- Dahl, A., Sherlock, B. R., Campos, J. J., & Theunissen, F. E. (2014). Mothers' tone of voice depends on the nature of infants' transgressions. *Emotion, 14*(4), 651–665. <https://doi.org/10.1037/a0036608>
- Dahl, A., & Turiel, E. (2019). Using naturalistic recordings to study children's social perceptions and evaluations. *Developmental Psychology, 55*, 1453–1460. <https://doi.org/10.1037/dev0000735>
- Dahl, A., & Waltzer, T. (2018). Moral disengagement as a psychological construct. *American Journal of Psychology, 131*, 240–246.
- Dahl, A., & Waltzer, T. (2020). Constraints on conventions: Resolving two puzzles of conventionality. *Cognition, 196*, Article 104152. <https://doi.org/10.1016/j.cognition.2019.104152>
- Davidson, P., Turiel, E., & Black, A. (1983). The effect of stimulus familiarity on the use of criteria and justifications in children's social reasoning. *British Journal of Developmental Psychology, 1*(1), 49–65. <https://doi.org/10.1111/j.2044-835X.1983.tb00543.x>
- Dodge, K. A., Coie, J. D., & Lynam, D. (2006). Aggression and antisocial behavior in youth. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 719–788). John Wiley & Sons, Inc.. <https://doi.org/10.1002/9780470147658.chpsy0312>
- Fiske, A. P., & Rai, T. S. (2014). *Virtuous violence: Hurting and killing to create, sustain, end, and honor social relationships*. Cambridge University Press.
- Gasser, L., Malti, T., & Gutzwiller-Helfenfinger, E. (2012). Aggressive and nonaggressive children's moral judgments and moral emotion attributions in situations involving retaliation and unprovoked aggression. *The Journal of Genetic Psychology, 173*(4), 417–439. <https://doi.org/10.1080/00221325.2011.614650>
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry, 23*(2), 101–124. <https://doi.org/10.1080/1047840X.2012.651387>
- Hawley, P. H. (2003). Strategies of control, aggression, and morality in preschoolers: An evolutionary perspective. *Journal of Experimental Child Psychology, 85*(3), 213–235. [https://doi.org/10.1016/S0022-0965\(03\)00073-0](https://doi.org/10.1016/S0022-0965(03)00073-0)
- Hay, D. F. (2006). Yours and mine: Toddlers' talk about possessions with familiar peers. *British Journal of Developmental Psychology, 24*(1), 39–52. <https://doi.org/10.1348/026151005x68880>
- Hay, D. F., Hurst, S.-L., Waters, C. S., & Chadwick, A. (2011). Infants' use of force to defend toys: The origins of instrumental aggression. *Infancy, 16*(5), 471–489. <https://doi.org/10.1111/j.1532-7078.2011.00069.x>
- Hay, D. F. (2005). The beginnings of aggression in infancy. In R. E. Tremblay, W. W. Hartup, & J. Archer (Eds.), *Developmental origins of aggression* (pp. 107–132). Guilford Press.
- Helwig, C. C., Zelazo, P. D., & Wilson, M. (2001). Children's judgments of psychological harm in normal and noncanonical situations. *Child Development, 72*(1), 66–81. <https://doi.org/10.1111/1467-8624.00266>
- Hox, J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.,). Routledge.
- Jambon, M., Colasante, T., & Malti, T. (2021). A longitudinal investigation of the happy victimizer tendency in childhood: A matter of control or care? *Developmental Psychology, 57*(5), 689–701. <https://doi.org/10.1037/dev0001176>
- Jambon, M., & Smetana, J. G. (2014). Moral complexity in middle childhood: Children's evaluations of necessary harm. *Developmental Psychology, 50*(1), 22–33. <https://doi.org/10.1037/a0032992>
- Jambon, M., & Smetana, J. G. (2018). Individual differences in prototypical moral and conventional judgments and children's proactive and reactive aggression. *Child Development, 89*(4), 1343–1359. <https://doi.org/10.1111/cdev.12757>

- Killen, M., Elenbaas, L., & Rizzo, M. T. (2018). Young children's ability to recognize and challenge unfair treatment of others in group contexts. *Human Development*, 61(4–5), 281–296. <https://doi.org/10.1159/000492804>
- Killen, M., Mulvey, K. L., Richardson, C., Jampol, N., & Woodward, A. (2011). The accidental transgressor: Morally-relevant theory of mind. *Cognition*, 119(2), 197–215. <https://doi.org/10.1016/j.cognition.2011.01.006>
- Kimble, N. B., Russo, S. A., Bergman, B. G., & Galindo, V. H. (2010). Revealing an empirical understanding of aggression and violent behavior in athletics. *Aggression and Violent Behavior*, 15(6), 446–462. <https://doi.org/10.1016/j.avb.2010.08.001>
- Kurthy, M., Lawford-Smith, H., & Sousa, P. (2017). Does ought imply can? *PLOS ONE*, 12(4), Article e0175206. <https://doi.org/10.1371/journal.pone.0175206>
- Kuyel, N., & Glover, R. J. (2010). Moral reasoning and moral orientation of U.S. and Turkish university students. *Psychological Reports*, 107(2), 463–479. <https://doi.org/10.2466/08.09.17.PR0.107.5.463-479>
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7(1), 19–40. <https://doi.org/10.1037/1082-989X.7.1.19>
- Martin, J. L., & Ross, H. S. (1996). Do mitigating circumstances influence family reactions to physical aggression? *Child Development*, 67(4), 1455. <https://doi.org/10.2307/1131711>
- Miller, J. G., Bersoff, D. M., & Harwood, R. L. (1990). Perceptions of social responsibilities in India and in the United States: Moral imperatives or personal decisions? *Journal of Personality and Social Psychology*, 58(1), 33–47.
- Miller, P., & Sperry, L. L. (1987). The socialization of anger and aggression. *Merrill-Palmer Quarterly*, 33(1), 1–31.
- Mintah, J. K., Huddleston, S., & Doody, S. G. (1999). Justifications of aggressive behavior in contact and semicontact sports. *Journal of Applied Social Psychology*, 29(3), 597–605. <https://doi.org/10.1111/j.1559-1816.1999.tb01403.x>
- Nisan, M., & Kohlberg, L. (1982). Universality and variation in moral judgment: A longitudinal and cross-sectional study in turkey. *Child Development*, 53(4), 865–876. <https://doi.org/10.2307/1129123>
- Noh, J. Y., Jambon, M., Smetana, J. G., Lee, I. J., & Killen, M. (2020). Children's evaluations of necessary harm: The role of maternal disapproval and friend relationship status. *Merrill-Palmer Quarterly*, 66(2), 145–173. <https://doi.org/10.13110/merrpalmquar1982.66.2.0145>
- Nucci, L. P., Turiel, E., & Roded, A. D. (2017). Continuities and discontinuities in the development of moral judgments. *Human Development*, 60, 279–341. <https://doi.org/10.1159/000484067>
- Nucci, L. P., & Weber, E. K. (1995). Social interactions in the home and the development of young children's conceptions of the personal. *Child Development*, 66(5), 1438–1452. <https://doi.org/10.2307/1131656>
- Pellegrini, A. D., & Smith, P. K. (1998). Physical activity play: The nature and function of a neglected aspect of play. *Child Development*, 69(3), 577–598. <https://doi.org/10.1111/j.1467-8624.1998.tb06226.x>
- Pnevmatikos, D. (2018). When do children start to take mitigating circumstances into account when judging the act of killing? *Cognitive Development*, 48, 94–104. <https://doi.org/10.1016/j.cogdev.2018.07.007>
- Rhodes, M., & Chalik, L. (2013). Social categories as markers of intrinsic interpersonal obligations. *Psychological Science*, 24(6), 999–1006. <https://doi.org/10.1177/0956797612466267>
- Rubin, K. H., Bukowski, W. M., & Parker, J. G. (2006). Peer interactions, relationships, and groups. In R. M. Lerner, & N. Eisenberg (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 571–645). Wiley & Sons. <https://doi.org/10.1002/9780470147658.chpsy0310>
- Siegal, M., & Storey, R. M. (1985). Day care and children's conceptions of moral and social rules. *Child Development*, 56(4), 1001–1008. <https://doi.org/10.2307/1130111>
- Smetana, J. G. (1985). Preschool children's conceptions of transgressions: Effects of varying moral and conventional domain-related attributes. *Developmental Psychology*, 21(1), 18–29.
- Smetana, J. G. (2013). Moral development: The social domain theory view. *Oxford Handbook of Developmental Psychology*, 1, 832–866.
- Smetana, J. G., & Ball, C. L. (2018). Young children's moral judgments, justifications, and emotion attributions in peer relationship contexts. *Child Development*, 89(6), 2245–2263. <https://doi.org/10.1111/cdev.12846>
- Smetana, J. G., Campione-Barr, N., & Yell, N. (2003). Children's moral and affective judgments regarding provocation and retaliation. *Merrill-Palmer Quarterly*, 49(2), 209–236. <https://doi.org/10.1353/mpq.2003.0010>
- Smetana, J. G., Daddis, C., Toth, S. L., Cicchetti, D., Bruce, J., & Kane, P. (1999). Effects of provocation on maltreated and nonmaltreated preschoolers' understanding of moral transgressions. *Social Development*, 8(3), 335–348. <https://doi.org/10.1111/1467-9507.00099>
- Smetana, J. G., Jambon, M., & Ball, C. L. (2018). Normative changes and individual differences in early moral judgments: A constructivist developmental perspective. *Human Development*, 61(4–5), 264–280. <https://doi.org/10.1159/000492803>
- Smetana, J. G., Kelly, M., & Twentyman, C. T. (1984). Abused, neglected, and nonmaltreated children's conceptions of moral and social-conventional transgressions. *Child Development*, 55(1), 277–287. <https://doi.org/10.2307/1129852>
- Smetana, J. G., Rote, W. M., Jambon, M., Tasopoulos-Chan, M., Villalobos, M., & Comer, J. (2012). Developmental changes and individual differences in young children's moral judgments. *Child Development*, 683–696. <https://doi.org/10.1111/j.1467-8624.2011.01714.x>
- Smetana, J. G., Yoo, H. N., Nguyen, N., & Ball, C. L. (2021). Moms know best? Children's evaluations of mothers' unfair responses to peer conflicts. *Journal of Experimental Child Psychology*, 201, Article 104993. <https://doi.org/10.1016/j.jecp.2020.104993>
- Smith, P. K., & Boulton, M. (1990). Rough-and-tumble play, aggression and dominance: Perception and behaviour in children's encounters. *Human Development*, 33(4–5), 271–282. <https://doi.org/10.1159/000276524>
- Srinivasan, M., Kaplan, E., & Dahl, A. (2019). Reasoning about the scope of religious norms: Evidence from Hindu and Muslim children in India. *Child Development*, 90, e783–e802. <https://doi.org/10.1111/cdev.13102>
- Tannock, M. T. (2008). Rough and tumble play: An investigation of the perceptions of educators and young children. *Early Childhood Education Journal*, 35(4), 357–361. <https://doi.org/10.1007/s10643-007-0196-1>
- Tisak, M. S. (1986). Children's conceptions of parental authority. *Child Development*, 57, 166–176.
- Tremblay, R. E., Vitaro, F., & Côté, S. M. (2018). Developmental origins of chronic physical aggression: A bio-psycho-social model for the next generation of preventive interventions. *Annual Review of Psychology*, 69, 383–407. <https://doi.org/10.1146/annurev-psych-010416-044030>
- Turiel, E., & Dahl, A. (2017). The development of domains of moral and conventional norms, coordination in decision-making, and the implications of social opposition. In K. Bayertz, & N. Boughley (Eds.), *The normative animal: On the biological significance of social, moral, and linguistic norms*. Oxford University Press.
- Turiel, E., Edwards, C. P., & Kohlberg, L. (1978). Moral development in Turkish children, adolescents, and young adults. *Journal of Cross-Cultural Psychology*, 9(1), 75–86. <https://doi.org/10.1177/002202217891006>
- Turiel, E. (2015). Moral development. In R. M. Lerner (Ed.), *Handbook of Child Psychology and Developmental Science* (Vol. 1, pp. 484–522). John Wiley & Sons, Inc.
- Vitaro, F., & Brendgen, M. (2005). Proactive and reactive aggression: A developmental perspective. In R. E. Tremblay, W. W. Hartup, & J. Archer (Eds.), *Developmental Origins of Aggression* (pp. 178–201). Guilford Press.
- Wainryb, C. (1991). Understanding differences in moral judgments: The role of informational assumptions. *Child Development*, 62(4), 840–851. <https://doi.org/10.2307/1131181>
- Wainryb, C., Brehl, B. A., & 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). <https://doi.org/10.1111/j.1540-5834.2005.00350.x>
- Ward, C. V. (2014). Stand your ground and self-defense. *American Journal of Criminal Law*, 42(2), 89–138.