Journal of Experimental Child Psychology 165 (2018) 161–182

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Differentiating "could" from "should": Developmental changes in modal cognition

Andrew Shtulman [a,*], Jonathan Phillips [b]

[a] Department of Psychology, Occidental College, Los Angeles, CA 90041, USA
[b] Department of Psychology, Harvard University, Cambridge, MA 02138, USA

ARTICLE INFO

Article history:
Available online 23 June 2017

Keywords:
Moral judgment
Possibility judgment
Social reasoning
Physical reasoning
Modality
Conceptual development

ABSTRACT

Young children have difficulty in distinguishing events that violate physical laws (impossible events) from those that violate mere physical regularities (improbable events). They judge both as "impossible." Young children also have difficulty in distinguishing events that violate moral laws (immoral events) from events that violate mere social regularities (unconventional events). They judge both as "wrong." In this set of studies, we explored the possibility that both difficulties arise from a more general deficit in *modal* cognition, or the way in which children represent and reason about possibilities. Participants (80 children aged 3–10 years and 101 adults) were shown impossible, improbable, unconventional, and immoral events and were asked to judge whether the events could occur in real life and whether they would be okay to do. Preschool-aged children not only had difficulty distinguishing law-violating events from regularity-violating events but also had difficulty distinguishing the two modal questions themselves, judging physically abnormal events (e.g., floating in the air) as immoral and judging socially abnormal events (e.g., lying to a parent) as impossible. These findings were replicated in a second study where participants (74 children and 78 adults) judged whether the events under consideration would require magic (a specific consequence of impossibility) or would require punishment (a specific consequence of impermissibility). Our findings imply that young children's modal representations clearly distinguish abnormal events from ordinary events but do not clearly distinguish different types of abnormal events from each other. That

* Corresponding author.
E-mail address: shtulman@oxy.edu (A. Shtulman).

http://dx.doi.org/10.1016/j.jecp.2017.05.012
0022-0965/© 2017 Elsevier Inc. All rights reserved.

is, the distinction between whether an event *could* occur and whether an event *should* occur must be learned.

© 2017 Elsevier Inc. All rights reserved.

Introduction

Predicting other people's behavior is a key part of daily life. We make these predictions on the basis of expectations that differentiate likely behaviors from unlikely behaviors. For instance, if we observe someone at a restaurant discover a fly in his soup, we expect him to ask the waiter for a new bowl of soup or possibly leave the restaurant. We do not expect him to eat the soup with the fly in it or to pull a new uncontaminated bowl of soup out of his backpack. We also do not expect him to swap his bowl of soup for that of his dinner companion when she is in the bathroom or to recite an incantation over the bowl to remove the fly by magic.

Expectations of this nature do not have to be learned situation by situation. Rather, they can be derived from two more general sources of information: social norms and physical regularities (Kalish, 1998; Levy, Taylor, & Gelman, 1995; Lockhart, Abrahams, & Osherson, 1977; Schmidt, Rakoczy, & Tomasello, 2011). Social norms dictate whether a behavior is socially acceptable, whereas physical regularities dictate whether a behavior is physically plausible. Moreover, some social norms specify behaviors that are unconventional but not necessarily immoral (e.g., eating the bowl of soup with the fly in it), whereas others specify behaviors that are patently immoral (e.g., surreptitiously swapping a contaminated bowl of soup with someone else's uncontaminated bowl). Likewise, some physical regularities specify behaviors that are improbable but not necessarily impossible (e.g., pulling a new bowl of soup out of one's backpack), whereas others specify behaviors that are patently impossible (e.g., using a magical incantation to remove the fly).

Most adults make categorical distinctions between these classes of events. They judge unconventional behaviors as less wrong than immoral behaviors (Haidt, Koller, & Dias, 1993; Smetana, 1989; Zalla, Barlassina, Buon, & Leboyer, 2011), and they judge improbable events as more plausible or more imaginable than impossible events (Barnes & Black, 2016; Nolan-Reyes, Callanan, & Haigh, 2016; Shtulman, 2009). Young children, on the other hand, do not make these same distinctions. Several studies (reviewed below) have found that preschool-aged children classify unconventional behaviors as morally wrong and classify improbable events as physically impossible. Children of this age know the social norms that make unconventional behaviors questionable and the physical regularities that make improbable events unlikely, but they do not robustly differentiate events that violate laws—either moral laws or physical laws—from those that violate mere regularities.

Here, we aimed to show that these two phenomena—mistaking improbable events for impossible events and mistaking unconventional events for immoral events—represent a general deficit in modal cognition, or the way in which children represent and reason about possibilities. We show not only that changes in reasoning about physical possibility develop in tandem with changes in reasoning about moral permissibility but also that preschoolers' reasoning about physical possibility is not well differentiated from their reasoning about moral permissibility as a whole. In other words, we show that as children learn to differentiate what could occur in a given situation from what could *not* occur, they also learn to differentiate what *should* occur from what *could* occur in that situation.

Developmental changes in reasoning about physical possibility

By the age of 4 years, children are adept at differentiating impossible events from ordinary events (Samuels & Taylor, 1994; Schult & Wellman, 1997; Sobel, 2004), judging that impossible events could not happen in real life but ordinary events could or that impossible events would require magic to happen but ordinary events would not. This recognition extends to several types of impossible events: physically impossible events, such as moving a marble just by thinking about it (Johnson & Harris,

1994), biologically impossible events, such as growing from an adult back into a child (Rosengren, Kalish, Hickling, & Gelman, 1994), and psychologically impossible events, such as knowing what a person ate just by looking at him (Browne & Woolley, 2004). Even preverbal infants appear to be sensitive to the distinction between impossible events and ordinary events, as revealed by several decades of looking-time experiments (Baillargeon, 2004; Spelke & Kinzler, 2007).

Such findings demonstrate that preschoolers are sensitive to the regularities that govern the physical world and recognize when those regularities have been violated. However, they do not demonstrate that preschoolers are sensitive to the differences between events that are strictly precluded from occurring by laws (e.g., laws of physics) and events that are unlikely to occur but not strictly precluded from occurring. Indeed, 4-year-olds claim not only that events that violate physical laws could not happen in real life but also that events that violate social conventions could not happen either (Browne & Woolley, 2004; Chernyak, Kushnir, Sullivan, & Wang, 2013; Kalish, 1998; Komatsu & Galotti, 1986; Lockhart et al., 1977). That is, they correctly claim that a person could not float in the air or walk through a wall, but they incorrectly claim that a person could not eat candy for dinner or take a bath with his shoes on.

The ability to differentiate events that violate physical laws (impossible events) from those that violate other kinds of statistical regularities (improbable events) develops much later than the ability to differentiate events that violate physical laws from those that violate no regularities whatsoever (ordinary events). Several studies have now shown that preschoolers deny the possibility of events that adults uniformly judge as possible such as owning a lion for a pet, drinking onion juice, or finding an alligator under the bed (Lane, Ronfard, Francioli, & Harris, 2016; Nolan-Reyes et al., 2016; Shtulman, 2009; Shtulman & Carey, 2007; Shtulman & Yoo, 2015; Weisberg & Sobel, 2012). Not until the age of 10 do children robustly differentiate improbable events from impossible events, both in what they judge as possible and in how they justify those judgments. Whereas 4-year-olds justify their judgments of why an event is impossible by citing events that would occur or could occur in place of the target event (e.g., walking through a wall is impossible because "you would hit your head" or because "you could walk through a door instead"), 10-year-olds justify their judgments by citing facts about the world that preclude the events from occurring (e.g., walking through a wall is impossible because "walls are hard and so are people").

Preschoolers' inability to differentiate impossible events from improbable events is not absolute; it varies by culture (Chernyak et al., 2013), by task (Weisberg & Sobel, 2012), and by context (Komatsu & Galotti, 1986). Nevertheless, preschoolers are several times less likely to affirm the possibility of an improbable event than adults are, and teaching preschoolers to differentiate improbable events from impossible events has proven largely ineffective (Lane et al., 2016; Shtulman & Carey, 2007). Preschoolers do not appear to question whether an improbable event violates a law or just a regularity. Instead, they view the fact that it violates a regularity as prima facie evidence of impossibility (see Woolley & Ghossainy, 2013).

Developmental changes in reasoning about moral permissibility

Just as preschoolers readily identify events that violate physical laws as impossible, they readily identify events that violate moral laws as impermissible (Cushman, Sheketoff, Wharton, & Carey 2013; Smetana, 1981). By the age of 4 years, they recognize that intentionally deceiving someone, intentionally hurting someone, or intentionally damaging someone's property is wrong and deserves punishment. Moreover, even 15-month-old infants show some sensitivity to unfair behavior (Schmidt & Sommerville, 2011), and 12-month-olds are sensitive to the distinction between helping someone and harming someone (Hamlin & Wynn, 2011; Hamlin, Wynn, & Bloom, 2007).

At the same time, preschoolers have trouble differentiating events that violate moral laws, such as hitting and stealing, from those that violate mere social regularities, such as not participating in show-and-tell and not sitting on the rug during story time. Although several studies have shown that preschoolers are able to differentiate moral transgressions from conventional transgressions when averaged across items and across children (Dahl & Kim, 2014; Haidt et al., 1993; Nucci, 1981; Turiel, 2008), the size of that distinction is small. Consider, for instance, Smetana's (1981) finding that preschoolers rate moral transgressions as more serious, more punishable, less rule dependent, and less

context dependent than conventional transgressions. Children as young as 3 years provided different ratings for the two types of transgressions, but they still rated conventional transgressions as wrong. That is, they typically rated conventional transgressions as "very bad" (vs. "very, very bad") and as deserving "a little punishment" (vs. "a lot of punishment"), and about a third of these children claimed that conventional transgressions would still be wrong if there were no rule prohibiting them or if they were performed in a different context (e.g., at home rather than at school).

In fact, children as old as 10 continue to view conventional transgressions as wrong. When Tisak and Turiel (1988) asked fifth graders about the permissibility of wearing pajamas to school, 59% claimed it would be wrong for the school to allow children to wear pajamas at school, 55% claimed it would be wrong for children to wear pajamas even if the school did not prohibit it, and 50% claimed it would be wrong for children to wear pajamas even if it were expressly allowed by the school principal. Thus, just as children often deny the possibility of events that are unusual but not impossible, they also often deny the permissibility of events that are unusual but not immoral.

Developmental changes across modal domains

Both of the domains reviewed above—the domain of physical possibility and the domain of moral permissibility—involve reasoning about non-actual events and, thus, both recruit children's capacity for modal cognition. In each domain, young children are able to discriminate between abnormal events and ordinary events, but they are not able to discriminate among separate classes of abnormal events, at least not robustly. They lump violations of physical regularities (improbable events) with violations of physical laws (impossible events), and they lump violations of social regularities (unconventional events) with violations of social laws (immoral events). We propose that this parallel is not a coincidence but rather a general feature of children's modal cognition. Specifically, children first come to understand whether or not an event can occur without representing the particular constraints that preclude particular events from occurring. Only later do they develop the ability to differentiate the reasons why such events cannot (or do not) occur, allowing them to distinguish more clearly between improbable and impossible events and between unconventional and immoral events.

If this is correct, and young children's modal cognition represents all precluded events as similar, it follows that young children should experience difficulty in making modal discriminations *across* domains as well as *within* domains. That is, they should have difficulty in distinguishing the permissibility of socially anomalous events from the permissibility of physically anomalous events, and they should have difficulty in distinguishing the possibility of physically anomalous events from the possibility of socially anomalous events.

Although this prediction may initially seem surprising, there are a number of reasons to believe that the distinction between physical possibility and moral permissibility might not be an inherent feature of modal cognition. First, across languages, there are modal terms that are consistently used both when talking about possibility and when talking about permissibility (Kratzer, 2012; Matthewson, 2016; Nauze, 2008). In English, for example, universal modal auxiliaries such as "must" and "have to" are used to make claims about physical necessity (e.g., "objects have to obey the laws of gravity") as well as claims about moral obligation (e.g., "children have to obey their parents"). Likewise, existential modal auxiliaries such as "can" and "may" are used to make claims about physical possibility (e.g., "there may be more connections in the human brain than stars in the galaxy") as well as claims about moral permissibility (e.g., "you may drive 45 miles per hour"). In general, children are not confused about the meaning of modal auxiliaries; they appropriately comprehend and produce these terms in a well-defined space of possibilities (Byrnes & Duff, 1989; Ozturk & Papafragou, 2015). However, the terms themselves cut across domains, appearing in both physical and moral contexts (Johnson-Laird, 1978).

Second, studies with adults have demonstrated a number of different ways in which reasoning about what can physically occur tracks reasoning about what is morally good (Barnes & Black, 2016; Bear & Knobe, 2016; Tworek & Cimpian, 2016; Young & Phillips, 2011). These findings suggest that the domains of physical possibility and moral permissibility might not be entirely separate even for adults. For instance, Shtulman and Tong (2013) asked college-educated adults to consider the

physical possibility of several extraordinary events (e.g., bringing an extinct species back to life, teleporting an object to a distant location) as well as the moral permissibility of several extraordinary actions (e.g., replacing a borrowed necklace with an exact copy, using an American flag to clean the toilet). They found that adults' tendency to judge physically extraordinary events as possible predicted their tendency to judge morally extraordinary actions as permissible even when controlling for other predictors of moral judgment (e.g., disgust sensitivity). They also found that participants' justifications for their judgments were correlated across domains, with some participants focusing on identifying principles violated by the events at hand and others focusing on identifying circumstances under which the events could occur or should occur. These findings suggest that reasoning about physical possibility and reasoning about moral permissibility draw on common cognitive resources even for adults.

Third, young children have been shown to conflate physical possibility and moral permissibility in judging the possibility of morally impermissible events. Phillips and Bloom (2017) asked children between the ages of 4 and 7 years to judge the possibility of three kinds of events: improbable events, such as being given 100 identical shirts, impossible events, such as turning a hat into a candy bar, and immoral events, such as taking a ball from another child. They found that children not only denied the possibility of improbable events (as shown in previous research) but also denied the possibility of immoral events. In fact, 4- and 5-year-olds denied the possibility of immoral events as often as they denied the possibility of impossible events. Similarly, when adults are forced to make judgments of possibility under severe time pressure, they also tend to judge immoral events as impossible (Phillips & Cushman, 2017).

These three sets of findings suggest that when considering non-actual events, modal cognition might not clearly differentiate what is physically impossible from what is morally wrong. We attempted to confirm this hypothesis by asking preschool- and elementary-school-aged children to judge *both* the physical possibility and the moral permissibility of various kinds of events: events that violate physical laws (impossible events), events that violate physical regularities but not physical laws (improbable events), events that violate moral laws (immoral events), events that violate social conventions but not moral laws (unconventional events), and events that violate no laws or regularities whatsoever (ordinary events). Our prediction was that young children would not only have difficulty differentiating law-violating events from regularity-violating events within a given domain (impossible vs. improbable and immoral vs. unconventional) but would also have difficulty differentiating events *across* domains, judging impossible and improbable events as morally wrong and judging immoral and unconventional events as impossible.

Study 1

Method

Participants

In total, 80 children and 101 adults participated in Study 1. The adults were recruited from introductory psychology and cognitive science courses at Occidental College and were compensated for their participation with extra credit in those courses. The children were recruited from local parks and were tested on-site. In terms of gender, 71% of the adults were female and 55% of the children were female. Participants were not asked for information about their race, but they were sampled from racially diverse populations. The adults were sampled from a population that was 52% Caucasian, 14% Latino, 14% Asian, 5% African American, and 15% "other," and the children were sampled from a population that was 42% Caucasian, 25% Latino, 10% Asian, 10% African American, and 13% "other."

The children ranged in age from 3.5 to 10.5 years. For the purposes of data analysis, we divided them into a younger group (3.5–5.8 years, $M = 4.7$ years, $n = 47$) and an older group (6.4–10.5 years, $M = 8.2$ years, $n = 33$). We chose 6 years as our dividing line because it roughly splits preschoolers from elementary school children. Some older 5-year-olds may have been in kindergarten at the time of testing, but they would not yet have completed a full year.

Materials

Participants were presented with eight vignettes, each of which consisted of a problem followed by five resolutions to that problem. The resolutions instantiated the five types of events mentioned above: ordinary, impossible, improbable, unconventional, and immoral. For each resolution, participants judged whether it was possible or whether or it was permissible. The vignettes were presented in one of four random orders, and each vignette was illustrated with a photograph depicting the problem at hand. The entire set of vignettes can be found in the Appendix.

As an illustration, one vignette was presented to participants as follows: "This is Melissa. Melissa doesn't want to go to school because she doesn't want to leave her mother. She always misses her mother a lot when she goes to school." Participants were then presented with five potential resolutions to this problem: "Melissa and her mother agree to do something special after school, and that makes Melissa feel happy" (an ordinary event); "Melissa snaps her fingers and suddenly it's Saturday, so she doesn't have to go to school" (an impossible event); "Melissa asks her mother to go to school with her, and her mother agrees and goes to all her classes" (an improbable event); "Melissa decides to wear her pajamas to school because wearing pajamas makes her feel happy" (an unconventional event); and "Melissa lies to her mother and tells her that school is closed today so that she doesn't have to go" (an immoral event). The events were presented in a random order, as they were for all vignettes.

For our improbable events, we chose events that were difficult to conceive but possible nonetheless—that is, *conceptually* improbable events. We focused on conceptual improbability rather than statistical improbability because young children are not confused about the distinction between probability and possibility in general. They recognize, for instance, that it is possible to pull a red marble from a bag containing ten blue marbles and one red marble but that it is impossible to pull a yellow marble from the same bag (Shtulman & Carey, 2007; see also Horobin & Acredolo, 1989; Sophian & Somerville, 1988). The notion of possibility that young children have difficulty grasping is possibly with respect to unlikely conceptual combinations (e.g., the possibility of finding an alligator under the bed).

For our unconventional events, we chose events that are socially anomalous but not overtly harmful. Some exemplified disregard for personal well-being (e.g., a child searching through the trash at a birthday party for leftover cake), some exemplified disregard for social conventions (e.g., a child cleaning her room by shoving her clothes under her bed), and some exemplified disregard for in-group norms (e.g., a child tattling on a classmate to obtain a ball that her classmate has monopolized). Although this collection may seem heterogeneous on its surface, it coheres at a deeper level in that none of our unconventional events instantiated direct, intentionally caused harm, as our immoral events did. This distinction was validated by our adult participants, who consistently judged the unconventional events as less wrong (Study 1) and less punishable (Study 2) than the immoral events.

Procedure

At the beginning of the experiment, participants were introduced to the task of making possibility or permissibility judgments in the context of a training phase. For possibility judgments, participants were told, "What we're going to talk about now is whether certain things can happen in the real world. Some things that happen in stories can also happen in the real world and some things cannot—they're impossible." They were then shown four pictures—a boy playing a trombone, a woman sitting on a cloud, a girl painting a picture, and a boy eating a plate of lightning—and were asked to decide whether each of the depicted events was possible or impossible. Correct answers were reaffirmed by the experimenter, and incorrect answers were corrected. For permissibility judgments, participants were told, "What we're going talk about now is whether certain things are okay for a person to do. Some things that happen in stories are okay for a person to do and some things are not—they're wrong." They were then asked to judge the permissibility of four events: a girl sharing her snack with a friend, a girl pulling another girl's hair, a boy playing a game with his friends, and a boy hitting another boy. Correct answers were reaffirmed, and incorrect answers were corrected. Overall, participants needed very little correction.

After completing the training phase, participants judged the possibility or permissibility of 40 events: five events for each of the eight vignettes. Possibility and permissibility judgments were

elicited in blocks. Half of the participants made possibility judgments before permissibility judgments, and half did the opposite. The vignettes were counterbalanced across blocks so that participants made only one type of judgment per vignette (e.g., half made possibility judgments for the vignette about Melissa not wanting to go to school, and half made permissibility judgments instead).

Both types of judgments were elicited using a two-step procedure. For possibility judgments, participants were first asked whether the event could occur in real life. If they answered "yes," the experimenter moved on to the next event. If they answered "no," the experimenter asked whether the event was "sorta impossible" or "very impossible." For permissibility judgments, participants were first asked whether the event was okay for someone to do in real life. If they answered "yes," the experimenter moved on to the next event. If they answered "no," the experimenter asked whether the event was "sorta wrong" or "very wrong." This protocol was used for both children and adults. That is, even the adults were tested in one-on-one interviews with an experimenter.

Our motivation for using a 3-point scale was to allow participants to make modal distinctions among events that might otherwise be conflated on a 2-point scale. Consider the immoral event of lying to your mother and the unconventional event of wearing pajamas to school. If we had asked participants only whether these events are wrong, many participants would likely have said "yes" to both because wearing pajamas is a violation of the dress code at most schools. By asking a follow-up question about degree of wrongness, participants were able to clarify that wearing pajamas to school is only sorta wrong, whereas lying to your mother is very wrong (if that was, in fact, what they believed).

That said, we were concerned that our youngest participants (preschoolers) might have difficulty in using a 3-point scale. Thus, we administered the scale in two parts so that children only ever had to make a dichotomous decision (e.g., "okay" vs. "wrong" followed by "sorta wrong" vs. "very wrong"). We also verified that children's answers to the first question were in line with previous research on children's possibility judgments and permissibility judgments and that children applied our scale to the control items in an adult-like manner, implying that they understood its meaning.

Scoring

For each possibility judgment, participants received a score of 0 if they claimed the event could happen in real life, a score of 1 if they claimed the event was "sorta impossible," and a score of 2 if they claimed the event was "very impossible." For each permissibility judgment, participants received a score of 0 if they claimed the event was okay to do in real life, a score of 1 if they claimed the event was "sorta wrong," and a score of 2 if they claimed the event was "very wrong." For each participant, these scores were then averaged across the four events within each category within each judgment block (e.g., scores for the four improbable events within the possibility judgment block), resulting in a mean score ranging from 0 to 2 for each type of event and each type of judgment.

Analysis plan

We analyzed participants' scores in three ways. First, we focused on how participants answered the first question in our two tiers of questioning, looking for developmental differences in how often they judged the four types of events as possible and permissible. This analysis was intended to establish that younger participants were less likely than older participants to recognize that violations of mere regularities are still possible and permissible (as found in previous studies).

Second, we investigated participants' ability to differentiate events that violate laws from those that do not *within* each domain. We did this by subtracting their scores for regularity-violating events from their scores for the corresponding law-violating events. For questions about physical possibility, participants' mean scores for judgments of immoral events, improbable events, and unconventional events were subtracted from their mean score for judgments of impossible events. This analysis resulted in three difference scores that indicate whether participants differentiated each kind of event from events that clearly violate physical laws. A similar approach was taken for judgments of permissibility. Participants' mean scores for judgments of impossible events, improbable events, and unconventional events were subtracted from their mean score for judgments of immoral events. This analysis resulted in three difference scores that indicate whether participants differentiated each kind of event from events that clearly violate moral rules. For both sets of difference scores, the higher the score, the more strongly participants distinguished events that violate a domain-specific law from

those that do not. We used this metric to assess developmental differences in participants' ability to make distinctions within the domains of physical possibility and moral permissibility.

Third, we computed difference scores *across* domains. For simplicity, we focused on immoral and impossible events. In each case, we took participants' permissibility scores and subtracted them from their possibility scores for the same kind of event. This resulted in two difference scores for each participant that indicate whether they thought immoral events were more immoral than impossible and, conversely, whether they thought impossible events were more impossible than immoral. On this metric, positive scores indicated stronger judgments that the event is impossible but not impermissible, and negative scores indicated stronger judgments that the event is impermissible but not impossible. We analyzed this metric as a way of assessing developmental differences in participants' capacity to differentiate the modal domains themselves.

Note that between-domain difference scores were computed within participants but that judgment type was counterbalanced across participants, meaning that for any given event participants judged its possibility or its impermissibility but not both. Thus, difference scores reflect participants' ability to differentiate possibility from permissibility in general rather than their ability to do so for particular events. Future research should explore whether children's ability to differentiate possibility from permissibility is improved by considering both questions for the same event.

Results

Control items

We treated participants' mean judgments for the ordinary events as a control. We compared those means with a score of 1.0, which corresponded to "sorta impossible" in the domain of physical possibility and "sorta wrong" in the domain of moral permissibility. In both domains, the mean score for all age groups was significantly less than 1.0 [younger children: mean possibility score = 0.25, $t(46) = 12.39$, $p < .001$, and mean permissibility score = 0.24, $t(46) = 14.45$, $p < .001$; older children: mean possibility score = 0.03, $t(32) = 53.67$, $p < .001$, and mean permissibility score = 0.13, $t(32) = 20.59$, $p < .001$; adults: mean possibility score = 0.01, $t(100) = 285.69$, $p < .001$, and mean permissibility score = 0.02, $t(100) = 121.82$, $p < .001$]. Even the younger children had no difficulty in answering questions about possibility and permissibility for ordinary events; they judged them as possible and permissible significantly more often than they did for the four types of abnormal events [for all comparisons, $t(46) > 4.00$, $p < .001$]. These findings confirm that participants of all ages understood the task and used the scale appropriately. All remaining analyses focus solely on the abnormal events.

Within-domain distinctions

Participants' tendency to judge each type of event as simply possible or permissible is displayed in Table 1. We used one-way analyses of variance (ANOVAs) to assess the effects of age group on each type of judgment for each type of event. Participants' tendency to judge particular events as possible

Table 1
Mean proportion of events judged to be possible (could happen in real life) or permissible (okay to do in real life) for each type of event and each age group in Study 1.

		Younger children	Older children	Adults	*F*
Possible	Impossible	.31	.14	.02	31.38***
	Improbable	.53	.79	.83	18.64***
	Unconventional	.56	.91	.97	67.07***
	Immoral	.53	.89	.96	54.59***
Permissible	Impossible	.48	.52	.64	3.28*
	Improbable	.57	.45	.58	2.19
	Unconventional	.39	.44	.54	5.19**
	Immoral	.16	.08	.02	12.87***

* $p < .05$.
** $p < .01$.
*** $p < .001$.

increased with age for improbable events, $F(2, 178) = 18.64$, $p < .001$, $\eta^2 = .17$, unconventional events, $F(2, 178) = 67.07$, $p < .001$, $\eta^2 = .43$, and immoral events, $F(2, 178) = 54.59$, $p < .001$, $\eta^2 = .38$, but *decreased* with age for impossible events, $F(2, 178) = 31.38$, $p < .001$, $\eta^2 = .26$. Contrast analyses confirmed that these effects varied linearly across age groups [impossible events: $F(1, 178) = 62.44$, $p < .001$, $\eta^2 = .26$; improbable events: $F(1, 178) = 36.28$, $p < .001$, $\eta^2 = .17$; unconventional events: $F(1, 178) = 131.07$, $p < .001$, $\eta^2 = .42$; immoral events: $F(1, 178) = 106.66$, $p < .001$, $\eta^2 = .37$].

Participants' tendency to judge particular events as permissible *increased* with age for impossible events, $F(2, 178) = 3.28$, $p = .040$, $\eta^2 = .04$, and unconventional events, $F(2, 178) = 5.19$, $p = .006$, $\eta^2 = .06$, but *decreased* with age for immoral events. $F(2, 178) = 12.87$, $p < .001$, $\eta^2 = .13$. Once again, contrast analyses confirmed that these effects were linear in nature [impossible events: $F(1, 178) = 5.64$, $p = .019$, $\eta^2 = .03$; unconventional events: $F(1, 178) = 9.40$, $p = .003$, $\eta^2 = .05$; immoral events: $F(1, 178) = 25.55$, $p < .001$, $\eta^2 = .13$]. The only set of judgments for which age effects were not observed were permissibility judgments for the improbable events, $F(2, 178) = 2.19$, $p = .115$, $\eta^2 = .02$; older children and adults judged these events as "sorta wrong" more than expected.

Difference scores between law-violating events and regularity-violating events are displayed in Fig. 1. For both types of judgments—possibility judgments and permissibility judgments—younger children differentiated law-violating events (e.g., lying to a parent) from the other types of events (e.g., wearing pajamas to school, bringing a parent to school) to a lesser degree than older children did, and older children differentiated law-violating events from the other types of events to a lesser degree than adults did. To confirm the reliability of this finding, we submitted each set of difference scores to a 3×3 mixed-measures ANOVA, in which event comparison (e.g., impossible/immoral vs.

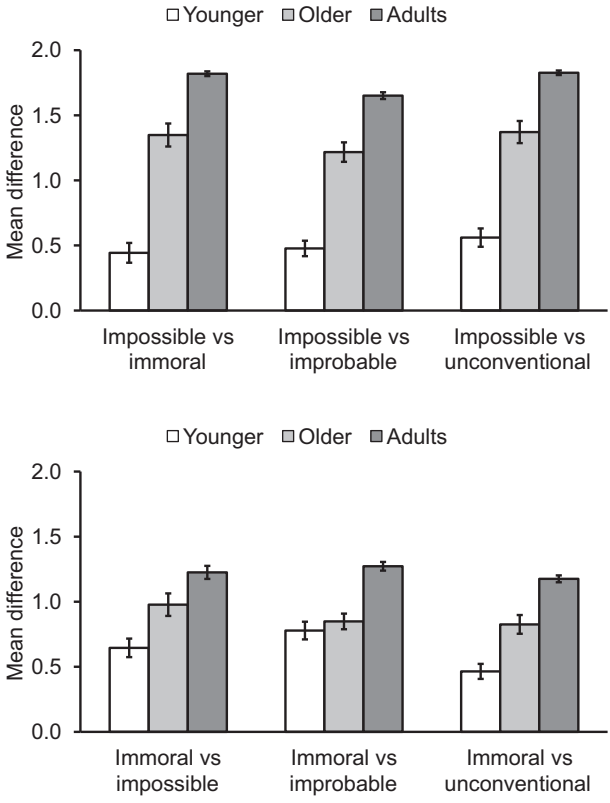

Fig. 1. Within-domain differences, Study 1: Mean difference in possibility scores between the impossible events and all other events by age group (top panel) and mean difference in permissibility scores between the immoral events and all other events by age group (bottom panel). Error bars represent standard errors.

impossible/improbable vs. impossible/unconventional) was entered as a within-participants factor and age group (younger children vs. older children vs. adults) was entered as a between-participants factor.

For possibility judgments, this analysis revealed main effects of event comparison, $F(2, 356) = 10.06$, $p = .001$, $\eta^2 = .01$, and age group, $F(2, 178) = 130.39$, $p < .001$, $\eta^2 = .54$, as well as a marginal interaction between them, $F(4, 356) = 2.35$, $p = .054$, $\eta^2 = .01$. With respect to the events, improbable events were differentiated from impossible events less than immoral events were ($M = 1.27$ vs. $M = 1.38$), $t(360) = -2.33$, $p = .020$, $d = 0.25$. Likewise, improbable events were differentiated from impossible events less than unconventional events were ($M = 1.27$ vs. $M = 1.42$), $t(360) = -1.99$, $p = .047$, $d = 0.21$, but unconventional events were differentiated from impossible events no less than immoral events were, $t(360) = -0.49$, $p = .625$, $d = 0.05$. With respect to age, younger children differentiated impossible events from other events less than older children did ($M = 0.49$ vs. $M = 1.31$), $t(238) = -9.35$, $p < .001$, $d = 1.23$, and older children differentiated impossible events from other events less than adults did ($M = 1.77$), $t(112) = -6.51$, $p < .001$, $d = 1.07$. The marginal interaction occurred because the degree of differentiation among event types was not equivalent across age groups.

The same basic pattern was observed for impermissibility judgments. A mixed-measures ANOVA revealed a main effect of event comparison, $F(2, 356) = 4.33$, $p = .014$, $\eta^2 = .01$, and age group, $F(2, 178) = 30.93$, $p < .001$, $\eta^2 = .17$, but no interaction between them, $F(4, 356) = 1.61$, $p = .17$, $\eta^2 = .01$. With respect to the events, unconventional events were differentiated from immoral events less than improbable events were ($M = 0.93$ vs. $M = 1.07$), $t(360) = -2.34$, $p = .020$, $d = 0.25$, and less than impossible events were ($M = 0.93$ vs. $M = 1.03$), $t(360) = -1.479$, $p = .140$, $d = 0.155$, but improbable events were differentiated from immoral events no less than impossible events were, $t(360) = 0.53$, $p = .594$, $d = 0.06$. With respect to age, younger children differentiated immoral events from other events less than older children did ($M = 0.63$ vs. $M = 0.88$), $t(238) = -3.097$, $p = .002$, $d = 0.41$, and older children differentiated immoral events from other events less than adults did ($M = 0.88$ vs. $M = 1.22$), $t(112) = -5.31$, $p < .001$, $d = 0.62$.

In sum, we found consistent effects of event type and age group for both possibility judgments and permissibility judgments. The effects of event type, however, were substantially smaller than the effects of age group (possibility judgments: $\eta^2 = .01$ vs. $\eta^2 = .54$; permissibility judgments: $\eta^2 = .01$ vs. $\eta^2 = .17$), indicating that participants differentiated regularity-violating events from law-violating events more or less equivalently for different types of regularity-violating events, but their ability to do so increased dramatically with age.

Between-domain distinctions

The above analyses establish that older participants were better than younger participants at differentiating events that violate laws (physical or moral) from those that do not when assessing the events' physical possibility or moral permissibility. But how well did participants differentiate the modal questions themselves—that is, the question of whether something is physically possible and the question of whether something is morally permissible?

To investigate, we assessed the effects of age on participants' between-domain difference scores (described above) using one-way ANOVAs (Fig. 2). These analyses confirmed that between-domain difference scores varied by age groups for both the impossible events, $F(2, 178) = 25.87$, $p < .001$, partial $\eta^2 = .23$, and the immoral events, $F(2, 178) = 57.44$, $p < .001$, partial $\eta^2 = .39$. Contrast analyses confirmed that these scores increased linearly across age groups for the impossible events, $F(1, 178) = 47.97$, $p < .001$, partial $\eta^2 = .21$, but decreased linearly across age groups for the immoral events, $F(1, 178) = 111.81$, $p < .001$, partial $\eta^2 = .39$. In other words, younger participants were less likely than older participants to differentiate the question of whether something is possible from the question of whether something is permissible, both when considering events that violate physical laws but not moral laws (e.g., traveling back in time) and when considering events that violate moral laws but not physical laws (e.g., intentionally tripping someone). This relationship was observed even when adults were excluded from the analysis and children were compared with each other in terms of their age in months. Children's age in months was correlated both with stronger impossibility scores relative to impermissibility scores for the impossible events, $r(80) = .24$, $p < .05$, and with stronger impermissibility scores relative to impossibility scores for the immoral events, $r(80) = -.55$, $p < .001$.

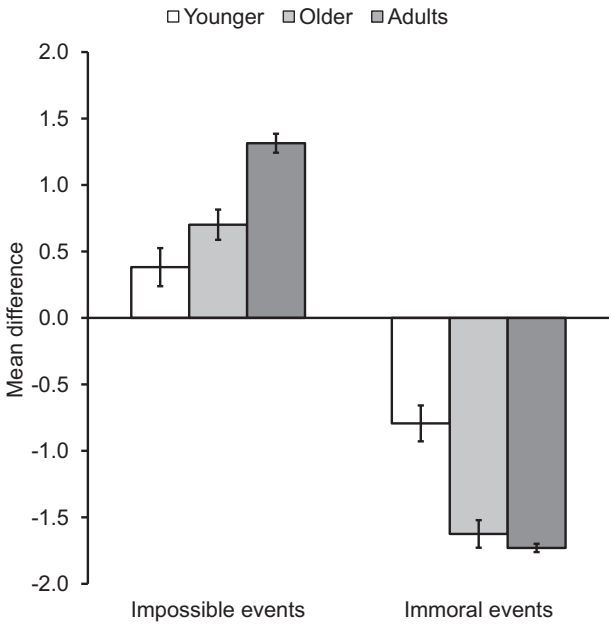

Fig. 2. Between-domain differences, Study 1: Mean difference between possibility scores and permissibility scores for impossible events and immoral events by age group. Positive scores indicate stronger judgments that the event is impossible; negative scores indicate stronger judgments that the event is impermissible. Error bars represent standard errors.

Discussion

For each modal question, adults strongly differentiated events that violate question-relevant laws from those that violate mere regularities. They also differentiated the questions themselves, judging impossible events as impossible but not impermissible and judging immoral events as impermissible but not impossible. Children's differentiation of event types and question types was less pronounced, particularly for younger children. Younger children, as a group, showed evidence of making within-domain distinctions and between-domain distinctions, but the size of those distinctions was small relative to the distinctions made by older participants.

Our findings replicate previous findings demonstrating that preschoolers have difficulty in making modal distinctions both within the domain of physical possibility (e.g., Shtulman & Carey, 2007) and within the domain of moral permissibility (e.g., Tisak & Turiel, 1988). They extend those findings by showing that (a) children's ability to make modal distinctions in the domain of physical possibility develops in tandem with their ability to make modal distinctions in the domain of moral permissibility and that (b) children must learn not only to distinguish different types of events within the same domain but also to distinguish the domains themselves. This latter finding is particularly striking. When younger children were presented with events such as floating in the air or conjuring a ball from nowhere, they claimed not just that those events were impossible but also that they were impermissible. Conversely, when presented with events such as stealing a candy bar or lying to a parent, they claimed not just that those events were impermissible but also that they were impossible. For many young children, the question of whether something could happen in real life and the question of whether something is okay to do merited surprisingly similar responses.

One unexpected finding is that many older children and adults judged impossible events and improbable events as "sorta wrong." These judgments may reflect the same disposition seen in young children but to a lesser degree; however, a more likely explanation is that older participants interpreted our questions in unintended ways. Take, for example, the vignette about Melissa not wanting

to go to school. The improbable event was that "Melissa asks her mother to go to school with her, and her mother agrees and goes to all her classes," and the impossible event was that "Melissa snaps her fingers and suddenly it's Saturday, so she doesn't have to go to school." In both cases, Melissa attempts to sidestep her problem rather than confront it head-on, and participants may have focused on the nature of the solution rather than the principles violated by the solution (see Study 2 below for evidence consistent with this interpretation).

Concerns of this nature could, in fact, apply to the developmental differences we observed as well. Whereas we have argued that the developmental changes observed in Study 1 reflect changes in how children evaluate the modal status of abnormal events, an alternative interpretation is that they reflect changes in how children understood the modal questions themselves. For example, it is possible that young children understood the question of whether something was possible as a question about general plausibility rather than possibility, and they may have understood the question about whether something was wrong as a question about general oddness rather than permissibility. In other worlds, young children's failure to distinguish between physical and social considerations when making modal judgments may have been a product of how they interpreted the particular questions we asked them.

To investigate this possibility, we repeated Study 1 using the same materials and the same age groups but changed our questions. Instead of asking whether something could happen in real life, we asked whether the event would take *magic* to happen. And instead of asking whether something was okay to do, we asked whether the actor who performed that action could be *punished* for it. Our rationale was that the terms "magic" and "punishment" unambiguously apply to physical and social domains, respectively. These terms are not particularly susceptible to alternative interpretations, and young children understand both (see, e.g., Johnson & Harris, 1994, on young children's understanding of magic).

Study 2

Method

Participants

In total, 74 children and 78 adults participated in Study 2. The adults were recruited from Amazon Mechanical Turk and were paid for their participation. The children were recruited from local parks and were tested on-site. In terms of gender, 59% of the adults were female and 70% of the children were female. The children ranged in age from 3.5 to 10.1 years and were once again divided into a younger group (3.5–5.9 years, $M = 4.9$ years, $n = 28$) and an older group (6.0–10.1 years, $M = 7.6$ years, $n = 46$) based on whether they were younger or older than 6 years.

Procedure

The materials and the protocol were identical to those in Study 1 with the exception that the question "could that happen in real life?" was changed to "would that take magic to happen?" and the question "is that okay for him/her to do?" was changed to "could he/she get punished for doing that?" As in Study 1, both questions were followed by a second question if the participants answered "yes." For the magic question participants were asked whether the event required "a little magic" or "a lot of magic," and for the punishment question participants were asked whether the actor deserved "a little punishment" or "a lot of punishment." One other point of difference between Studies 1 and 2 is that the adults in Study 2 completed the survey online rather than in person. The ordering and content of the events, however, were the same.

Scoring and analysis

For magic judgments, participants received a score of 0 if they claimed the event would not require magic, a score of 1 if they claimed it would require a little magic, and a score of 2 if they claimed it would require a lot of magic. For punishment judgments, participants received a score of 0 if they claimed the actor did not deserve punishment, a score of 1 if they claimed the actor deserved a little

punishment, and a score of 2 if they claimed the actor deserved a lot of punishment. For each participant, these scores were then averaged across the four events within each category within each judgment block. As in Study 1, we then analyzed scores in three ways: (a) by computing how often the events were judged as not magical and not punishable, (b) by computing within-domain differences between domain-relevant law violations and the other three kinds of events, and (c) by computing between-domain differences between magic scores and punishment scores for the impossible events and immoral events.

Results

Control items

As in Study 1, we compared judgments for the ordinary events with a score of 1.0, which corresponded to claiming that the event required either "a little magic" or "a little punishment." For both types of judgments, the mean score for all age groups was significantly less than 1.0 [younger children: mean magic score = 0.38, $t(27) = 5.38$, $p < .001$, and mean punishment score = 0.30, $t(27) = 7.09$, $p < .001$; older children: mean magic score = 0.12, $t(45) = 21.13$, $p < .001$, and mean punishment score = 0.21, $t(45) = 14.26$, $p < .001$; adults: mean magic score = 0.02, $t(77) = 186.34$, $p < .001$, and mean punishment score = 0.01, $t(77) = 260.91$, $p < .001$]. These data indicate that participants of all ages understood the task and the scale, claiming that ordinary events require neither magic nor punishment. Once again, our remaining analyses focus solely on judgments for the abnormal events.

Within-domain distinctions

Participants' tendency to judge each type of event as magical or punishable is displayed in Table 2. We used one-way ANOVAs to assess the effects of age group on each type of judgment for each type of event. Participants' tendency to judge that particular events would not require magic (i.e., that they were possible) increased with age for unconventional events, $F(2, 149) = 8.50$, $p < .001$, $\eta^2 = .10$, and immoral events, $F(2, 149) = 12.02$, $p < .001$, $\eta^2 = .14$, but decreased with age for impossible events, $F(2, 149) = 31.18$, $p < .001$, $\eta^2 = .30$. Contrast analyses confirmed that these effects varied linearly across age group [impossible events: $F(1, 149) = 52.65$, $p < .001$, $\eta^2 = .25$; unconventional events: $F(1, 149) = 16.99$, $p < .001$, $\eta^2 = .10$; immoral events: $F(1, 149) = 24.00$, $p < .001$, $\eta^2 = .14$]. No reliable age effects were observed for the improbable events, $F(2, 149) = 2.86$, $p = .06$, $\eta^2 = .04$) because adults claimed these events would require "a little magic" unexpectedly often (25% of the time).

Participants' tendency to judge that particular events did not require punishment (i.e., that they were permissible) increased with age for impossible events, $F(2, 149) = 21.91$, $p < .001$, $\eta^2 = .23$, and improbable events, $F(2, 149) = 8.16$, $p < .001$, $\eta^2 = .10$, but decreased with age for immoral events, $F(2, 149) = 17.03$, $p < .001$, $\eta^2 = .19$. Contrast analyses confirmed that all three effects were linear in nature [impossible events: $F(1, 149) = 25.49$, $p < .001$, $\eta^2 = .13$; improbable events: $F(1, 149) = 10.82$,

Table 2
Mean proportion of events judged as possible (would not require magic) or permissible (would not require punishment) for each type of event and each age group in Study 2.

		Younger children	Older children	Adults	*F*
Possible	Impossible	.43	.27	.01	15.27***
	Improbable	.72	.82	.69	2.86
	Unconventional	.76	.90	.95	8.50***
	Immoral	.75	.91	.96	12.02***
Permissible	Impossible	.65	.67	.95	21.91***
	Improbable	.60	.63	.79	8.16***
	Unconventional	.46	.44	.53	1.94
	Immoral	.29	.14	.03	17.03***

* $p < .05$.
** $p < .01$.
*** $p < .001$.

$p = .001$, $\eta^2 = .07$; immoral events: $F(1, 149) = 32.78$, $p < .001$, $\eta^2 = .18$]. No age effects were observed for the unconventional events, $F(1, 149) = 1.94$, $p = .147$, $\eta^2 = .03$, with younger children claiming that these events were permissible (i.e., not punishable) at higher rates than they did in Study 1.

Difference scores between law-violating events and regularity-violating events are displayed in Fig. 3. From a normative perspective, only the impossible events violate physical laws, and thus only impossible events should be judged as requiring magic. Likewise, only the immoral events violate moral laws, and thus only immoral events should be judged as requiring punishment. We assessed how well participants were able to make these distinctions by submitting participants' difference scores to mixed-measures 3 × 3 ANOVAs with event comparison (e.g., impossible/immoral vs. impossible/improbable vs. impossible/unconventional) entered as the within-participants factor and age group (younger children vs. older children vs. adults) entered as the between-participants factor.

For magic judgments, this analysis revealed main effects of event comparison, $F(2, 300) = 26.85$, $p = .001$, $\eta^2 = .02$, and age group, $F(2, 150) = 84.94$, $p < .001$, $\eta^2 = .51$), as well as a significant interaction between them, $F(4, 300) = 9.22$, $p < .001$, $\eta^2 = .01$. Improbable events were differentiated from impossible events less than immoral events were ($M = 1.19$ vs. $M = 1.41$), $t(304) = -2.72$, $p = .007$, $d = 0.31$, and less than unconventional events were ($M = 1.19$ vs. $M = 1.40$), $t(304) = -2.53$, $p = .012$, $d = 0.29$, but unconventional events were differentiated from impossible events to the same degree as immoral events were, $t(304) = 0.12$, $p = .906$, $d = 0.01$. With respect to age, younger children differentiated impossible events from other events less than older children did ($M = 0.53$ vs. $M = 1.08$), $t(213) = -6.87$, $p < .001$, $d = 0.89$, and older children differentiated impossible events from other events less

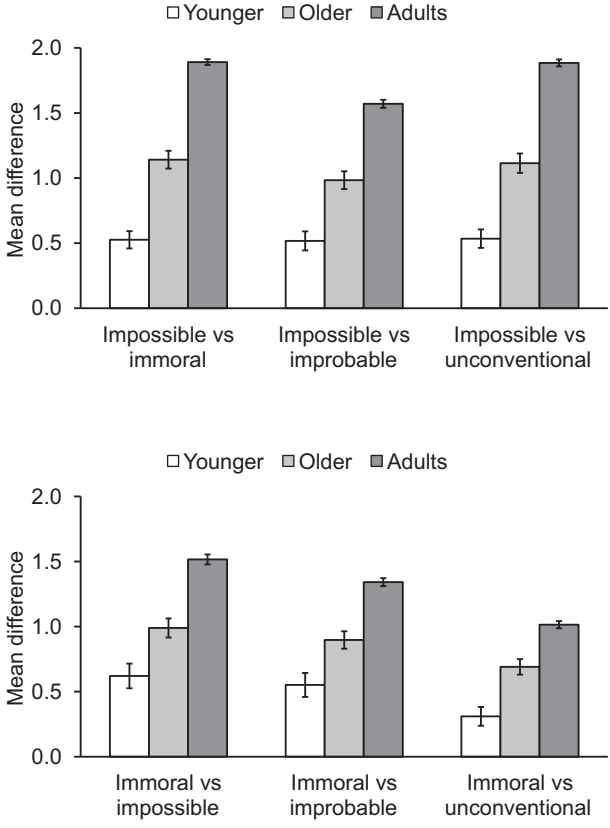

Fig. 3. Within-domain differences, Study 2: Mean difference in magic scores between the impossible events and all other events by age group (top panel) and mean difference in punishment scores between the immoral events and all other events by age group (bottom panel). Error bars represent standard errors.

than adults did ($M = 1.08$ vs. $M = 1.78$), $t(185) = -11.29$, $p < .001$, $d = 1.40$. The interaction between event type and age group occurred because adults differentiated improbable events from impossible events less strongly than they differentiated unconventional and immoral events from impossible events, but this pattern was less pronounced in older children and was absent in younger children.

For punishment judgments, the mixed-measures ANOVA revealed main effects of event comparison, $F(2, 300) = 56.17$, $p < .001$, $\eta^2 = .07$, and age group, $F(2, 150) = 31.93$, $p < .001$, $\eta^2 = .26$. Unconventional events were differentiated from immoral events less than improbable events were ($M = 0.78$ vs. $M = 1.06$), $t(304) = -4.17$, $p < .001$, $d = 0.47$, and less than impossible events were ($M = 0.783$ vs. $M = 1.19$), $t(304) = -5.70$, $p < .001$, $d = 0.65$. With respect to age, younger children differentiated immoral events from other events less than older children did ($M = 0.49$ vs. $M = 0.86$), $t(223) = -4.06$, $p < .001$, $d = 0.56$, and older children differentiated immoral events from other events less than adults did ($M = 0.86$ vs. $M = 1.29$), $t(370) = -7.53$, $p < .001$, $d = 0.81$. The interaction between event type and age group was marginal, $F(4, 300) = 2.283$, $p = .060$, $\eta^2 = .006$, and it occurred for the same reason as the interaction described above, that is, because participants' differentiation of immoral events from other events varied by event for some age groups slightly more than others.

In sum, these findings parallel those of Study 1; we found consistent effects of event type and age group for both types of domain-specific judgments. We also found that the effects of event type were substantially smaller than the effects of age group (possibility judgments: $\eta^2 = .02$ vs. $\eta^2 = .51$; permissibility judgments: $\eta^2 = .07$ vs. $\eta^2 = .26$). These findings further confirm that the ability to differentiate law-violating events from regularity-violating events varies only slightly by event but varies substantially with age.

Between-domain distinctions

For impossible events and immoral events, how well did participants differentiate the extent to which they require magic from the extent to which they require punishment? To answer this question, we subtracted punishment scores from magic scores. The mean difference between these scores is displayed in Fig. 4 as a function of age group.

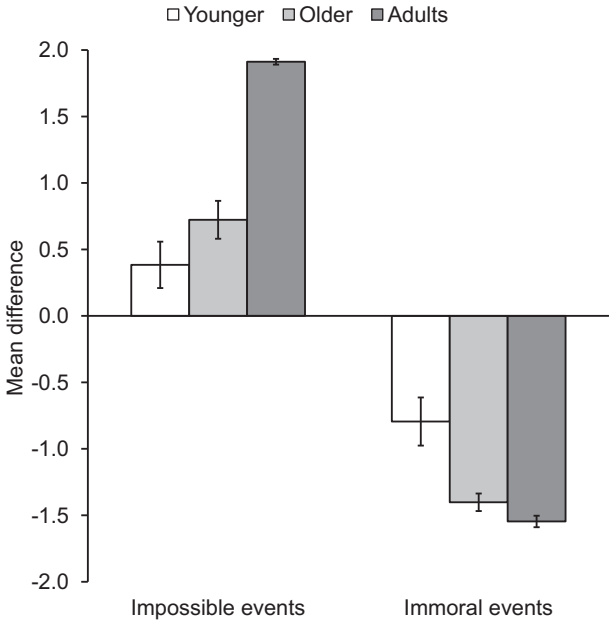

Fig. 4. Between-domain differences, Study 2: Mean difference between magic scores and punishment scores for impossible events and immoral events by age group. Positive scores indicate stronger judgments that the event would require magic; negative scores indicate stronger judgments that the actor deserves punishment. Error bars represent standard errors.

One-way ANOVAs revealed that difference scores varied by age group for both the impossible events, $F(2, 149) = 76.53$, $p < .001$, partial $\eta^2 = .51$, and the immoral events, $F(2, 149) = 19.99$, $p < .001$, partial $\eta^2 = .21$. Contrast analyses revealed that difference scores increased linearly across age groups for the impossible events, $F(1, 149) = 108.44$, $p < .001$, partial $\eta^2 = .42$, but decreased linearly across age groups for the immoral events, $F(1, 149) = 39.66$, $p < .001$, partial $\eta^2 = .21$. That is, younger participants were less likely than older participants to judge that events requiring magic (impossible events) did not also require punishment and to judge that events requiring punishment (immoral events) did not also require magic. This finding was not driven solely by differences between children and adults. Among just the children, age (in months) was correlated with stronger magic scores relative to punishment scores for the impossible events, $r(74) = .27$, $p < .05$, and stronger punishment scores relative to magic scores for the immoral events, $r(74) = -.56$, $p < .001$. Thus, older children viewed questions about magic and questions about punishment as more distinct than younger children did for both immoral and impossible events.

Discussion

The findings of Study 2 closely replicate those of Study 1 despite the change in how the judgments of possibility and permissibility were elicited. The ability to differentiate events that violate domain-relevant laws from other types of abnormal events increased with age, as did the ability to differentiate the modal domains themselves. These findings are particularly noteworthy when considering that children's prior experience with punishment would have pertained exclusively to violations of moral principles and that children's prior experience with the concept of magic would have pertained exclusively to violations of physical principles. Yet, despite such differences in prior experience, most younger children claimed that it would take magic for a person to do something immoral (e.g., lie, steal) and that a person who does something impossible (e.g., float in the air, conjure an object out of nowhere) deserves punishment.

Consistent with the idea that the questions used in Study 2 were more domain-specific than those used in Study 1, we found that older children and adults were less inclined to judge impossible, improbable, and unconventional events as punishable than to judge them as wrong. That is, older children's punishment scores for events that did not violate any moral laws averaged 0.7 in Study 2, whereas their impermissibility scores for those same events averaged 0.9 in Study 1, and adults' punishment scores for events that did not violate any moral laws averaged 0.3 in Study 2, whereas their impermissibility scores for those same events averaged 0.6 in Study 1. These data confirm our suspicion that the unexpectedly high impermissibility ratings among older children and adults in Study 1 may have been an artifact of the question.

General discussion

When making judgments of physical possibility and judgments of moral permissibility, both of which recruit modal cognition, how well do young children differentiate events that violate domain-relevant laws from events that violate other kinds of regularities? And how well do they differentiate the domain of physical possibility from the domain of moral permissibility in general? In two studies, we have shown that preschool-aged children are not adept at either kind of differentiation. They claim that events that are improbable but do not violate any physical laws, such as getting into a movie for free or convincing a sibling to clean one's room, are impossible and would require magic to occur. Likewise, they claim that behaviors that are unconventional but do not violate any moral rules, such as eating a bug or wearing pajamas to school, are wrong and should be punished if they occur. Finally, when preschoolers regard an event as impossible, they also often claim that it is impermissible, and when they regard an event as impermissible, they also often claim that it is impossible.

This latter finding is particularly telling of the way in which young children conceive of events that are precluded by physical or social principles. They recognize that these events will not occur or should not occur, but they seem not to identify the specific principle that precludes them from occur-

ring. Without identifying that principle, an event that is precluded for physical reasons (e.g., floating in the air) is also viewed as morally precluded, and an event that is precluded for moral reasons (e.g., lying to a parent) is also viewed as physically precluded. Adults draw a sharp distinction between these two senses of preclusion, but young children apparently do not.

It is important to keep in mind that conflating questions of physical possibility and moral permissibility is not the same as conflating physical causation and psychological causation. Children as young as 3 recognize that some actions are mediated by physical causes (e.g., lifting an object depends on how heavy the object is) and others are mediated by psychological causes (e.g., finding a bowl in a kitchen depends on knowing where the bowls are kept) (see Kalish, 1998; Schult & Wellman, 1997; Sobel, 2004). Thus, when children in the current study claimed that floating in the air is wrong or that lying to a parent is impossible, it is unlikely that they were confused about the mechanisms that might bring about these events. Rather, they appear to be deficient in their ability to represent an event as meeting one set of modal criteria but not another set. Children's apparent default for thinking about non-actual events is to represent them either as something that can happen or as something that cannot happen. Identifying the particular reason why an event cannot happen appears to require additional cognitive capacity or effort. Children may know implicitly that physically abnormal events and morally abnormal events violate different kinds of regularities, but they do not appear to consult that knowledge when making judgments that involve non-actual possibilities.

One of the more striking aspects of what we have uncovered is that young children judge immoral events as impossible even in cases where they are likely to have committed these acts themselves (e.g., lying to a parent, taking something from a classmate). This aspect of our findings cuts sharply against what would be predicted by constructivist theories of moral development. We suspect that one likely explanation of this pattern is that children are not conceiving of *their own* actions as moral or immoral (or at least not to the same extent that they are moralizing others' actions). This hypothesis is supported by research showing that young children often believe that moral norms apply to others' behavior more than to their own behavior, predicting that others will act in accordance with moral norms more than they themselves will (and more than they in fact do) (see Chernyak & Kushnir, 2014; Smith, Blake, & Harris, 2013). Future work should directly explore this issue by asking children about the possibility of their own (vs. others') immoral actions.

At the same time, the fact that children *do know* something about physical regularities and social norms potentially explains why their conflation of modal distinctions is relative rather than absolute. In neither study did we observe a complete lack of differentiation between target events or target questions; all mean difference scores were greater than zero even for our youngest participants. Although children's ability to make modal distinctions increases with age, they exhibit some initial competence. What might account for that competence?

One possibility is that our youngest participants had already begun developing the capacity to make within- and between-domain distinctions in the same way as older children and adults do—that is, by tagging the event as a violation of some particular law or regularity—but they were not practiced at doing so. Thus, younger children may have successfully identified the nature of the violation on some occasions but defaulted to a conflated sense of modality on most other occasions.

Another possibility is that young children were aided by the content of the events themselves. Within a domain, the distinction between law-violating events and regularity-violating events is implicitly conveyed by the egregiousness of the violation. For example, improbable events must strike young children as abnormal, but impossible events must strike them as even more abnormal. Young children may be sensitive to this degree of abnormality and use it as a way of assessing the extent to which an event violates domain-relevant regularities. Between-domain distinctions are also implicitly conveyed by the content of the events themselves. Impossible events are likely to activate knowledge of physical regularities more so than social regularities, and immoral events are likely to activate knowledge of social regularities more so than physical regularities. Such differences in knowledge activation may have effectively focused children's attention on one set of modal considerations (e.g., those relevant to physical possibility) rather than another set of modal considerations (e.g., those relevant to moral permissibility).

In line with this view, many of the physical violations we used as stimuli were sufficiently separated from the social domain that even young children may have thought that social considerations

did not apply. A ball instantaneously appearing on a person's head, for instance, has little to do with mental states or interpersonal relations, and children may have viewed such an event as simply beyond the purview of moral rules (as it is). That said, the robust developmental differences observed between our younger participants and our older participants suggests that the ability to make clear and consistent modal distinctions requires explicit identification of the principle being violated and its status with respect to the modal question at hand.

On a broader note, our findings have implications not just for the development of modal cognition but also for the architecture of mature modal cognition. As noted in the Introduction, there are deep similarities between the language used to express physical possibility and the language used to express moral permissibility (Kratzer, 2012; Portner, 2009). There are also deep similarities between adults' reasoning about physical possibility and their reasoning about moral permissibility (Bear, 2016; Shtulman & Tong, 2013). Our findings suggest that those similarities are not coincidental. Modal distinctions—either within a domain or between domains—appear to be secondary to the more general (and simpler) distinctions between events that are precluded by prior knowledge and those that are not. Put differently, people of all ages may focus on events that are *normal* in a given situation and ignore the events that are not, including impossible events, improbable events, unconventional events, and immoral events. Adults are able to distinguish among abnormal events if pressed to do so, but those distinctions may be less fundamental to everyday modal reasoning than simply distinguishing what is normal from what is abnormal. Indeed, Phillips and Cushman (2017) found that even adults judge immoral events as impossible when they are forced to make their judgments quickly and do not have enough time to reflect on the reasons why a particular event is or is not precluded.

This emerging picture of modal cognition aligns closely with the capacity that has been theoretically posited by researchers working on high-level judgments. For instance, researchers working on causation have suggested that causal judgments rely on some representation of the alternative events that could have occurred but did not occur (Lewis, 1973; Pearl, 2000; Woodward, 2006). Critically, however, people do not treat all alternative events equally when making causal judgments; they tend to represent alternative possibilities only when those possibilities are not immoral, improbable, unconventional, or impossible (Gerstenberg & Tenenbaum, 2017; Halpern & Hitchcock, 2015; Icard, Kominsky, & Knobe, 2017; Kominsky, Phillips, Gerstenberg, Lagnado, & Knobe, 2015; Phillips & Cushman, 2017). A similar capacity has also been independently posited by researchers working on moral judgments, judgments of freedom, and modality in natural languages (Stalnaker, 2002; Kratzer, 2012; Phillips, Luguri, & Knobe, 2015). Accordingly, the form of modal cognition that adults have been posited to rely on may be deeply related to the modal reasoning we have observed in young children (for a theoretical argument along these lines, see Phillips & Knobe, in press). Future empirical research should continue to explore the centrality of modal cognition to other forms of higher-order cognition.

Conclusion

The ability to differentiate what *could* occur from what *should* occur develops gradually over the first decade of life. It does not come for free with an intuitive theory of physical laws or social rules. To be able to differentiate the events that cannot happen physically from those that should not happen socially, children must learn to tag events not only as unusual or unlikely but also as precluded by specific kinds of principles that impose specific kinds of constraints. Piaget (1948/1997) once speculated that "[social] rules are naturally placed by the child on the same plane as actual physical phenomena. One must go to bed at night, have a bath before going to bed, etc., exactly as the sun shines by day and the moon by night, or as pebbles sink while boats remain afloat. All these things are and must be so; they are as the World-Order decrees that they should be, and there must be a reason for it all" (p. 191). Our findings support Piaget's speculation; the youngest children in our study treated violations of physical rules similarly to how they treated violations of moral rules and vice versa. Distinguishing one type of violation from the other is not an inherent feature of modal cognition; it is a developmental achievement.

Acknowledgments

This research was supported by a James S. McDonnell Understanding Human Cognition Scholar Award to Andrew Shtulman and by Grant N00014-14-1-0800 from the Office of Naval Research to Jonathan Phillips. We thank Gillian Binnie, Tangier Davis, Julia Hamilton, Harper Hayes, Sarah Hennessy, Jennifer Hichar, Nicholas Hung, Jessica Ingle, Allison Powers, Thomas Selstad, Ilana Share, Rosie Silber-Marker, Debra Skinner, and Lea Theodorou for their help with data collection and data analysis.

Vignettes used in the studies

This is Henry. Henry is standing in line at a grocery store when he sees some candy he really wants. He asks his mother if she will buy him the candy, but his mother says no. So ...

- Henry waits to get home to eat his favorite snack. [ordinary]
- Henry throws his hat in the air and it turns into the candy he wants. [impossible]
- Henry tells the man at the store that he should get all the candy in the store for free, and the man gives it to him. [improbable]
- Henry buys the candy with his own money and then waits until he is alone in his room to eat it. [unconventional]
- Henry takes the candy without paying and doesn't tell his mother. [immoral]

This is Annie. This is Lisa. Lisa is playing with her ball, and Annie wants to play with the ball too. But Lisa says that she wants to play with the ball by herself. So ...

- Annie asks Lisa if she wants to play on the swings with her, and Lisa agrees. [ordinary]
- Annie claps her hands loudly and a new ball appears on top of her head. [impossible]
- Annie tells Lisa that balls aren't fun to play with, so Lisa gives the ball to Annie forever. [improbable]
- Annie tells the teacher that Lisa won't share, and the teacher makes Lisa give Annie the ball. [unconventional]
- Annie gets mad and takes Lisa's ball away when Lisa is playing with it. [immoral]

This is Melissa. Melissa doesn't want to go to school because she doesn't want to leave her mother. She always misses her mother a lot when she goes to school. So ...

- Melissa and her mother agree to do something special after school, and that makes Melissa feel happy. [ordinary]
- Melissa snaps her fingers and suddenly it's Saturday, so she doesn't have to go to school. [impossible]
- Melissa asks her mother to go to school with her, and her mother agrees and goes to all her classes. [improbable]
- Melissa decides to wear her pajamas to school because wearing pajamas makes her feel happy. [unconventional]
- Melissa lies to her mother and tells her that school is closed today so that she doesn't have to go. [immoral]

This is Tom. This is Elliot. Tom and Elliot both want to play on their school's basketball team, but there's room for only one of them. Elliot is better at basketball, so Tom thinks that Elliot will be picked for the team. So ...

- Tom decides to try out for the soccer team instead of the basketball team. [ordinary]
- Tom jumps into the air and floats above the ground so he becomes taller and better at basketball. [impossible]

- Tom tells Elliot that doing chores is more fun than basketball, so Elliot decides not to try out for the team. [improbable]
- Tom tells the basketball coach about something bad that Elliot did, and the coach won't let Elliot try out for the team. [unconventional]
- Tom trips Elliot so that Elliot falls and hurts his knee and can't play basketball anymore. [immoral]

This is Jimmy. Jimmy sees a boy standing in line at a movie theater with a ticket in his back pocket. Jimmy also wants to see a movie at the movie theater, but he doesn't have a ticket. So ...

- Jimmy goes back home and watches a movie on TV instead. [ordinary]
- Jimmy finds a tree with movie tickets growing on it and picks one. [impossible]
- Jimmy asks the ticket person if he can see the movie for free, and she lets him. [improbable]
- Jimmy tells the boy how the movie ends and the boy decides he doesn't want to see the movie anymore and gives Jimmy his ticket. [unconventional]
- Jimmy steals the boy's ticket when he is buying popcorn. [immoral]

This is Lauren. Lauren wants to play with her friends, but her parents say she has to clean her room first. So ...

- Lauren tells her friends that she needs to clean her room, and her friends help her so they can play together sooner. [ordinary]
- Lauren turns the lights off in her room, and when she turns them back on her room is instantly clean. [impossible]
- Lauren pays her sister a penny to clean her room for her, and her sister agrees. [improbable]
- Lauren cleans her room by shoving all her clothes and toys under the bed and then goes to play with her friends. [unconventional]
- Lauren lies to her mother and says she cleaned her room and goes to play with her friends without cleaning up. [immoral]

This is Ralph. This is John. Ralph is at John's birthday party. When it's time for cake, Ralph finishes his slice and wants to eat more, but the rest of the cake is gone. So ...

- Ralph eats a slice of watermelon instead. [ordinary]
- Ralph waves his hands over his plate and a new slice of cake instantly appears there. [impossible]
- Ralph tells John's mother that he is still hungry, and she brings him a cake of his very own. [improbable]
- Ralph waits until everyone clears their plates and then searches the trash for any leftover cake. [unconventional]
- When John is not looking, Ralph steals John's cake off his plate and shoves it in his mouth. [immoral]

This is Sophie. Sophie is making a loaf of bread to sell to her neighbor. Just as she's about to finish the dough, a big hairy spider falls in. So ...

- Sophie throws the dough away and makes new dough from scratch. [ordinary]
- Sophie uses a time machine to go back in time and catch the spider before it falls into the dough. [impossible]
- Sophie tells her neighbor that a spider fell into the dough, and her neighbor buys the bread anyway. [improbable]
- Sophie bakes the dough anyway and eats the bread herself. [unconventional]
- Sophie bakes the dough anyway and sells the loaf of bread to her neighbor without telling him about the spider. [immoral]

References

Baillargeon, R. (2004). Infants' physical world. *Current Directions in Psychological Science, 13*, 89–94.

Barnes, J., & Black, J. (2016). Impossible or improbable: The difficulty of imagining morally deviant worlds. *Imagination, Cognition, and Personality, 36*, 27–40.

Bear, A. & Knobe, J. (2016). Normality: Part descriptive, part prescriptive. Cognition. Advance online publication. doi:10.1016/j.cognition.2016.10.024.

Browne, C. A., & Woolley, J. D. (2004). Preschoolers' magical explanations for violations of physical, social, and mental laws. *Journal of Cognition and Development, 5*, 239–260.

Byrnes, J. P., & Duff, M. A. (1989). Young children's comprehension of modal expressions. *Cognitive Development, 4*, 369–387.

Chernyak, N., & Kushnir, T. (2014). The self as a moral agent: Preschoolers behave morally but believe in the freedom to do otherwise. *Journal of Cognition and Development, 15*, 453–464.

Chernyak, N., Kushnir, T., Sullivan, K. M., & Wang, Q. (2013). A comparison of American and Nepalese children's concepts of freedom of choice and social constraint. *Cognitive Science, 37*, 1343–1355.

Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition, 127*, 6–21.

Dahl, A., & Kim, L. (2014). Why is it bad to make a mess? Preschoolers' conceptions of pragmatic norms. *Cognitive Development, 32*, 12–22.

Gerstenberg, T., & Tenenbaum, J. B. (2017). Intuitive theories. In M. R. Waldmann (Ed.), *Oxford handbook of causal reasoning*. Oxford, UK: Oxford University Press.

Haidt, J., Koller, S., & Dias, M. G. (1993). Affect, culture, and morality, or Is it wrong to eat your dog? *Journal of Personality and Social Psychology, 65*, 613–628.

Halpern, J. Y., & Hitchcock, C. (2015). Graded causation and defaults. *British Journal for the Philosophy of Science, 66*, 413–457.

Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development, 26*, 30–39.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450*, 557–559.

Horobin, K., & Acredolo, C. (1989). The impact of probability judgments on reasoning about multiple possibilities. *Child Development, 60*, 183–200.

Icard, T. F., Kominsky, J. F., & Knobe, J. (2017). Normality and actual causal strength. *Cognition, 161*, 80–93.

Johnson, C. N., & Harris, P. L. (1994). Magic: Special but not excluded. *British Journal of Developmental Psychology, 12*, 35–51.

Johnson-Laird, P. N. (1978). The meaning of modality. *Cognitive Science, 2*, 17–26.

Kalish, C. (1998). Reasons and causes: Children's understanding of conformity to social rules and physical laws. *Child Development, 69*, 706–720.

Komatsu, L. K., & Galotti, K. M. (1986). Children's reasoning about social, physical, and logical regularities. *Child Development, 57*, 413–420.

Kominsky, J. F., Phillips, J., Gerstenberg, T., Lagnado, D., & Knobe, J. (2015). Causal superseding. *Cognition, 137*, 196–209.

Kratzer, A. (2012). *Modals and conditionals: New and revised perspectives*. Oxford, UK: Oxford University Press (Oxford Studies in Theoretical Linguistics, Vol. 36).

Lane, J. D., Ronfard, S., Francioli, S. P., & Harris, P. L. (2016). Children's imagination and belief: Prone to flights of fancy or grounded in reality? *Cognition, 152*, 127–140.

Levy, G. D., Taylor, M. G., & Gelman, S. A. (1995). Traditional and evaluative aspects of flexibility in gender roles, social conventions, moral rules, and physical laws. *Child Development, 66*, 515–531.

Lewis, D. (1973). Causation. *Journal of Philosophy, 70*, 556–567.

Lockhart, K. L., Abrahams, B., & Osherson, D. N. (1977). Children's understanding of uniformity in the environment. *Child Development, 48*, 1521–1531.

Matthewson, L. (2016). Modality. In M. Aloni & P. Dekker (Eds.), *Cambridge handbook of formal semantics* (pp. 525–559). Cambridge, UK: Cambridge University Press.

Nauze, F. (2008). *Modality in typological perspective*. Doctoral dissertation, Institute for Logic, Language, and Computation, Amsterdam. Retrieved from https://www.illc.uva.nl/Research/Publications/Dissertations/DS-2008-08.text.pdf.

Nolan-Reyes, C., Callanan, M. A., & Haigh, K. A. (2016). Practicing possibilities: Parents' explanations of unusual events and children's possibility thinking. *Journal of Cognition and Development, 17*, 378–395.

Nucci, L. P. (1981). The development of personal concepts: A domain distinct from moral or societal concepts. *Child Development, 52*, 114–121.

Ozturk, O., & Papafragou, A. (2015). The acquisition of epistemic modality: From semantic meaning to pragmatic interpretation. *Language Learning and Development, 11*, 191–214.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. New York: Cambridge University Press.

Phillips, J., & Bloom, P. (2017). Do children believe immoral events are magical? Manuscript under review.

Phillips, J., & Cushman, F. (2017). Morality constrains the default representation of what is possible. *Proceedings of the National Academy of Sciences of the United States of America, 114*, 4649–4654.

Phillips, J., & Knobe, J. (in press). The psychological representation of modality. Mind & Language.

Phillips, J., Luguri, J. B., & Knobe, J. (2015). Unifying morality's influence on non-moral judgments: The relevance of alternative possibilities. *Cognition, 145*, 30–42.

Piaget, J. (1997). *The moral judgment of the child*. New York: Free Press (Original work published 1948).

Portner, P. (2009). *Modality*. Oxford, UK: Oxford University Press.

Rosengren, K. S., Kalish, C. W., Hickling, A. K., & Gelman, S. A. (1994). Exploring the relation between preschool children's magical beliefs and causal thinking. *British Journal of Developmental Psychology, 12*, 69–82.

Samuels, A., & Taylor, M. (1994). Children's ability to distinguish fantasy events from real life events. *British Journal of Developmental Psychology, 12*, 417–427.

Schmidt, M. F., Rakoczy, H., & Tomasello, M. (2011). Young children attribute normativity to novel actions without pedagogy or normative language. *Developmental Science, 14*, 530–539.

Schmidt, M. F., & Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *PLoS ONE, 6*(10), e23223.
Schult, C. A., & Wellman, H. M. (1997). Explaining human movements and actions: Children's understanding of the limits of psychological explanation. *Cognition, 62*, 291–324.
Shtulman, A. (2009). The development of possibility judgment within and across domains. *Cognitive Development, 24*, 293–309.
Shtulman, A., & Carey, S. (2007). Improbable or impossible? How children reason about the possibility of extraordinary events. *Child Development, 78*, 1015–1032.
Shtulman, A., & Tong, L. (2013). Cognitive parallels between moral judgment and modal judgment. *Psychonomic Bulletin & Review, 20*, 1327–1335.
Shtulman, A., & Yoo, R. I. (2015). Children's understanding of physical possibility constrains their belief in Santa Claus. *Cognitive Development, 34*, 51–62.
Smetana, J. G. (1981). Preschool children's conceptions of moral and social rules. *Child Development, 52*, 1333–1336.
Smetana, J. G. (1989). Toddlers' social interactions in the context of moral and conventional transgressions in the home. *Developmental Psychology, 25*, 499–508.
Smith, C. E., Blake, P. R., & Harris, P. L. (2013). I should but I won't: Why young children endorse norms of fair sharing but do not follow them. *PLoS ONE, 8*(3), e59510.
Sobel, D. (2004). Exploring the coherence of young children's explanatory abilities: Evidence from generating counterfactuals. *British Journal of Developmental Psychology, 22*, 37–58.
Sophian, C., & Somerville, S. C. (1988). Early developments in logical reasoning: Considering alternative possibilities. *Cognitive Development, 3*, 183–222.
Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science, 10*, 89–96.
Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy, 25*, 701–721.
Tisak, M. S., & Turiel, E. (1988). Variation in seriousness of transgressions and children's moral and conventional concepts. *Developmental Psychology, 24*, 352–357.
Turiel, E. (2008). Thought about actions in social domains: Morality, social conventions, and social interactions. *Cognitive Development, 23*, 136–154.
Tworek, C. M., & Cimpian, A. (2016). Why do people tend to infer "ought" from "is"? The role of biases in explanation. *Psychological Science, 27*, 1109–1122.
Weisberg, D. S., & Sobel, D. M. (2012). Young children discriminate improbable from impossible events in fiction. *Cognitive Development, 27*, 90–98.
Woodward, J. (2006). Sensitive and insensitive causation. *Philosophical Review, 115*, 1–50.
Woolley, J. D., & Ghossainy, M. E. (2013). Revisiting the fantasy–reality distinction: Children as naïve skeptics. *Child Development, 84*, 1496–1510.
Young, L., & Phillips, J. (2011). The paradox of moral focus. *Cognition, 119*, 166–178.
Zalla, T., Barlassina, L., Buon, M., & Leboyer, M. (2011). Moral judgment in adults with autism spectrum disorders. *Cognition, 121*, 115–126.