Do children believe immoral events are possible?

Jonathan Phillips

Harvard University, Department of Psychology

Paul Bloom

Yale University, Department of Psychology

Word count: 8,992

Abstract

Previous research found that young children judge improbable events (e.g., counting all the hairs on a dog's tail) to be impossible and to require magic. One possibility is that this effect occurs because young children are unable to simulate how they could occur. We argue here for an alternative, which is that these children have an *undifferentiated* representation of possibility, one that doesn't distinguish between events that are statistically improbable, physically impossible, or morally prohibited. In two studies, we ask children about the possibility of immoral but otherwise ordinary events (e.g., taking a toy from another child). If children are relying on a capacity for simulation, they should judge them to be possible, but if children have an undifferentiated representation of possibility, they should not. We find that young children judge immoral events to be impossible, much like violations of physics (Study 1) and judge that immoral events would require magic to occur (Study 2). Both studies also find that children's representation of possibility becomes more differentiated over the course of development.

Keywords: possibility, morality, probability, intuitive physics, modality

Do children believe immoral events are possible?

There are many things that you will never do. You will never turn back time and you will

never win the lottery. While both of these will never happen, we think of them as being

categorically different: not turning back time is a matter of strict impossibility, but not winning

the lottery is a matter of the odds simply being stacked against you.

For adults, this is clear enough. However, there is some intriguing evidence that young

children fail to make this kind of distinction. Previous research suggests that 4- to 5-year-old

children don't clearly distinguish between physically impossible and statistically improbable

events—they regard both kinds of events to be impossible (Shtulman & Carey, 2007). In fact,

they think that the occurrence of statistically improbable events, like physically impossible

events, would require magic (Shtulman, 2009).

There are at least two theories of this phenomenon. According to the first, young children

judge improbable events to be impossible because they have difficulty simulating how these

events could actually occur (Brown & Woolley, 2004; Shtulman & Carey, 2007). According to

the second, young children have a different understanding of possibility than adults: their

representation of possibility does not differentiate between different kinds of prohibited events,

but treats all of them as belonging to a single category.

We test these two theories by asking young children and adults whether immoral events,

like lying to one's parents, are 'impossible' and whether their occurrence would require 'magic'.

These immoral events are easy to imagine, and thus simulation-based accounts should predict

that children will judge them to be possible and not require magic. In contrast, the

undifferentiated representation account predicts that children will not differentiate between

violations of physical constraints and moral constraints, just as they do not differentiate between

violations of statistical and physical constraints. Thus, the undifferentiated account predicts that

young children will judge moral violations to be impossible and to require magic.

We will begin by reviewing the previous research on children's understanding of

possibility and explain in more detail how both theories can account for the previous work on

children's judgments of possibility. We will then consider the predictions of these two theories

for immoral events, and present our tests of these two accounts.

Previous research on children's judgments of possibility

There is considerable evidence that young children are able to distinguish ordinary events

from those that are impossible (Komatsu & Galotti, 1986; Samuels & Taylor, 1994; Sobel, 2004;

Schult & Wellman, 1997). This capacity has been demonstrated not only when children were

asked about events that violate simple physical laws (Johnson & Harris, 1994), but also when

asked about events that occur in more complex causal domains such as biology (e.g, growing

younger, Rosengren, Kalish, Hickling, & Gelman, 1994) and psychology (e.g., telepathy,

Browne & Woolley, 2004).

What about improbable events? Shtulman and Carey (2007) explored 4- to 8-year-old

children's ability to distinguish between the possibility of different kinds of events. In addition to

asking children about the possibility of actions that violated physical laws (e.g., eating lightning

for dinner) and actions that were completely ordinary (e.g., wearing a baseball cap), they asked

children about the possibility of actions that were highly unlikely to occur but did not violate any

physical law (e.g., finding an alligator under one's bed). Younger children judged that both the physically impossible and the statistically unlikely events were impossible. Older children, by contrast, distinguished between the possibility of improbable and physically impossible events. Subsequent research (Shtulman, 2009) employed an alternative method (developed by Browne & Woolley, 2004) and demonstrated that young children also judge that statistically unlikely events would require 'magic' to occur. This tendency also decreases as children develop.

This research raises two related questions about the development of reasoning about possibility. The first is why young children believe that statistically unlikely events are impossible and require magic. The second concerns what changes over the course of development. There are at least two theoretical approaches to answering these questions.

Simulation-based theories. One possibility is that children may determine whether an event is possible by trying to simulate or imagine the circumstances (and causal background) that would allow that event to occur (Browne & Woolley, 2004; Nolan-Reyes, Callanan, & Haigh, 2016). According to this proposal, young children judge that highly unlikely events are impossible because they have difficulty simulating how they could actually occur (in much the same way that they would have difficulty simulating how events that violate physics could actually occur). Naturally, this approach suggests that what changes over the course of development is children's capacity for simulating the circumstances that would allow for the improbable event to occur.

Empirical research has provided mixed support for this account. On the one hand, a number of studies have shown that children find improbable events to be difficult to imagine (Lane, Ronfard, Francioli & Harris, 2016, Nolan-Reyes, Callanan, & Haigh, 2016). Moreover, in a study of children and parent's conversations about the possibility of improbable events, Nolan-

Reyes et al. (2016) found that children's judgments of the possibility of events was predicted by whether or not their parents discussed the mechanisms that would allow the event to occur (or prevent it from occurring) On the other hand, evidence against this account comes from Lane and colleagues, who first asked children to try to imagine various events occurring and then subsequently asked them if it was possible for the event to occur in real life (Lane, Ronfard, Francioli & Harris, 2016). They found that young children continued to judge that improbable events were impossible even when they themselves had succeeded in imagining how that event could occur, contrary to the predictions of the simulation theory.

   Undifferentiated representation theories. An alternative approach is to argue that children's representation of possibility differs in an important way from that of adults. In particular, children's representation of possibility may not clearly distinguish between violations of different kinds of constraints (physical, statistical), instead lumping all of these kinds of violations together as separate from what is possible (see Phillips & Knobe, *forthcoming,* for a theoretical proposal of precisely this kind of representation of *possibility*).

   A natural way to think about this proposal is by analogy to the adult representation of what is *abnormal*. Representations of abnormality are sensitive to both statistical norm violations and violations of physical laws. That is, events that are extremely improbable (e.g., finding an alligator under your bed) are abnormal in much the same way that events that violate physical laws are (e.g., walking through a wall). It is not as though adults have two senses of abnormality – one for statistical norm violations and one for violations of physical laws. Rather, they have a single sense of abnormality that treats both kinds of violations as abnormal and distinguishes both of them from events that are completely ordinary (Bear & Knobe, *forthcoming*). It may be

that the representation of *possibility* that young children possess functions in much the same way.

In other words, the suggestion is that young children represent events as possible to the extent that they are seen as falling within the intersection of the set of events that are consistent with physics and the set of events that do not have an extraordinarily low probability. In short, this account suggests that children's representation of possibility is *undifferentiated* (Phillips & Knobe, *forthcoming*).

On this approach, the change that occurs across the course of development is that children develop an understanding of 'possibility' that is increasingly focused on the event's physical properties. Such a representation would then be *differentiated*: it would allow for events to be represented as possible in virtue of their consistency with physical, but not statistical, constraints.

Immoral events as a test of these accounts

While both of these accounts can capture the pattern of young children's judgments of improbable events, they make conflicting predictions about how children will understand the possibility of immoral events.

On the one hand, young children are likely to have participated in and witnessed many immoral but otherwise ordinary events (e.g., being cruel to a friend or taking something that belongs to someone else). Thus, unlike physically impossible or statistically improbable events, it should be comparatively easy for children to simulate or imagine conditions under which these immoral events could occur. Consequently, simulation-based theories ought to predict that ordinary moral infractions should not typically be represented as impossible.

At the same time, judging that an immoral event is 'possible' requires the capacity to assess an event's possibility in a way that is differentiated—to distinguish violations of moral constraints from other violations of other kinds of constraints (e.g., physical or statistical norms). Accordingly, if young children's representation of possibility is undifferentiated, they should be inclined to judge that immoral events are impossible, much like they judge statistically improbable events to be impossible. It is only once children's understanding of 'possibility' begins to differentiate between different constraints, that they should begin to regard both immoral and improbable events as possible in virtue of their physical properties.

Previous research on children's understanding of immoral events. Several studies provide suggestive evidence supporting the idea that young children may in fact represent immoral events as impossible. Three- to 5-year-old children, for example, tend to say that others *cannot* do immoral or socially unacceptable actions (Chernyak, Kushnir, Sullivan, & Wang, 2013; Chernyak, Kushnir, & Wellman, 2010; Kalish, 1998; Levy, Taylor & Gelman, 1995). In one of these studies, for instance, Chernyak and colleagues told children about various actions that other children wanted to take (e.g., 'Pat wants to say something that will make his friend cry'), and then asked them to make a judgment about whether or not those children could do that thing (e.g., 'Can Pat say something that will make his friend cry?'). When these actions were morally wrong, young children in both the United States and Nepal answered negatively, judging that others 'cannot' do these actions (Chernyak et al., 2013).

While intriguing, an alternative way to account for these findings is to argue that they merely reflect a confusion about the language being used. All of this previous research explored children's intuitions by asking them questions involving the English modal auxiliaries 'can' or 'could'. The trouble is that, even for adults, these linguistic constructions can be used both to

indicate that some action is actually impossible (e.g., 'You can't fly.'), and to indicate that some action is morally impermissible (e.g., 'You can't hit your sister.') (Matthewson, 2016; Nauze, 2008; Portner, 2009). And so, when children responded by saying that one 'can't' do something immoral, they could be interpreting the question as one that's specifically about morality, not possibility. The developmental difference that exists, then, would not reflect different conceptions about what's possible, but rather different pragmatic abilities—adults, but not children, appreciate the question is intended to ask about possibility, not morality.

On the other hand, it may be that children's understanding of possibility actually does exclude immoral events. Here, we take up a more direct test, which doesn't suffer from the problems described above. Across a series of studies, we ask children whether they believe immoral events to be 'impossible' or to require 'magic'.

The present studies

We first examine whether young children regard immoral events to be 'impossible' and compare these judgments to judgments about physically impossible, statistically improbable, and completely ordinary events. Secondly, we ask whether young children also regard immoral events to require 'magic' to occur. For both kinds of judgments, we further investigated the observed developmental changes by asking which features of these events best predict impossibility and magic judgments at different stages in development.

<div align="center">

Study 1a: Judgments of possibility

</div>

Participants

We recruited 40 4- to 5-year-olds (M_{age} = 5.10, SD_{age} = 0.45, 20 females), 40 6- to 7-year-olds (M_{age} = 6.91, SD_{age} = 0.57, 21 females), and 80 adult participants (M_{age} = 30.19, SD_{age} =

9.75, 27 females, 1 unreported). Children were recruited from the Yale Developmental

Psychology child database and at the Yale Peabody Museum of Natural History. Adult

participants were recruited through Amazon Mechanical Turk (www.mturk.com).

Methods

The study began with a series of six familiarization trials in which children judged the

possibility of simple events by answering the question, 'Is that possible or impossible?' These

familiarization trials involved events that were either clear violations of physics (e.g., the

experimenter levitating and flying to the far corner of the room) or were clearly possible (e.g.,

the experimenter crawling across the floor to the far corner of the room). Children responded by

pressing one of two buttons on an iPad labeled 'Possible' and 'Impossible' and were corrected if

they made errors on these trials, though few did (< 10% of trials).

After these familiarization trials, children were introduced to eight different scenarios in

random order (e.g., a child who doesn't want to go to school because she will miss her dad); each

scenario was accompanied by a simple visual illustration of the characters in the scenario and the

situation they were in. After being introduced to each scenario, children were then asked about

the possibility of a subsequent event occurring in that scenario. On each trial, the events were

randomly selected from one of the following four categories:

1. a violation of intuitive physics (e.g., the girl snaps her fingers and suddenly it's Saturday, so she doesn't have to go to school)

2. a highly improbable event (e.g., the girl's dad agrees to go to school with her and attends all of her classes)

3. an immoral action (e.g., the girl lies to her dad and says she's sick so that she doesn't have to go to school)

4. none of these kinds of violations (e.g., she takes a picture of her dad to school with her so she can look at the picture when she misses her dad)

After the event was presented, children indicated whether that event was possible or impossible by pressing one of two buttons on an iPad (as in the training trails). For comparison, we also collected judgments from 80 adult participants who completed an identical task online; adults were not given feedback on the training trials. All stimuli, data, analysis code and other supporting materials are available at: https://github.com/phillipsjs/deonticMagic.

Results

Our primary analysis concerns whether young children's judgments of possibility were affected by the morality and probability of an event more than older children's and adults' judgments were. We tested for this interaction using generalized linear mixed effects models (Barr, Levy, Scheepers, & Tily, 2013). We observed a main effect of Age, $\chi^2(2) = 60.26$, $p < .001$, such that younger children generally judged events to be impossible more than adults, and a main effect of Violation Type, $\chi^2(3) = 40.139$, $p < .001$, such that some kinds of violations (e.g., physical violations) were generally judged to be impossible more than others (e.g., ordinary events). More importantly, the analysis also revealed the predicted Age × Violation Type interaction effect, $\chi^2(6) = 28.46$, $p < .001$ (Figure 1).

To investigate this interaction effect, we switched to analyzing the data at the level of the 32 different events that participants were asked about.[1] For each age group, we calculated the

---

[1] This analysis approach helps to communicate that our effect was not driven by a small number of items in each category that were judged to be highly (im)possible. As suggested by the overall analysis (which is analyzed at the

mean possibility judgment for each of the events (with a score of 100 indicating that it was

always judged to be impossible, and a score of 0 indicating that it was always judged to be

possible). Pairwise comparisons between types of violations were then used to characterize how

the ratings for the different kinds of events differed from one another and changed over the
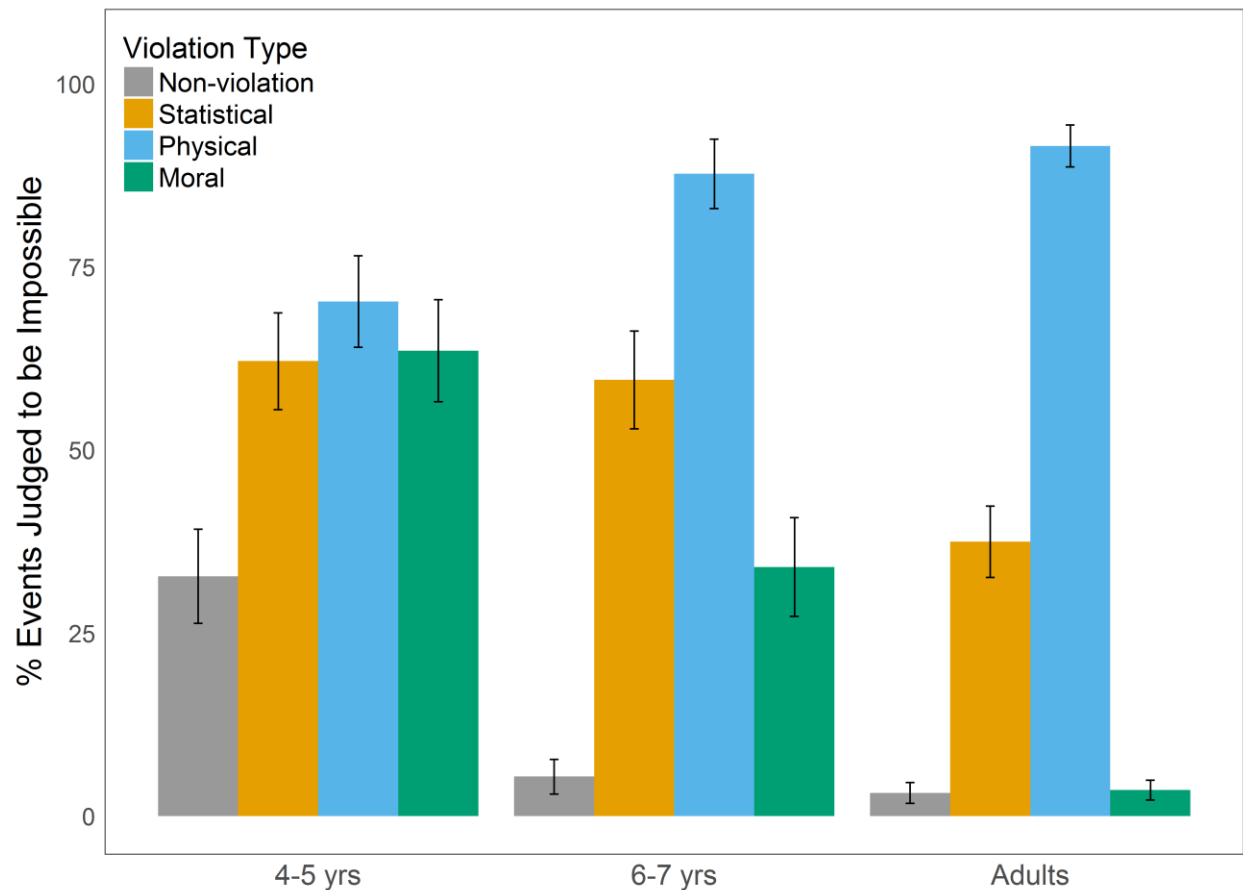
course of development.

Figure 1. Participants' mean judgments of the possibility of the events involving each kind of

violation type as a function of age group. Error bars indicate +/- 1 SEM.

Focusing first on the youngest children's judgments, we found that for 4- to 5-year-olds, the immoral events were regarded as impossible ($M = 62.64$, $SD = 18.83$) more frequently than the events without any kind of violation ($M = 34.57$, $SD = 15.71$), $t(14) = 3.24$, $p = .006$, $d = 1.62$ (and also greater than chance, $t(7) = 1.90$, $p < .050$). Moreover, events involving physical violations were not regarded as impossible ($M = 71.73$, $SD = 11.27$) significantly more than events involving moral violations, $t(14) = -1.17$, $p = .261$, $d = 0.59$. In line with previous research (Shtulman & Carey, 2007), children also tended to regard events involving probability violations as impossible ($M = 66.94$, $SD = 18.83$) more than ordinary events, $t(14) = 3.49$, $p = .004$, $d = 1.75$ (and more than chance, $t(7) = 2.28$, $p = .028$).

Six- to seven-year-olds regarded immoral events as impossible ($M = 40.81$, $SD = 17.22$) significantly less than 4- to 5-year-old children, $t(14) = -2.42$, $p = .030$, $d = 1.21$. However, they continued to regard immoral events as impossible more than events not involving any violations ($M = 7.28$, $SD = 10.91$), $t(14) = 4.65$, $p < .001$, $d = 2.33$. Six- to seven-year-olds also continued to judge that events involving probability violations were impossible ($M = 61.46$, $SD = 24.38$) more than events without any violations, $t(14) = 5.74$, $p < .001$, $d = 2.87$. Unlike 4-to 5-year-olds though, 6- to 7-year-olds regarded events involving physical violations to be impossible ($M = 90.73$, $SD = 9.51$) more than events involving either probability violations, $t(9.08) = 3.16$, $p = .011$, $d = 1.58$, or moral violations, $t(14) = 7.18$, $p < .001$, $d = 3.59$.

Finally, as expected, adult participants were significantly less likely to regard immoral events as impossible ($M = 4.64$, $SD = 4.09$) than 6- to 7-year-old children, $t(7.79) = -5.78$, $p < .001$, $d = 2.89$, and did not regard immoral events to be impossible more than events without any kind of violation ($M = 4.52$, $SD = 5.23$), $t(14) = 0.05$, $p = .961$, $d = 0.02$. Unexpectedly, adult participants did judge improbable events to be impossible ($M = 34.57$, $SD = 22.19$) more than

events not involving any kind of violation, $t(7.78) = 3.73$, $p = .006$, $d = 1.86$, suggesting that they may have occasionally perceived these as requiring that some physical law to be violated.

Interim Discussion

These results provide evidence that young children's judgments of possibility differ from that of adults, in that they judge immoral (and improbable) events to be impossible to roughly the same extent that they regard events that violate physics to be impossible. In contrast, older children and adults clearly judged immoral events to be possible and judged improbable events to be more possible than those that violated physics (though they continued to judge such events as impossible to some degree).

The finding that younger children judged the immoral events to be impossible is particularly striking considering that (1) they were trained on the terms 'possible' and 'impossible' using only events that violated physical constraints, and (2) these children are likely to have both witnessed and done many of these actions themselves (e.g., taking something from a sibling, lying to one's parent). These results make it unlikely that children's judgments of impossibility are primarily relying on a capacity for simulating how an event could occur. Instead, these results offer support for an account according to which children are relying on an undifferentiated representation of possibility, which does not clearly distinguish between different kinds of violations but treats them all similarly.

While our results show clear differences between different kinds of violations, it is also worth noting that there was significant variation within the different categories of violations. This was most striking for improbable events, some of which were rarely judged to be impossible (e.g., 11% of adults reported that it was impossible for a coach to change the rules of

a children's basketball game), while others were more consistently judged to be impossible (e.g., 75% of adults judged that it was impossible for a child to convince a friend's parents to move to a different country). One possible explanation for much of this variation is that some of our events may have been seen as violating more than one constraint. For example, participants may have viewed some of the unlikely events as also involving physical violations. Similarly, it is likely that some of the physically impossible events were seen as not only impossible but also as morally wrong (e.g., making a person fly up into the air by stamping one's foot), and that some of the immoral events were seen as not only morally wrong but also statistically improbable (e.g., a child stealing a candy bar).

To explore this issue, we next collected a set of independent ratings of the extent to which each of the events we used was (1) physically impossible, (2) statistically improbable, and (3) morally wrong. We then used these ratings to test our a priori characterization of the violation categories we created.

Study 1b: Event ratings

Participants

We recruited 76 adult participants ($M_{age} = 31.93$, $SD_{age} = 8.91$; 33 females) through Amazon Mechanical Turk (www.mturk.com).

Methods

For consistency with Study 1a, participants were first asked to identify their gender, and then were shown the eight scenarios from Study 1a, with protagonists who matched their gender. On each trial, participants were randomly assigned to make judgments of either (1) whether or not each of the four events related to that scenario was *physically impossible*, (2) whether or not each

event was *extremely unlikely*, or (3) whether each event was *very morally wrong*. Thus, on some trials, participants made four dichotomous judgments about whether four different events in a given scenario were morally wrong (or not), while on other trials, they made four judgments of whether four events were physically impossible (or not), etc. In each case, participants were always given the option to indicate that the question was 'Not Applicable' for any one of the four events. After completing this task for all 8 scenarios, participants completed a brief demographic questionnaire. All stimuli, data, analysis code and other supporting materials are available at: https://github.com/phillipsjs/deonticMagic.

Results

We excluded all responses where participants indicated that the question was not applicable.[2] For each question, we then assigned judgments that an event *did* violate the relevant norm (physical, statistical, moral) a score of 1 and assigned judgments that an event did *not* violate that norm a score of -1. Then, for each event, we calculated the mean judgment that it was physically impossible, statistically unlikely and morally wrong. Mean ratings for each event are available in Table 1 in the Supporting Materials.

We first used these ratings to characterize how the different a priori categories of violations differed from one another (see Figure S1 in the Supporting Materials). For physical impossibility, participants judged physical violations to be significantly more physically impossible than statistical violations, ($p<.001$, $d=8.81$), moral violations ($p<.001$, $d=14.78$), or non-violations ($p<.001$, $d=13.39$). For judgments of whether an event was extremely unlikely,

---

[2] The rate of not applicable responses was low ($< 15\%$) for all question $\times$ violation-type pairs, except three: many participants reasonably thought that the morality and probability questions were not applicate to physically impossible events. In addition, a smaller but still substantial number also thought the moral questions were not applicable to the statistically improbable events.

we found that participants judged both statistical violations and physical violations to be extremely unlikely, but did not distinguish between the two ($p=.107$, $d=0.86$). In comparison, participants judged moral violations and non-violations to be much more likely than statistical violations ($p<.001$, $d=5.34$ and $p<.001$, $d=9.74$, respectively). For moral wrongness, we found that participants judged moral violations to be significantly more wrong than physical violations ($p=.001$, $d=2.31$), statistical violations ($p<.001$, $d=2.67$), and non-violations ($p<.001$, $d=9.54$).

This analysis broadly supports the a priori categorization of the events used in Study 1a. That is, they provide reason for thinking that the statistical violations were *primarily* judged to be impossible because they were improbable, and moral violations were primarily judged to be impossible because they were morally wrong. At the same time, these rating also confirm that there was significant variation within the different categories. For example, participants judged that the improbable events involved violations of physics ~20% of the time, and that moral violations were highly unlikely ~25% of the time. Accordingly, a stronger test of our hypothesis would be to ignore our a priori categorization of the events altogether, and instead use only the moral, statistical and physical ratings of each event to predict judgments of possibility over the course of development. We take this analytical approach next.

Combined analyses. We combined the data from Study 1a and 1b, and then, for each age group, tested whether judgments of impossibility were uniquely predicted by the extent to which the events violated moral, statistical or physical norms. To do this, we built a separate linear model for each age group that predicted the events' impossibility using all three types of ratings (morality, statistical probability, physical impossibility). Then, to test the unique contribution of any one factor (e.g., morality), we compared this full model to one in which the factor in question was removed. In the case of morality, for example, this approach allows us to ask,

'Above and beyond any effect of statistical improbability and physical impossibility, does an event's morality uniquely predict whether the event will be judged to be possible?'. A similar question can then be asked for statistical probability and physical possibility as well.

These analyses revealed that for 4- to 5-year-olds, morality and probability uniquely predicted their judgments of possibility ($\chi^2(1) = 6.897$, $p = .009$; $\chi^2(1) = 5.79$, $p = .016$, respectively), while the extent to which the event violated physics did not ($\chi^2(1) = 0.43$, $p = .510$). The latter result suggests that for 4- to 5-year-olds, the events' physical impossibility was not predictive above and beyond the events' statistical improbability. For 6- to 7-year-olds, morality continued to play a significant, though smaller, role ($\chi^2(1) = 5.01$, $p = .025$), probability played a more substantial, role ($\chi^2(1) = 14.503$, $p < .001$), and physical impossibility also began to play a unique role ($\chi^2(1) = 6.388$, $p = .011$). For adults, we found that morality played no unique role ($\chi^2(1) < 0.01$, $p = .986$), probability continued to play a significant role ($\chi^2(1) = 16.289$, $p < .001$), and physical impossibility now played a dominant role ($\chi^2(1) = 56.257$, $p < .001$).

These results provide even stronger evidence that the uniquely moral features drove 4- to 5-year-olds' (and to a lesser extent 6- to 7-year olds') judgments of possibility, but that over time, judgments of possibility came to be increasingly predicted by whether the event violated physics. That is, they provide more evidence that young children were relying on an undifferentiated representation of possibility, which became more differentiated over the course of development.

Differentiation over development. Given these results, we next sought to quantitatively characterize the pattern of differentiation that occurs over the course of development. To do this, we calculated and analyzed two scores for each event of the 32 events we tested. The first was a

measure of the extent to which each event did vs. did not involve *any* violation, which was calculated by summing across the scores for whether the event was physically impossible, statistically unlikely, and morally wrong. The second was a measure of whether the event was thought to have involved specifically physical violations. We then investigated how these two scores changed in the extent to which they predicted judgments of impossibility over the course of development.

A comparison of linear mixed-effects models revealed interaction effects between both (1) age group and whether the event involved violations, $\chi^2(2) = 6.88$, $p = .032$, and (2) between age group and whether the event involved was physically impossible, $\chi^2(2) = 29.50$, $p < .001$. As predicted, these interactions arose from the fact that younger children's judgments of impossibility were completely driven by the extent to which the event involved *any* kind of violation, and not by whether the event was specifically physically impossible (see Figure 2a), but that this pattern fully reversed itself over the course of development (see Figure 2b). By adulthood, participants exhibited a more differentiated representation of possibility, judging events to be impossible mostly in virtue of whether or not they would violate physical laws.
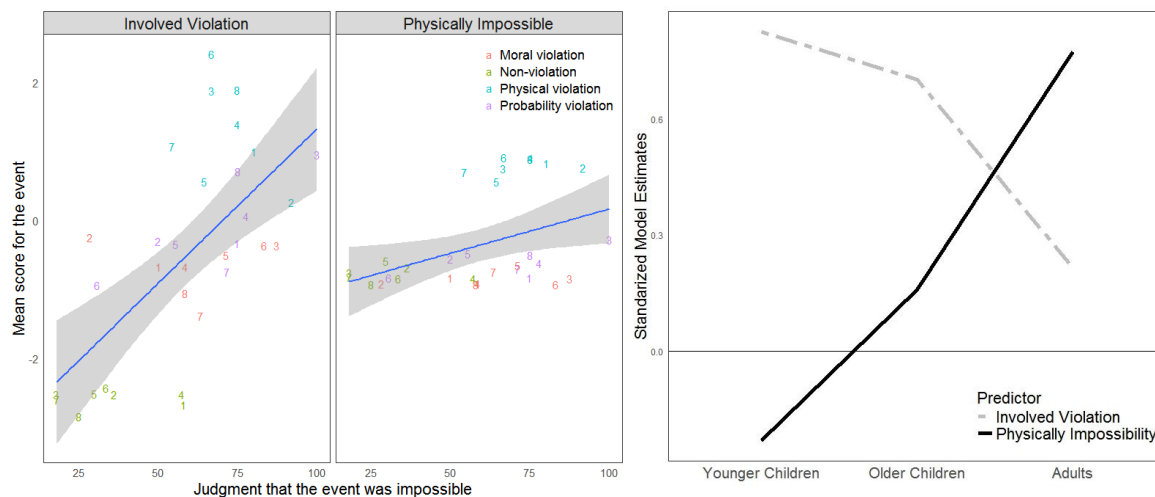
Figure 2. (a) Depiction of the relationship between 4- to 5-year-olds' judgments of possibility and the extent to which that event involved any kind of violation (left) and the extent to which that event was physically impossible (right). (b) Standardized model estimates for the extent to which the impossibility of younger children, older children, and adults are predicted by whether the event was physically impossible (solid dark line) or involved any kind of violation (dashed grey line).

Discussion

Taken together, these results suggest that young children possess an undifferentiated representation of possibility that becomes increasingly differentiated over the course of development. Study 1a found that young children judged that immoral, improbable, and impossible events are all impossible. Study 1b provided a stronger test by capitalizing on the natural variation in the extent to which each event violated physical, statistical, and moral norms. This approach revealed two key findings. First, we found evidence that young children's representation of possibility allows them to represent an event as possible only insofar as it does not violate any relevant constraint. Second, we found that over the course of development this representation becomes more differentiated: older children and adults increasingly exhibit the capacity to represent an event as physically possible even if (statistically speaking) it will never occur or (morally speaking) it should not occur. However, even for adults, we found that their representation of possibility is not perfectly differentiated – we return to this point in the General Discussion.

Study 2a: Judgments of Magic

One concern with the approach taken in Study 1a-b is that young children may be relatively unfamiliar with the terms 'possible' and 'impossible'. While we sought to ground children's understanding of the task by introducing these terms in the context of events that were physically impossible, it is conceivable that the children's lack of familiarity with the term still led them to misunderstand the task they were asked to complete. A stronger test would adopt the methods used in previous work, and ask children whether or not the occurrence of immoral events would require 'magic' (see, e.g., Browne & Wolley, 2004; Shtulman, 2009).

This concern with judgments of 'possibility' should not extend to judgments of 'magic'. Even young children have a relative familiarity with magic (Johnson & Harris, 1994). Moreover, young children's experience of 'magic' often involves instances in which magic is associated with good rather than evil (the tooth fairy, Harry Potter, genies, etc.). Accordingly, it is unlikely that children would be inclined to judge that immoral events require magic simply as a way of indicating that they are wrong. We pursue this approach next.

Participants

We recruited 20 3-year-olds (M_{age} = 3.66, SD_{age} = 0.27, 10 females), 40 4- to 5-year-olds (M_{age} = 4.95, SD_{age} = 0.57, 19 females), 40 6- to 7-year-olds (M_{age} = 7.04, SD_{age} = 0.65, 21 females), and 79 adult participants (M_{age} = 30.57, SD_{age} = 9.53, 29 females). As in Study 1, children were recruited from the Yale Developmental Psychology child database and at the Yale Peabody Museum of Natural History. Adult participants were recruited through Amazon Mechanical Turk (www.mturk.com).

Methods

In addition to having children make judgments of whether an event would require 'magic',

the methods in this study differed from those used in Study 1 in two ways. First, before

conducting the study, we decided to included a group of younger children (3-year-olds), as the

task was made substantially less demanding by not using the terms 'possible' and 'impossible'.

Second, because even adults judged a surprising number of the probability violations to be

impossible (and sometimes thought they involved physical violations), these items were replaced

with less extreme probability violations to allow for a better test of the changes that occur

throughout development. For example, rather than the girl's dad attending school with her, her

dad says it is okay if she doesn't go to school today. All stimuli and other supporting materials

are available at: https://github.com/phillipsjs/deonticMagic.

Before beginning the study, children were again trained on the task of making 'magic'

judgments using the same six training events from Study 1, which clearly violated physics (e.g.,

stretching one's arm long enough to reach the ceiling) or were clearly possible (e.g., touching the

ceiling after climbing a tall ladder). Consistent with prior work on children's intuitive grasp of

magic (Johnson & Harris, 1994), even they youngest children very rarely needed correction.

Results

The data were first analyzed as in Study 1a using generalized linear mixed-effects

models. We again observed a main effect of Age, $\chi^2(3) = 43.98$, $p < .001$, such that children

overall tended to judge events to require magic more than adults, and a main effect of Violation

Type, $\chi^2(3) = 141.98$, $p < .001$, such that some violations tended to be judged to require magic

more than others (e.g., physical violations vs. non-violations). More importantly, we again

observed a significant Violation Type $\times$ Age interaction effect, $\chi^2(9) = 130.60$, $p < .001$ (Figure

3). As in Study 1, we next investigated this interaction effect by analyzing the data at the level of
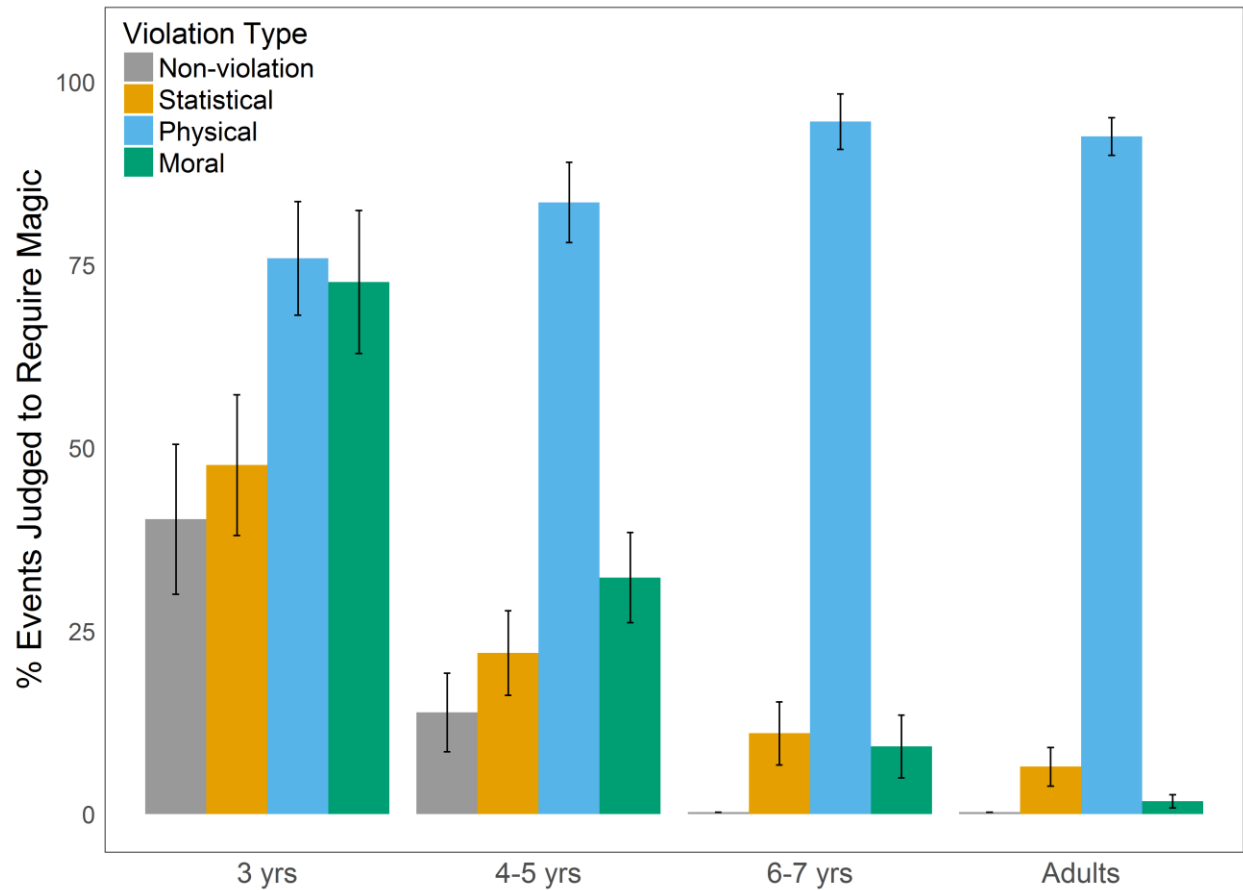
the 32 different events.

Figure 3. Participants' mean judgments of whether magic would be required for the events

involving each kind of violation type as a function of age group. Error bars indicate +/- 1 SEM.

Focusing first on the youngest children's judgments of possibility, we found that 3-year-

olds regarded immoral events as requiring magic to occur ($M = 72.69$, $SD = 41.40$) significantly

more than events without any kind of violation ($M = 40.28$, $SD = 43.37$), $t(14) = 2.92$, $p = .011$, d

$= 1.46$ (and more than chance, $t(7) = 2.20$, $p = .032$). There was also a weak trend such that they

regarded immoral events to require magic more than the new improbable events ($M = 47.69$, SD

= 40.81), $t(14) = 1.76$, $p = .099$, $d = 0.88$. Moreover, they did not tend to judge that events involving physical violations required magic ($M = 75.93$, $SD = 32.95$) more than events involving moral violations, $t(14) = 0.91$, $p = .377$, $d = 0.46$.

Four- to five-year-olds were significantly less likely to judge that events involving moral violations required magic ($M = 32.31$, $SD = 36.86$) than 3-year-old children were, $t(14) = -3.33$, $p = .005$, $d = 1.67$, and judged that these events required less magic than events involving physical violations ($M = 83.56$, $SD = 33.25$), $t(14) = -5.77$, $p < .001$, $d = 2.88$. However, 4- to 5-year-olds continued to judge that events involving moral violations more often required magic than events not involving any violations ($M = 13.89$, $SD = 30.81$), $t(14) = 2.23$, $p = .043$, $d = 1.12$. They did not regard improbable events to require significantly less magic ($M = 11.04$, $SD = 26.21$) than immoral events, $t(14) = -1.00$, $p = .333$, $d = 0.50$, or significantly more magic than events that did not involve any kind of violation, $t(14) = 1.58$, $p = .136$, $d = 0.79$.

Six- to seven-year-olds were significantly less likely to judge that events involving moral violations required magic ($M = 9.70$, $SD = 10.54$) than 4- to 5-year-old children were, $t(14) = -2.29$, $p = .038$, $d = 1.15$. However, we continued to observe that immoral events were regarded as requiring magic more than events that did not involve any kind of violation ($M = 0.00$, $SD = 0.00$), $t(7) = 2.60$, $p = .035$, $d = 1.30$. We also observed a marginal effect such that improbable events were regarded as requiring magic ($M = 10.31$, $SD = 12.89$) more than events that did not involve any kind of violation, $t(7) = 2.26$, $p = .058$, $d = 1.13$. Events involving physical violations were judged to require magic ($M = 97.30$, $SD = 3.80$) more than all other event types, p's $< .001$, d's > 11.05.

Adults' judgments of magic for immoral events ($M = 3.71$, $SD = 5.36$) did not differ significantly from those of 6- to 7-year-old participants, $t(14) = -1.43$, $p = .174$, $d = 0.72$. Adult

participants did not regard immoral events as requiring magic more than events without any kind of violation ($M = 0.00$, $SD = 0.00$), $t(7) = 1.96$, $p = .091$, $d = 0.98$, but did judge that improbable events more required magic ($M = 6.93$, $SD = 3.93$), $t(7) = 4.99$, $p = .002$, $d = 2.49$. Unsurprisingly, events involving physical violations were regarded as requiring magic ($M = 91.63$, $SD = 5.70$) more than all other event types, p's $< .001$, d's > 17.30.

Interim Discussion

As predicted, we found that young children judge immoral events to require magic just as much as events that violate physics, and that even older children judge immoral events to require magic more than ordinary events. At a qualitative level, these results provide further support for an account on which children possess an undifferentiated representation of possibility, which becomes more differentiated over the course of development.

At the same time, we also observed a notable difference in children's judgments of magic and their judgments of impossibility (from Study 1). Specifically, judgments of magic exhibited a shifted developmental trajectory, with 4- to 5-year-old children less frequently judging that moral violations would require magic than that they were impossible. One plausible explanation for this difference is that children's early and frequent exposure to the concept of magic facilitates the development of domain-specific strategies for understanding when magic is required. To the extent that children recruit these domain-specific strategies, they will not need to rely on their more general representation of what is possible. Three-year-olds likely have not yet developed such strategies, and thus must rely on their general representation of possibility, which treats all violations as distinct from the possible. We return to this hypothesis after further investigating the developmental trajectory of children's judgments of magic.

As in Study 1, we next collected independent ratings of the extent to which each of the events we used is (1) physically impossible, (2) statistically improbable, and (3) morally wrong. These ratings allowed us to conduct three further sets of analyses. First, we again used these ratings to characterize the a priori groups, including the new statistically improbable events. Second, we then used the ratings of each event's morality, probability, and physical possibility (rather than the a priori categories) to predict children's and adults' judgments of magic. Third, and finally, we used these ratings to create an index of the extent to which each violated *any* norm vs. a specifically physical norm, and then investigated the process of differentiation that occurs across development.

Study 2b: Event ratings

Participants

We recruited 76 adult participants ($M_{age} = 32.50$, $SD_{age} = 9.10$; 25 females) through Amazon Mechanical Turk (www.mturk.com).

Methods

The methods for this study were identical with those of Study 1b, except that we used the events judged in Study 2a rather than the events judged in Study 1a. Data, analysis code and other supporting materials are available at: https://github.com/phillipsjs/deonticMagic.

Results

We again excluded all responses where participants indicated that the question was not applicable.[3] We then assigned judgments that an event *did* violate the relevant norm (physical, statistical, moral) a score of 1 and assigned judgments that an event did *not* violate the norm a score of -1. For each event, we calculated the mean judgment that it was physically impossible, statistically unlikely and morally wrong. Mean ratings for each event are available in Table 2 in the Supporting Materials.

Participants' judgments of physical impossibility revealed that physical violations were judged to be significantly more physically impossible than statistical violations, ($p<.001$, $d=17.84$), moral violations ($p<.001$, $d=19.22$), or non-violations ($p<.001$, $d=26.74$). For judgments of whether an event was extremely unlikely, we found that participants judged physical violations to be the most unlikely and more unlikely than statistical violations ($p<.001$, $d=5.35$). In comparison, participants judged moral violations and non-violations to be much more likely than statistical violations ($p<.001$, $d=2.39$ and $p<.001$, $d=3.98$, respectively). For moral wrongness, we found that participants judged moral violations to be significantly more wrong than physical violations ($p<.001$, $d=2.77$), statistical violations ($p<.001$, $d=7.89$), and non-violations ($p<.001$, $d=13.24$). See Figure S2 in the Supporting Materials.

Of particular interest in this study were the new statistically improbable events. As expected, these events were still seen as much more improbable than the immoral or ordinary events, but they were also judged to be much less improbable than the events used in Study 1, and were no longer seen as requiring violations of physics. These analyses thus offer general

---

[3] Similar to Study 1b, the rate of 'not applicable' responses was low ($< 9\%$) for all question $\times$ violation-type pairs, except that participants again often thought that the morality and probability questions were not applicate to physically impossible events.

support for the a priori categorization of the events used in Study 2a, suggesting again that the

unique features of the immoral and improbable events were responsible for judgments that these

events required magic.

Setting our a priori categories to one side, we next used the variation in the three types of

ratings across the different events to predict judgments of magic for each age group. This

approach again provides a stronger test of whether the specifically moral, statistical, or physical

features of each event uniquely explained judgments of magic, and allowed us to test how this

changes across development.

Combined analyses. We combined the data from Study 2a and 2b, and again built a separate

linear model for each age group, which predicted the mean judgment of magic for each event

using all three types of ratings (ratings of morality, statistical probability, physical impossibility).

Then, to test the *unique* contribution of any one factor (e.g., morality), we compared this full

model to one in which the factor in question was removed, and asked whether the second model

predicted judgments of magic less well.

These analyses revealed that for 3-year-olds, morality uniquely predicted judgments of

magic ($\chi^2(1) = 9.170$, $p = .002$). In contrast, probability and physical possibility did not uniquely

predict judgments of magic ($\chi^2(1) = 0.325$, $p = .569$; $\chi^2(1) = 0.767$, $p = .381$, respectively),

indicating that both ratings could account for much of the same variance in 3-year-olds'

judgments of magic. For 4- to 5-year-olds, morality continued to uniquely predict their

judgments of magic ($\chi^2(1) = 4.856$, $p = .028$), and probability continued to not uniquely predict

their judgments of magic ($\chi^2(1) = 2.29$, $p = .130$). In contrast to 3-year-olds however, judgments

of physical possibility did uniquely predict judgments of magic for 4- to 5-year-olds ($\chi^2(1) =$

7.59, $p = .006$). For 6- to 7-year-olds, morality's unique contribution was only a trend ($\chi^2(1) =$

3.01, $p = .083$), probability played a more substantial, role ($\chi^2(1) = 8.35$, $p = .004$), and physical impossibility played a strong unique role ($\chi^2(1) = 38.714$, $p < .001$). For adults, morality played no unique role ($\chi^2(1) = 0.034$, $p = .855$), probability played a smaller but still significant role ($\chi^2(1) = 4.895$, $p = .027$), and physical impossibility played an overwhelmingly strong role ($\chi^2(1) = 75.66$, $p < .001$).

These results provide stronger evidence that the uniquely moral features drove 3-year-olds' (and to a lesser extent 4- to 7-year olds') judgments of magic, but that over time, judgments of possibility came to be primarily predicted by whether the event violated physics. In short, judgments of magic provide strong evidence that young children rely on a representation of possibility that treats moral violations as similar to physical violations, but that judgments of magic quickly become clearly differentiated over the course of development.

Differentiation over development. As in Study 1, we next sought to quantitatively characterize the changes that occur over the course of development. To do this, we again calculated and analyzed both an index of the extent to which each of the 32 events violated *any* norm (moral, statistical, physical), and an index of the extent to which the event involved a specifically physical violation. We then asked how well these two scores predicted magic judgments over the course of development.

A comparison of linear mixed-effects models again revealed the predicted interaction effects between both (1) age group and whether the event involved any violation, $\chi^2(3) = 10.74$, $p = .013$, and (2) between age group and whether the event involved was physically impossible, $\chi^2(3) = 39.57$, $p < .001$. As with judgments of possibility, these interactions arose from the fact that the youngest children's judgments of magic were completely driven by the extent to which the event involved any kind of violation, and not by whether the event was specifically

physically impossible (see Figure 4a), but that this pattern again switched over the course of development (see Figure 4b). By the time of adulthood, judgments of magic are primarily driven by whether the event involved a violation of some specifically physical constraint.
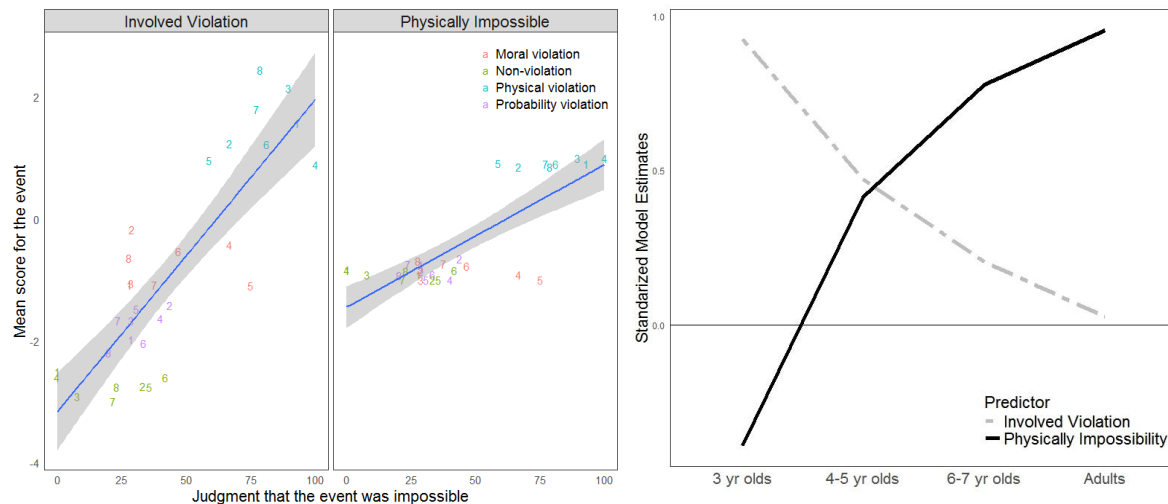

Figure 4. (a) Depiction of the relationship between younger (3- 5-year-old) children's judgments of magic and the extent to which that event involved any kind of violation (left) and the extent to which that event was physically impossible (right). (b) Standardized model estimates for the extent to which the magic judgments of 3-year-olds, 4- to 5-year olds, 6- to 7-year-olds and adults are predicted by whether the event was physically impossible (solid dark line) or involved any kind of violation (dashed grey line).

Discussion

Study 2a found that the youngest children judged that immoral events, much like physically impossible events, require magic. Study 2b provided a stronger test by capitalizing on the variation in the extent to which each event violated physical, statistical and moral norms. This approach again revealed two key findings. First, we found evidence that young children's representation of possibility allows them to represent an event as possible only to the extent that

it does not violate any relevant constraint. Secondly, we found that over the course of development judgments of magic become much more differentiated.

This complementary set of analyses also provides further support for our earlier hypothesis that children develop domain-specific strategies for making judgments of 'magic' earlier than for making judgments of 'possibility'. Four- to 5-year-old children's judgments of magic were partially predicted by whether the event involved some specifically physical violation, and 6- to 7-year-old children's judgments were primarily predicted by whether the event involved a physical violation (Figure 4b). Judgments of 'possibility', in contrast, showed a much more delayed developmental time course. (Figure 2b).

General Discussion

Earlier research provided evidence that young children believe that highly improbable events are impossible (Shtluman & Carey, 2007; Shtulman, 2009). A plausible explanation for this result was that young children had difficulty simulating how these events could actually occur (Brown & Woolley, 2004; Shtulman & Carey, 2007). In two two-part studies, we proposed and tested an alternative explanation of this finding which suggests that children differ from adults, not in their capacity for simulating how various events might occur, but in their representation of possibility. In support of this alternative, we found that young children believe that immoral events are impossible and that their occurrence would require magic.

Unlike earlier research on children's understanding of immoral or socially unacceptable events (Chernyak et al., 2013, 2010; Kalish, 1998), these results cannot be explained by an appeal to linguistic ambiguity, as we did not employ modal auxiliaries such as 'cannot' or 'could not.' Thus, the present studies provide positive reason for believing that young children

understand immoral events to be impossible, and consequently, that young children possess an undifferentiated representation of possibility, which gradually becomes more differentiated throughout the course of development.

Understanding children's representation of immoral events as impossible

One radical interpretation of our findings is that they indicate that children believe immoral acts are impossible in much the same way that adults believe acts that violate physics are impossible. But our analysis is much more modest. For adults, who have developed a more differentiated representation of possibility, believing that an act is impossible means believing that its occurrence would violate some specifically *physical* constraint. Yet, this is precisely the differentiated notion of possibility we believe that young children lack. Thus, rather than interpreting our results to show that children believe immoral events are impossible in the sense that they violate *physical* constraints, we take our results to show that children's understanding of possibility differs in a much more general way from adults' understanding.

Specifically, our proposal is that young children's representation of possibility differs from adults' in that it is *undifferentiated*. As we suggested earlier, it may be helpful to think of this representation by analogy to the adult understanding of what is *abnormal*. Just like children's understanding of possibility, adult's understanding of *abnormality* does not clearly distinguish statistical violations from physical violations, or either of these from moral violations—all of these violations are represented as abnormal, and are clearly differentiated from what is normal. Much the same seems to hold for young children's understanding of what is possible.

This analogy helps explain why we reject the radical interpretation. When considering adult's understanding of what is abnormal, we would not take the fact that this specific representation does not clearly distinguish between different kinds of violations to provide evidence that adults have no other ways of distinguishing between these different kinds of events. In the same way, when considering children's understanding of possibility, we do not take the fact that children find moral, statistical, and physical violations to all be impossible to show that they do not have any other ways of representing these events which allow them to distinguish between them (*cf*. Kalish, 1998).

Indeed, there is good evidence that children are able to distinguish different kinds of violations from one another under certain circumstances. For example, Weisberg and Sobel (2012) presented children with fictional stories that involved a series of improbable events happening and then asked children to propose ways that the story might continue. Young children tended to propose both ordinary and improbable events, but tended not to propose events that violated physics, which suggests that they had some way of distinguishing between improbable and impossible events in these fictional contexts (Weisberg & Sobel, 2012). This kind of result fits well with the theoretical proposal we have offered: while young children do not differentiate between these events in terms of their *possibility* in real life, they clearly have other ways of representing these events that allow them to realize that within these fictional contexts, improbable events will occur, while physically impossible ones will not.

At the same time, though, it is worth emphasizing that there is another way in which our findings are likely to have surprisingly widespread implications. Specifically, we take our results to show that children's actual representation of which events are possible to be constrained not only be descriptive norms (e.g., physics, probability) but also prescriptive norms (e.g., morality).

If this is correct, then any part of their cognition that depends on a general representation of possibility will be similarly affected. In Study 2, for example, we showed that the constraints of morality extend to their judgments of 'magic'. And, in much the same way, we would predict that many other judgments that children make (predicting others' future behavior, guessing which things happened given some evidence, and so on) will be similarly affected.

Relation to modal cognition in adults.

Many of the high-level judgments that have interested cognitive scientists (e.g., causal reasoning, language comprehension, moral judgment) rely on a capacity for understanding and reasoning about the set of events that are possible in a given situation (Lewis, 2013; Pearl, 2009; Stalnaker, 2002; Kratzer, 2012; Kane, 2001, Hart & Honore, 1985). This capacity, often referred to as 'modal cognition', has played a key role in both theoretical work on these aspects of cognition, and more recently, in emerging empirical research in adult cognition (Gerstenberg & Tenenbaum, in press; Kominsky, Phillips, Gerstenberg, Lagnado & Knobe, 2015; Icard, Kominsky & Knobe, 2017; Phillips, Luguri, & Knobe, 2015; Young & Phillips, 2012).

Critically, across all of these cases, research has shown that the representation of possibilities, on which adults implicitly rely, is one that consistently excludes both improbable and immoral events. To illustrate, consider people's judgments of whether an agent was *forced* to do an action. For example, we may be trying to decide whether someone was forced to take out a second mortgage to pay her medical bills. Intuitively, it will only be the case that she was forced to do this if it was not possible for her to do something else instead. While this much is intuitive, an important question is what notion of 'possible' we are relying on when making these judgments. Helping to answer this question, research has shown that when people ordinarily make judgments of 'force', they tend to not only reject alternative possibilities when they

involve violations of physics (e.g., growing money on trees), but also when they involve violations of morality (e.g., stealing the money to pay the medical bills) or violations of probability (e.g., winning the lottery and paying the bills) (Phillips & Cushman, 2017; Young & Phillips, 2011). That is, they only tend to agree that the agent was not forced when there were alternative possibilities that were physically possible, statistically likely, and morally good. As a whole, what this research has suggested is that even adults may be implicitly relying on an undifferentiated representation of possibility when making high-level judgments about force, causation, moral responsibility, and so on (for a general theoretical account, see Phillips & Knobe, forthcoming).

Other empirical work has also sought to more directly investigate adults' understanding of possibility and provided some evidence that it too shows signs of being constrained by both descriptive norms (statistics, physics) and by prescriptive norms (morality, rationality) (Barnes & Black, 2016; Phillips & Cushman, 2017; Shtulman & Tong, 2013). One recent finding is that when adult participants were forced to make judgments of the possibility of various events under time pressure, they frequently judged that immoral, improbable, and physically impossible events were all impossible, and only differentiated between these constraints when they had time to reflect. (Phillips & Cushman, 2017).

Plainly, this connects to the present results. In particular, the undifferentiated representation of possibility we have observed in young children may be preserved at an implicit level even into adulthood. If this is correct, then the developmental shift we observed would not be due to the maturation of a single system or way of representing possibilities, but rather due to the emergence of an independent, differentiated way of representing possibilities. This possibility provides a promising avenue for future research.

Acknowledgements

References

Barnes, J., & Black, J. (2016). Impossible or improbable: The difficulty of imagining morally

   deviant worlds. *Imagination, Cognition, and Personality*. 36(1), 27–40.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for

   confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68* (3),

   255–278.

Bear, A. & Knobe, J. (in press). Normality: Part descriptive, part prescriptive. *Cognition*.

Browne, C. A., & Woolley, J. D. (2004). Preschoolers' magical explanations for violations of

   physical, social, and mental laws. *Journal of Cognition and Development*, *5* (2), 239–260.

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 (7), 1343–1355.

Chernyak, N., Kushnir, T., & Wellman, H. (2010). Developing notions of free will:

   Preschoolers' understanding of how intangible constraints bind their freedom of choice. In

   Proceedings of the 32nd Annual Conference of the Cognitive Science Society (pp. 2602–

   2606). Cognitive Science Society.

Gerstenberg, T. & Tenenbaum, J. B. (in press). Intuitive Theories. In M. Waldmann (Ed.) *Oxford

   Handbook of Causal Reasoning*. Oxford University Press.

Icard, T. F., Kominsky, J. F., & Knobe, J. (2017). Normality and actual causal strength.

   Cognition, 161, 80–93. http://doi.org/10.1016/j.cognition.2017.01.010

Johnson, C. N., & Harris, P. L. (1994). Magic: Special but not excluded. *British Journal of*

   Developmental Psychology, *12* (1), 35–51.

Kalish, C. (1998). Reasons and causes: Children's understanding of conformity to social rules

   and physical laws. *Child development*, *69* (3), 706–720.

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* (Vol. 36). Oxford

   University Press.

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. (1995). Traditional and evaluative aspects of

   flexibility in gender roles, social conventions, moral rules, and physical laws. *Child

   Development*, *66*(2), 515–31.

Komatsu, L. K., & Galotti, K. M. (1986). Children's reasoning about social, physical, and logical

   regularities. *Child Development*, 57, 413-420.

Matthewson, L. (2016). Modality. In M. Aloni & P. Dekker (Eds.), *Cambridge Handbook of

   Formal Semantics*. Cambridge: Cambridge University Press.

Nauze, F. (2008). *Modality in typological perspective*. (Doctoral dissertation). 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.

Phillips, J. & Cushman, F. (2017). Morality constrains the default representation of what is

possible. *Proceedings of the National Academy of Sciences*.

Phillips, J. & Knobe, J. (forthcoming). 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.

Portner, P. (2009). *Modality*. OUP Oxford.

Samuels, A., & Taylor, M. (1994). Children's ability to distinguish fantasy events from real life

events. B*ritish Journal of Developmental Psychology*, 12, 417-427.

Shtulman, A. (2009). The development of possibility judgment within and across domains.

Cognitive Development, *24*(3), 293–309. http://doi.org/10.1016/j.cogdev.2008.12.006

Shtulman, A., & Carey, S. (2007). Improbable or impossible? how children reason about the

possibility of extraordinary events. *Child Development*, *78* (3), 1015–1032.

Shtulman, A., & Tong, L. (2013). Cognitive parallels between moral judgment and modal

judgment. *Psychonomic bulletin & review*, *20* (6), 1327–1335.

Weisberg, D. S., & Sobel, D. M. (2012). Young children discriminate improbable from

impossible events in fiction. *Cognitive Development, 27*, 90-98.