

The explanatory structure of unexplainable events: Causal constraints on magical reasoning

Andrew Shtulman¹ · Caitlin Morgan¹

Published online: 7 February 2017
© Psychonomic Society, Inc. 2016

Abstract A common intuition, often captured in fiction, is that some impossible events (e.g., levitating a stone) are “more impossible” than others (e.g., levitating a feather). We investigated the source of this intuition, hypothesizing that graded notions of impossibility arise from explanatory considerations logically precluded by the violation at hand but still taken into account. Studies 1–4 involved college undergraduates ($n = 357$), and Study 5 involved preschool-aged children ($n = 32$). In Studies 1 and 2, participants saw pairs of magical spells that violated one of 18 causal principles—six physical, six biological, and six psychological—and were asked to indicate which spell would be more difficult to learn. Both spells violated the same causal principle but differed in their relation to a subsidiary principle. Participants’ judgments of spell difficulty honored the subsidiary principle, even when participants were given the option of judging the two spells equally difficult. Study 3 replicated those effects with Likert-type ratings; Study 4 replicated them in an open-ended version of the task in which participants generated their own causal violations; and Study 5 replicated them with children. Taken together, these findings suggest that events that defy causal explanation are interpreted in terms of explanatory considerations that hold in the absence of such violations.

Keywords Causal reasoning · High order cognition · Cognitive development · Concepts

✉ Andrew Shtulman
shtulman@oxy.edu

¹ Department of Psychology, Occidental College, Los Angeles, CA 90041, USA

Introduction

The animator Walt Disney is best known for the cartoons he created in the later part of his career, not those he created earlier. Disney’s early cartoons were surreal: Clothes jumped from clothes lines and ran around the yard; sausages jumped from grills and danced in a kick line; pianos turned insolent and bit their players. Before making feature-length films, Disney decided that his cartoons had to be “plausibly impossible”: they could violate some of the audience’s real-world expectations but not too many. In the movie *Snow White*, for instance, it was deemed plausibly impossible for forest animals to communicate with Snow White but implausibly impossible for them to double in size or to ooze through keyholes (Lane, 2006).

The idea that something could be *plausibly* impossible is paradoxical. An impossible event, from a psychological point of view, is an event that violates a property of the world deemed true by necessity—a property that holds in all conceivable permutations of reality (Nichols, 2006; Shtulman, 2009; Shtulman & Carey, 2007). This kind of violation is presumably black-and-white; an event either violates a “law of nature” and cannot occur in the real world or violates no such law and can occur, whatever the likelihood. Considerations of plausibility—or subjective probability of occurrence—make sense when applied to events that can occur but makes much less sense when applied to events that cannot.

Still, fictional narratives frequently make distinctions between impossible events that are plausible and those that are not. In the fictional world of Star Wars, the Jedi master Yoda teaches the Jedi apprentice Luke Skywalker to levitate stones before he teaches him to levitate a starship. In the fictional world of Harry Potter, the potions instructor Severus Snape teaches his students how to brew forgetfulness before he teaches them how to brew endurance. And in the fictional world of Cinderella, Cinderella’s fairy godmother turns a

pumpkin into a stagecoach and a horse into a coachman rather than turning a horse into a stagecoach and a pumpkin into a coachman.

Starships are heavier than rocks; endurance requires more mental effort than forgetfulness; and pumpkins resemble stagecoaches more than they resemble coachmen. But why should these considerations influence our perception of patently impossible events? The magic of levitation severs the connection between weight and liftability; the magic of potions severs the connection between effort and ability; and the magic of transfiguration severs the connection between physical resemblance and mutability. Nevertheless, the genre of magical realism is full of impossible events that honor severed connections like these. Honoring such connections is arguably what constitutes the “realism” in magical realism.

The patterns described above are common in fiction (Lewis 1978; Ryan, 1980), and they suggest that our notions of impossibility are graded rather than categorical. Graded notions of impossibility are of interest to the study of causal reasoning and explanation for at least two reasons. First, they signify a quirk in our causal-explanatory knowledge in the same way that visual illusions signify a quirk in our visual systems. When we apply causal-explanatory knowledge to events that violate the foundations of that knowledge, we appear to register the violation while also failing to register it, similar to how we can perceive a triangle in the Kanizsa illusion (the illusion in which three black circles appear to be partially occluded by a white triangle) while also failing to perceive it. Our inconsistent response to impossible events suggests that these events reveal fault lines in the underlying structure of causal-explanatory knowledge.

Second, graded notions of impossibility play an explanatory role in causal narratives about fictional events. Impossible events, by definition, are relegated to the realm of fiction. We would never encounter such events in real life, but we spend an inordinate amount of time thinking about such events in the context of fictional worlds that violate our everyday understanding of causation (Taylor, 2013). When engaged in those worlds, we share intuitions that could neither be learned nor verified through observation of the real world. That is, we share intuitions about events that no one has ever experienced nor *could* ever experience—intuitions about which impossible events are plausible, how those events relate to ordinary (possible) events, where those events can happen, when those events can happen, who can make those events happen, and so forth. Impossible events are central to the plots of many fictional narratives (such as *Star Wars*, *Harry Potter*, and *Cinderella*), and it is unclear how or why they play such a role in these shared cognitive representations.

Psychologists have not, to our knowledge, studied graded notions of impossibility, but they have studied notions of impossibility in general. Following the work of Keil (1979) and Boyer (1999), impossible events have typically been

characterized as *ontological violations*. An ontology is a basic category of existence, such as “object” or “animal,” and an ontological violation is a violation of one of the category’s core properties, such as solidity (in the case of objects) or mortality (in the case of animals). Ontological violations have been shown to affect an event’s memorability in that events that violate one or two ontological commitments (e.g., a tree that never dies) are more memorable than events that violate no ontological commitments (e.g., a tree that bears fruit) or events that violate several ontological commitments (e.g., a tree that never dies, can speak, and floats in the air) (Atran & Norenzayan, 2004; Banerjee, Haque, and Spelke, 2013; Barrett & Nyhof, 2001; Boyer & Ramble, 2001). The memorability of an ontological violation affects, in turn, how often and how reliably these ideas are passed from one person to another and from one generation to another.

Such effects tell us that an event’s memorability is shaped by its conformity to an underlying ontology, but they tell us less about an event’s *plausibility*. Impossible events that violate a minimal number of ontological commitments are not always plausible (in the sense of being believable), and impossible events that are plausible do not always violate a minimal number of ontological commitments (Gervais & Henrich, 2010; Purzycki & Willard, 2016). For example, fictional characters like vampires and zombies violate a “cognitively optimal” number of ontological commitments but are not commonly believed to exist, whereas religious entities like the Abrahamic God—an omniscient, omnipotent, omnibenevolent, omnipresent, invisible, and immortal being—violate several ontological commitments but are commonly believed to exist (Shtulman, 2008; Shtulman & Lindeman, 2016).

The factors responsible for a concept’s plausibility may be distinct from those responsible for its memorability—namely, its ontological structure—but plausibility and ontological structure are not entirely unrelated, as shown by analyses of the representation of impossible events in fiction (Kelly & Keil, 1985; Norenzayan, Atran, Faulkner, & Schaller, 2006; Upal, 2011). Some impossible events appear in fiction more often than others, and many scholars have used frequency of representation as a measure of plausibility, on the assumption that the more often an event is represented, the better it resonates with readers or the better it fits within a larger narrative. Kelly and Keil (1985) applied this logic to the dozens of magical transformations described in Ovid’s *Metamorphoses* and Grimms’ *Fairy Tales*, and found that transformations across close ontological boundaries (e.g., humans turning into animals) were overrepresented relative to transformations across more distant ones (e.g., humans turning into inanimate objects), implying that the former are deemed more plausible than the latter.

Building on this work, Griffiths (2015) created materials in which he varied the direction of the transformation (e.g.,

humans turning into plants vs. plants turning into humans) and asked people which version was more interesting. He found that they systematically preferred transformations resulting in greater animacy or intelligence (plants turning into humans) to those that resulted in the opposite (humans turning into plants). These results imply that, when contemplating magical transformations, we not only prefer close ontological changes to distant ones but also prefer ontological “promotions” to ontological “demotions.”

It thus appears that we honor ontological commitments even when entertaining transformations that explicitly violate those commitments. Here, we explore the possibility that such findings are symptomatic of a more general phenomenon: that we honor causal principles of all types (ontological or otherwise) when entertaining magical events of all types (metamorphoses or otherwise). The intuition that rocks are easier to levitate than starships, for instance, cannot be explained in terms of the distance between ontological boundaries or the directionality of an ontological transformation. Rocks and starships are both inanimate objects, and the violation at hand is not a change in identity (from one object to another) but a change in location. Likewise, the intuition that a potion for forgetfulness is easier to brew than a potion for endurance has no ontology-based explanation. Forgetfulness and endurance are both mental states, and the violation at hand is not a change in identity but a change in the means by which those states are attained.

In short, the phenomenon of the “plausible impossible” has been studied in specific instantiations—religious concepts, folktales, magic tricks—but has not been studied as a phenomenon unto itself. Here, we do so by eliciting people’s intuitions about the plausibility of a wide range of causal violations, including changes in ontological properties or ontological identity but not limited to those changes. We also employ multiple tasks (forced-choice judgments, Likert-type ratings, free response) and multiple domains (physics, biology, psychology).

The cover story for all studies was a highly popular fictional context: the wizarding world of Harry Potter. Participants were asked to review the curriculum for Hogwarts School of Witchcraft and Wizardry (Harry’s school), and to evaluate the difficulty of various spells on the Hogwarts’ curriculum. The spells were constructed in pairs such that both spells violated the same deep-seated causal principle—and were thus impossible for the same reason—but differed in surface-level properties relevant to a subsidiary principle. For instance, the spells “making a basketball float in the air” and “making a bowling ball float in the air” both violate the principle that unsupported objects fall, but they differ in relation to the subsidiary principle that heavier objects are more difficult to lift (by hand) than lighter ones. Of interest was whether participants’ responses would honor causal principles not explicitly suspended in the spell descriptions but dependent on the suspended principles

nonetheless. In all studies, we found that participants relied on causal-explanatory considerations relevant to the spells’ outcomes but precluded by the spells’ causal structure nonetheless.

Study 1

Method

The participants in Studies 1 through 4 were undergraduates at Occidental College recruited from introductory psychology courses and compensated with extra credit. They completed the study in the form of an online questionnaire.

Sixty-four undergraduates participated in Study 1. They were shown 18 pairs of spells, ostensibly culled from the Hogwarts’ curriculum, and asked to determine which spell in each pair would be more difficult to learn. Six pairs of spells involved physical principles (e.g., a spell for making a bush invisible vs. a spell for making a tree invisible); six involved biological principles (e.g., a spell for growing an extra toe vs. a spell for growing an extra eye); and six involved psychological principles (e.g., a spell for making someone smile vs. a spell for making someone laugh). All 18 pairs can be found in Table 1.

Within each pair, the spells were designed to differ as minimally as possible. Basketballs and bowling balls, for instance, differ in weight but not size or shape (attributes that might also influence an object’s liftability). It should also be noted that the outcome of each spell was not always impossible but would be impossible to produce instantaneously or extemporaneously. For example, “making someone smile” and “making someone laugh” are not impossible, in principle, but are impossible to do instantaneously, without some kind of mediating stimulus.

Participants received the 18 pairs of spells in one of two random orders. Half were asked to indicate which spell would be more difficult to learn (Study 1A), and half were asked to do the same but were given the option of indicating that both spells were “equally difficult” to learn (Study 1B). By including an “equally difficult” option, Study 1B was a more stringent test of participants’ willingness to evaluate the plausibility of impossible events, as participants could opt out of that evaluation if they saw no basis for discriminating the two options.

To analyze the data (in Study 1, as well as in Studies 2 and 5), we compared participants’ selection of the more “extreme” version of each spell to chance-level responding, defined as the observed number of selections relative to the expected number. We did not compare participants’ selections across domains, as this comparison would be confounded with item effects. We expected participants to honor logically irrelevant causal principles in their judgments of spell difficulty across

Table 1 The six pairs of spells in each domain, grouped by the irrelevant causal constraint they embody

Domain	Causal constraint	Spell
Physics	Object size	Making a (bush, tree) invisible
	Object weight	Making a (basketball, bowling ball) float in the air
	Object shape	Turning a broom into a (shovel, bucket)
	Object complexity	Shrinking a (chair, computer) to half its size
	Object density	Walking through a wall made of (wood, stone)
	Object value	Turning a lump of coal into a lump of (silver, gold)
Biology	Evolutionary similarity	Turning a person into a (monkey, pig)
	Developmental similarity	Turning an adult back into a (teenager, child)
	Ailment severity	Curing a person's (hiccups, arthritis)
	Organ size	Mending a broken (finger, arm)
	Organ complexity	Growing an extra (toe, eye)
	Organ plasticity	Making a person's (hair, teeth) grow longer
Psychology	Knowledge entrenchment	Making a person forget his own (phone number, name)
	Knowledge complexity	Teaching a monkey to do (arithmetic, calculus)
	Skill difficulty	Teaching a cow how to (skip, tap dance)
	Affect intensity	Making someone (smile, laugh)
	Trait stability	Increasing a person's (memory, intelligence)
	Language comprehension	Teaching a person to (read, speak) a foreign language

all domains, but we had no expectation as to whether those judgments would differ by domain, nor did we design our materials to test for such differences.

Results

In Study 1A, participants' judgments of spell difficulty aligned with the spells' implicit causal ordering 85% of the time for physical spells, 86% of the time for biological spells, and 82% of the time for psychological spells (see Table 2). All three percentages were significantly greater than expected by chance (50%, physics: $t(31) = 10.88$, $p < .001$; biology: $t(31) = 14.04$, $p < .001$; psychology: $t(31) = 10.86$, $p < .001$). In addition, most participants (88%) demonstrated the anticipated effect for a significant number of spell pairs (13 or more, binomial probability $< .05$), and all spell pairs elicited the anticipated effect for a significant number of participants (21 or more, binomial probability $< .05$).

In Study 1B, participants' judgments of spell difficulty aligned with the spells' implicit causal ordering 57% of the time for physical spells, 72% of the time for biological spells, and 66% of the time for psychological spells. These percentages are lower than those from Study 1A, but participants were given three options rather than two. Accordingly, all percentages were significantly greater than expected by chance (33%, physics: $t(31) = 5.49$, $p < .001$; biology: $t(31) = 11.67$, $p < .001$; psychology: $t(31) = 9.38$, $p < .001$). Likewise, most participants (72%) demonstrated the

anticipated effect for a significant number of spell pairs (10 or more, binomial probability $< .05$), and most spell pairs (78%) elicited the anticipated effect for a significant number of participants (15 or more, binomial probability $< .05$).

Discussion

Participants demonstrated a robust sensitivity to the spells' implicit causal ordering, basing their judgments of spell difficulty on logically irrelevant causal principles in all three domains. In other words, participants revealed a consistent, domain-general intuition that some impossible events are "more impossible" than others. Note that what we label here as "illogical" is the intuition, not the participant. Participants responded sensibly in light of the task demands; the responses themselves, however, defy a normative interpretation of the notion of impossibility (as discussed above). That said, a concern with Study 1 is that participants' responses were purely an artifact of the task. The task highlighted a single difference between each pair of spells, and participants may have heeded that difference only because it was highlighted for them. One point against this interpretation is that participants continued to heed that difference even when given the option of indicating that both spells would be equally difficult (Study 1B). Nevertheless, we attempted to provide stronger evidence of causality-based reasoning in Study 2 by asking participants to explain their judgments.

Table 2 The proportion of adults (Studies 1A, 1B, and 2) and children (Study 5) who honored an irrelevant causal constraint in their judgments of spell difficulty. Proportions greater than that expected by chance (50% for Studies 1A and 5, 33% for Studies 1B and 2) are marked with asterisks

Domain	Causal constraint	Study 1A	Study 1B	Study 2	Study 5
Physics	Object size	.91*	.53*	.47*	.69*
	Object weight	.81*	.84*	.69*	.78*
	Object shape	.91*	.47*	.59*	.44
	Object complexity	.91*	.31	.50*	.53
	Object density	.81*	.84*	.47*	.69*
	Object value	.75*	.22	.19	.41
Biology	Evolutionary similarity	.88*	.50*	.53*	.44
	Developmental similarity	.84*	.72*	.63*	.53
	Ailment severity	.91*	.78*	.78*	.69*
	Organ size	.66*	.47*	.28	.63
	Organ complexity	.94*	.69*	.63*	.72*
Psychology	Organ plasticity	.97*	.81*	.94*	.69*
	Knowledge entrenchment	.81*	.84*	.81*	.50
	Knowledge complexity	1.00*	.91*	.84*	.75*
	Skill difficulty	.69*	.69*	.56*	.78*
	Affect intensity	.88*	.66*	.63*	.34
	Trait stability	.78*	.13	.47*	.44
	Language comprehension	.78*	.25	.47*	.66*

Study 2

The goal of Study 2 was to replicate study Study 1B, while also verifying that participants based their judgments on explanatory considerations precluded by the causal violation at hand. Accordingly, we asked participants to explain their judgment, and we coded those explanations for the causal constraints listed in Table 1.

Method

The participants were 32 undergraduates. For each pair of spells listed in Table 1, they judged whether the first spell would be more difficult to learn, the second would be more difficult to learn, or the two spells would be equally difficult to learn. The ordering of the pairs was randomized, as was the ordering of the spells within each pair. Participants were asked to provide an explanation for all judgments, and we coded those explanations for references to the target causal constraint.

For instance, with respect to the spells “turning a person into a monkey” and “turning a person into a pig,” we coded participants’ explanations for references to the evolutionary or genetic similarity between humans, monkeys, and pigs. The explanation “pigs have less DNA in common with humans than do monkeys” was coded as meeting that criterion, whereas the explanation “both spells involve transforming a person into an animal [and] there is nothing inherent about either animal that would cause either option to be more difficult”

was not. With respect to the spells “teaching a cow how to skip” and “teaching a cow how to tap dance,” we coded participants’ explanations for references to the physical difficulty of skipping versus tap dancing. The explanation “tap dancing involves rhythmic variation and a sequence of changing motions, whereas skipping is really one repeated motor program” was coded as meeting that criterion, whereas the explanation “cows don’t do either action naturally, so both are forced contortion of the cow’s body” was not.

Coding was accomplished by two judges who coded all 576 explanations independently. They agreed on 88% of their codes (Cohen’s kappa = .76); all disagreements were resolved through discussion.

Results

Replicating Study 1B, participants’ judgments of spell difficulty aligned with the spells’ implicit causal ordering more often than expected by chance (33%): 48% of the time for physical spells ($t(31) = 3.18, p < .001$), 63% of the time for biological spells ($t(31) = 7.42, p < .01$), and 58% of the time for psychological spells ($t(31) = 6.30, p < .001$). The proportion of participants who demonstrated the anticipated effect is broken down by spell in Table 2. Most participants (59%) demonstrated the anticipated effect for a significant number of pairs (10 or more, binomial probability $< .05$), and most pairs (89%) elicited the anticipated effect for a significant number of participants (15 or more, binomial probability $< .05$).

With respect to explanations, participants mentioned the irrelevant causal constraint listed in Table 1 more often than not: 59% of the time for physical spells, 66% for biological spells, and 64% for psychological spells. Critically, participants' mentioning of the irrelevant causal constraint was correlated with their evaluation of spell difficulty. When participants judged the more causally extreme spell as more difficult, they cited the target constraint in their explanation 96% of the time. When they did not make the predicted judgment—i.e., they judged the other spell as more difficult or judged both spells as equally difficult—they cited the target constraint only 19% of the time. A paired t test, comparing the proportion of causality-based judgments followed by a causality-based explanation to the proportion of other judgments followed by a causality-based explanation, confirmed that this difference was significant ($t(31) = 20.78, p < .001$).

Discussion

The findings of Study 2 validate the task and the materials used in Study 1. When participants judged the more causally extreme version of each spell as more difficult to learn, they did so by consulting the causal constraints we had intended them to consult—the constraints listed in Table 1. And when participants judged the spells as equally difficult or difficult in the opposite direction, they did not reference those same constraints, indicating that graded notions of impossibility derive from specific explanatory considerations and are not entertained in the absence of those considerations.

It is an open question why some participants entertained the target explanatory considerations and others did not, and we return to that question in the General Discussion. First, however, we present evidence that the effects documented in Studies 1 and 2 are not specific to the format of the task—i.e., that they can be obtained not just when participants evaluate the plausibility of each spell in the context of a closely-matched alternative but also when they evaluate the plausibility of each spell in isolation.

Study 3

In Study 3, we sought to replicate the findings of Studies 1 and 2 using Likert-type ratings rather than pairwise comparisons. We collected those ratings in one of three ways: by asking participants to rate the two versions of each spell back-to-back (Study 3A), by asking participants to rate the two versions of each spell intermixed with other spells (Study 3B), and by asking participants to rate either the more-extreme version of each spell or the less-extreme version but not both (Study 3C). The three conditions provide increasingly stronger evidence that plausibility judgments are not context-

specific. The similarity between the two versions of each spell was less apparent in Study 3B (where the spells were presented at random) than Study 3A (where the two versions of each spell were presented in tandem), and that similarity was not at all apparent in Study 3C (where only a single version of each spell was presented).

Method

The participants were 128 undergraduates; 32 participated in Study 3A, 32 in Study 3B, and 64 in Study 3C. They rated the difficulty of each spell on a seven-point scale, from “slightly difficult” to “extremely difficult.” Participants were encouraged to review the entire list of spells before making their ratings so as not to become anchored on a rating that was too low or too high early in the questionnaire. Studies 3A and 3B employed within-participants design, whereas Study 3C employed a between-participants design (hence, the inclusion of twice as many participants). In this study, our measure of whether participants' judgments reflected causality-based reasoning was the difference in ratings between the two versions of each spell within the same pair. We did not compare participants' ratings to chance, as it was unclear what measure would constitute chance-level responding in this task.

Results

In Study 3A, participants' difficulty ratings for the more-extreme spells were significantly higher than their difficulty ratings for the less-extreme spells in all three domains (physics: $M = 4.1$ vs. $3.6, t(31) = 4.95, p < .001$; biology: $M = 4.7$ vs. $4.0, t(31) = 7.56, p < .001$; psychology: $M = 4.2$ vs. $3.6, t(31) = 7.21, p < .001$). These differences remained significant even when the two versions of each spell were randomly intermixed with other spells (Study 3B), making their comparison less explicit (physics: $M = 4.3$ vs. $3.8, t(31) = 5.75, p < .001$; biology: $M = 5.3$ vs. $4.4, t(31) = 11.29, p < .001$; psychology: $M = 4.5$ vs. $3.8, t(31) = 7.65, p < .001$). And they remained significant even when participants rated the less-extreme version of each spell or the more-extreme version but not both (Study 3C), as assessed with independent-samples t tests rather than paired t tests (physics: $M = 4.1$ vs. $3.4, t(62) = 3.02, p < .01$; biology: $M = 5.1$ vs. $3.8, t(62) = 6.64, p < .001$; psychology: $M = 4.4$ vs. $3.2, t(62) = 5.30, p < .001$).

These effects were observed at the item level as well (see Table 3). In Study 3A, participants rated the more-extreme spell in each pair as significantly more difficult than the less-extreme spell for 94% of pairs (17 of 18). This was true for 78% of pairs (14 of 18) in Study 3B and for 83% of pairs (15 of 18) in Study 3C.

Table 3 Mean differences in difficulty ratings between more and less extreme versions of the same spell. Differences reliably greater than zero ($t(31) > 2.0$, $p < 0.05$) are marked with asterisks

Domain	Causal constraint	Study 3A	Study 3B	Study 3C
Physics	Object size	0.56*	0.47*	0.81*
	Object weight	0.72*	0.81*	0.94*
	Object shape	0.28*	0.38*	0.50
	Object complexity	0.44*	0.25	0.50
	Object density	0.53*	0.53*	0.78*
	Object value	0.22*	0.75*	0.72*
Biology	Evolutionary similarity	0.28*	-0.16	0.38
	Developmental similarity	0.25	0.41*	0.75*
	Ailment severity	2.13*	2.50*	3.00*
	Organ size	0.50*	0.66*	0.91*
	Organ complexity	0.66*	1.16*	0.94*
	Organ plasticity	0.81*	1.09*	2.13*
Psychology	Knowledge entrenchment	0.56*	1.56*	1.63*
	Knowledge complexity	0.94*	0.81*	1.47*
	Skill difficulty	0.69*	0.59*	0.78*
	Affect intensity	0.25*	0.22	0.78*
	Trait stability	0.59*	0.34	1.06*
	Language comprehension	0.56*	0.47*	1.38*

Discussion

Participants' intuitions about spell difficulty honored logically irrelevant causal principles even when those intuitions were assessed spell-by-spell, rather than in comparison to its causal match. These findings further confirm that the findings of Studies 1 and 2 were not an artifact of the task. Still, all findings were obtained using experimenter-generated spells. In Study 4, we expanded the scope of our investigation by asking participants to generate their own spells.

Study 4

In Study 4 we sought to replicate the findings of Studies 1-3 using an open-ended version of the spell-difficulty task in which participants generated their own examples of easy spells and difficult spells rather than assess the difficulty of examples given to them.

Method

The participants were 133 undergraduates. They were told that they were to revise the Hogwarts' curriculum and that the curriculum was divided into eight subjects: levitation, divination, transmutation, teleportation, enchantment, necromancy, conjuring, healing, and hexing. These subjects were selected to exemplify a mixture of physically impossible events (levitation, transmutation,

teleportation, conjuring), biologically impossible events (necromancy, healing), and psychologically impossible events (divination, enchantment, hexing). The participants' task was to devise one introductory spell, one intermediate spell, and one advanced spell per subject.

For each type of spell, participants were given a frame and asked to generate three examples of a category that would complete the frame. Levitation spells, for instance, were elicited with the frame "making a [object] float in the air." The items were coded by two judges blind to their source. The judges ordered those items along a causal dimension not specified in the spell frame but still predicted to have influenced participants' responses. For example, participants' responses to the frame "bringing a dead [animal] back to life" were ordered by size, from smallest to largest. All spell frames and coding criteria are listed in Table 4.

Sample responses for each type of spell are presented in Table 5. At the coding stage, participants' responses were randomly scrambled (within the triad), and judges were given no information as to who generated those responses or for what purpose. Both judges coded all 1,064 response triads independently. They assigned an order to each triad using only the coding criteria listed in Table 4 (the "causal constraint"). They agreed on 84% of their orderings (Cohen's kappa = .81), and disagreements were resolved through discussion. To assess whether participants honored the unmentioned causal constraint when generating spells, we compared participants' orderings to the judges' orderings.

Table 4 The frames presented to participants in Study 4, and the causal constraints that participants were expected to honor when using those frames to generate spells of varying difficulty

Spell type	Frame	Requested item	Causal constraint
Levitation	Making a ___ float in the air	Object	Weight
Divination	Predicting when the next ___ will occur	Event	Probability
Transmutation	Turning ___ into gold	Material	Density
Teleportation	Teleporting a package from LA to ___	Location	Proximity to LA
Enchantment	Enchanting a person to like ___	Food	Disgustingness
Necromancy	Bringing a dead ___ back to life	Animal	Size
Conjuring	Conjuring a ___ out of thin air	Object	Size
Healing	Making a potion that cures ___	Disease/ailment	Severity
Hexing	Hexing a person to lose their ___	Possession	Personal value

Results

Given that there are six ways to order three items (ABC, ACB, BAC, BCA, CAB, CBA), participants' orderings should have matched the judges' orderings 17% of the time by chance alone. In actuality, they matched 90% of the time for levitation spells, 62% for divination spells, 49% for transmutation spells, 96% for teleportation spells, 44% for enchantment spells, 88% for necromancy spells, 71% for conjuring spells,

Table 5 Examples of spell triads (introductory, intermediate, advanced) generated in Study 4

Spell type	Examples		
Levitation	Pencil	Spoon	Shoe
	Book	Chair	Lamp
	Person	Table	Truck
Divination	Hurricane	Train arrival	Snowfall
	Plane crash	Thunderstorm	Pop quiz
	Ice age	Nuclear war	Celebrity death
Transmutation	Yarn	Silver	Rock
	Sand	Wood	Coal
	Water	Meat	Flour
Teleportation	Malibu	Sacramento	San Francisco
	Pittsburgh	Canada	Paris
	Moscow	Antarctica	The moon
Enchantment	Baby food	Sour milk	Salt
	Bird seed	Maggots	Garbage
	Cat food	Human flesh	Poison
Necromancy	Fly	Slug	Rat
	Mouse	Goldfish	Tiger
	Elephant	Kitten	Dinosaur
Conjuring	Burrito	Comb	Flower
	Horse	Keyboard	Notebook
	House	Museum	Computer
Healing	Mononucleosis	Toothache	Hiccups
	Schizophrenia	Headache	Pneumonia
	Cancer	Broken heart	Leprosy
Hexing	Hair	Hat	Keys
	Pet	Wallet	Home
	Limb	Car	Mind

79% for healing spells, and 72% for hexing spells. All percentages were significantly greater than that expected by chance (17%, binomial probability < .001).

Discussion

Generating spells of varying difficulty yielded the same result as rating the difficulty of pre-generated spells: evaluations of spell difficulty were driven by causal considerations that should not have applied to those spells. The effects cut across tasks and across domains, and, in one final study, we show that they cut across development as well.

Study 5

In Study 5, we assessed whether preschool-aged children view magical events through a causal lens like adults do. We have argued that adults view magical events in this way because they are unable to fully suspend the causal knowledge precluded by those events, but it is also possible that adults' views of magic have been shaped by extensive exposure to magical narratives in books, television shows, and films. Here, we assess whether a group of participants who have had significantly less exposure to such narratives (preschoolers) hold the same views.

Method

The participants were 32 children between the ages of 4 and 5 years (M age = 4 years 8 months). Children of this age are adept at distinguishing possible events from impossible events, labeling only impossible events as "magic" (Phelps & Woolley, 1994; Rosengren, Kalish, Hickling, & Gelman, 1994), but it is unclear whether they believe some kinds of impossible events are more impossible than others.

The children were recruited from a preschool in Southern California and tested on site. They were shown the same 18

pairs of spells used in Studies 1–3 with two exceptions: “teaching a person to (read, speak) a foreign language” was replaced with “teaching a baby to (speak, read)” and “teaching a monkey to do (arithmetic, calculus)” was replaced with “teaching a baby to (count, add and subtract),” so as to render those spells more child-friendly. Each spell was visually depicted on a laminated index card. Participants were presented the cards in pairs and asked to sort them into each of two containers. One container was labeled with a picture of Harry Potter and the other was labeled with a picture of Professor Dumbledore. Participants were instructed to use the containers as follows:

“Have you ever heard of Harry Potter? Harry Potter is a boy in a story who has magical powers and goes to a school called Hogwarts School of Witchcraft and Wizardry, where he learns how to cast different spells. Here are two buckets with pictures of wizards on them. This bucket shows Harry who is a young wizard and not very good at magic yet, and this one shows Professor Dumbledore, one of Harry’s teachers at the school. He is an older wizard who is much better at magic. Over here I have a pile of cards with different spells on them. Some of the spells are easy, which means that even young wizards can do them, but some of the spells are hard, which means that only older wizards can do them. Can you help me figure out which spells are easy and which spells are hard? You will put the easy spells in this bucket that shows Harry, and you will put the hard spells in this bucket that shows Dumbledore.”

The experimenter verified that participants understood the instructions by asking them to point to the bucket for easy spells and then point to the bucket for hard spells. The experimenter then presented participants with each pair of spells in one of two random orders. Feedback was not provided during the sorting process, though participants who struggled with the task were assured that there are no correct answers and encouraged to try their best.

Results

Participants sorted the spells in accordance with their implicit causal ordering 62% of the time for physical spells, 58% for biological spells, and 58% for psychological spells. All percentages were significantly greater than that expected by chance (50%, physics: $t(31) = 2.96, p < .01$; biology: $t(31) = 2.22, p < .05$; psychology: $t(31) = 1.96, p < .05$). At the item level, participants showed the anticipated effect for at least half of the spells within each domain (see Table 2). The effect, pooled across items, was weaker for children than it was for adults (Cohen’s $d = 0.74$ vs. 2.70), but it strengthened with the child’s age. The older a child was (in months), the more likely

he/she sorted the less-extreme spell into the Harry-Potter container and the more-extreme spell into the Dumbledore container ($r(30) = 0.44, p < .05$).

Discussion

By age 5 years, children appear to hold causality-based intuitions about the plausibility of impossible events, assessing how difficult it would be to bring about a magical event on the basis of causal principles that should not logically pertain to that event. These intuitions are in place not only before children have received formal instruction on causal principles but also before they have received much exposure to magical narratives in fiction. Certainly, children of this age have had *some* exposure—all of them were familiar with the story of Harry Potter—but it is unlikely that such exposure could account for our findings, as our spells were novel and our task was novel as well.

One unexpected finding was that children’s sensitivity to the causal structure of our stimuli increased with age, from the beginning of age 4 to the end of age 5 years. This is a period during which children develop sensitivity to a related phenomena: the difference between a magic trick and “real” magic (Chandler & Lalonde, 1994; Rosengren & Hickling, 1994; Subbotsky, 2004). Whereas 5-year-olds tend to recognize that magic tricks are just illusions, brought about by sleight of hand or special apparatuses, 4-year-olds tend to claim that these tricks are genuinely impossible and require some kind of supernatural power (“magic”). This developmental change is likely driven by an increased awareness of the causal constraints on real-world events, and our task may have tapped into the same change, albeit less directly.

That said, the mere fact that children are willing and able to judge the plausibility of impossible events tells us that graded notions of impossibility emerge early, even in the absence of mature causal knowledge, and that such notions probably do not require effortful causal inference. It may actually be the reverse: *suppressing* notions of graded possibility may require effortful causal inference. Case in point, the creators of Star Wars did not have to convince their audience that starships are easier to levitate than rocks. We accept that quirk in the Star Wars plot without it needing to be mentioned, which suggests that convincing us otherwise—i.e., convincing us that levitation is *not* constrained by weight—may require some additional effort or elaboration.

General discussion

In five studies, we found that participants used real-world causal knowledge to interpret events that fall beyond the scope of such knowledge: magical events. We obtained these findings regardless of whether our task involved pairwise

comparisons, Likert-type ratings, or open-ended responses; regardless of whether our stimuli involved physical principles, biological principles, or psychological principles; and regardless of whether our participants were preschool-aged children or college-educated adults. As a whole, these findings imply that causal-explanatory knowledge is not easily suspended. Events like “walking through a wall made of wood” and “walking through a wall made of stone” have no explanation—they are deemed impossible by the laws of nature—but we still rely on explanatory considerations like hardness and density to interpret those events, perceiving the latter as more difficult than the former.

Our findings accord with a growing body of literature demonstrating that products of the imagination—fiction, fantasy, pretense, or superstition—are structured by causal constraints on reality (Harris, 2000; Legare & Souza, 2012; Lindeman & Aamio, 2007). For instance, it has been shown that we prefer stories that violate contingent truths (that Washington DC is the capital of the US) or conventional truths (that picking your nose is rude) to stories that violate mathematical truths (that two plus two equals four) or moral truths (that murder is wrong) (Barnes & Black, 2016; Gendler, 2000; Weisberg & Goodstein, 2009). We prefer extraterrestrial creatures that honor the biological properties of terrestrial creatures (e.g., bilateral symmetry, dual appendages, cephalization) to those that do not (Smith, Ward, & Schumacher, 1993; Ward, 1994). And we prefer divine agents that honor the properties of human psychology (e.g., that knowledge depends on perception, that perception depends on attention, that attention depends on interest) to agents that violate such expectations (Heiphetz, Lane, Waytz, & Young, 2015; Lane, Evans, Brink, & Wellman, 2016; Purzycki, 2013).

Our findings also accord with the emerging consensus that causal constraints on imagination operate even in young children (Buchsbaum, Bridgers, Weisberg, & Gopnik, 2012; Cook & Sobel, 2011; Weisberg & Gopnik, 2013). Two-year-olds, for instance, recognize that, when pretending to pour tea from an empty teapot into an empty cup, tea has been transferred from the teapot to the cup and will spill out of the cup if the cup is overturned (Harris & Kavanaugh, 1993; Harris, Kavanaugh, & Meredith, 1994). Four-year-olds can distinguish the properties of one pretend world (the world of Batman) from another (the world of SpongeBob) and keep those properties separate when pretending to interact with those worlds (Skolnick Weisberg & Bloom, 2009) or drawing inferences about what is likely to be true in those worlds (Skolnick & Bloom, 2006). And 4-year-olds can also distinguish fictional stories that resemble reality (stories about finding ladybugs and climbing trees) from those that do not (stories about finding fairies and talking to trees), privileging the former as a more secure source of information (Richert & Smith, 2011; Walker, Gopnik, & Ganea, 2015).

Each of these imaginative activities can be understood as selective rule-breaking: we break some of the rules of reality

but conform to others. Why do we break some rules but not all? Are we just conservative in our rule-breaking? Or does our rule-breaking follow discernible patterns, unfolding along some lines but not others? We believe the notion of plausibility helps make sense of this behavior. In five studies, we have shown that people find some impossible events more plausible than others, and their basis for making these discriminations appears to be causal knowledge *related to* the event but not explicitly *violated by* the event.

Researchers have implicitly acknowledged the role of plausibility in imaginative activity when claiming that imagination is “structured” (e.g., Harris, 2000; Ward, 1994), but explicit recognition is needed to advance our understanding of the structure itself. Take, for example, the act of pretending to pour a cup of tea. Our pretense follows one sequence of events (tea is transferred from the teapot to the cup) but could have followed others (the tea misses the cup, the tea evaporates, the cup falls off the table, the cup shatters). Only the first sequence seems plausible, but the source of that intuition is not obvious. Many have pointed to causal knowledge (e.g., Cook & Sobel, 2011; Phelps & Woolley, 1994; Rosengren et al., 1994; Walker et al., 2015; Weisberg & Gopnik, 2013), but the means by which causal knowledge is activated and applied to events that explicitly violate that knowledge is not yet understood.

Relations to causal knowledge

What aspects of causal knowledge might give rise to intuitions about the plausibility of impossible events? One proposal is that, when entertaining an impossible event, we bring to mind possible events that closely resemble the target event in content or structure and then assess the similarity between the possible event and the target event. The greater the similarity, the more plausible we judge the target event. For instance, when contemplating a spell that would reverse aging, we bring to mind instances of normal aging, and realize that, if the age-related differences between children and adults are greater than those between teenagers and adults, then turning an adult back into a child should be “more impossible” than turning an adult back into a teenager. On this account, the role of causal knowledge is implicit—implicit in the real-world analogues chosen for comparison and the similarity metric used to make that comparison—which accords with the finding from Study 5 that children who lack explicit knowledge of causal principles (by and large) can still assess the plausibility of events that violate those principles.

A limitation of this account, however, is that it is hard to apply to impossible events with no obvious real-world analogue. For example, it is not clear what real-world analogue is relevant to deciding whether it would be easier to grow an extra toe or an extra eye, as humans cannot grow any new

appendages. The same concern holds for deciding whether it would be easier to walk through a wall made of wood or a wall made of stone, as humans cannot walk through any solid material. Real-world analogues are still accessible—e.g., growing something other than an appendage (hair, skin, nails) or walking through something other than a solid (water, mud, fog)—but the conceptual distance between the analogue and the target event may be too great to support robust intuitions. It is not obvious, for instance, that walking through water is more similar to walking through wood than walking through stone, as wood and stone are more similar to each other than either is to water.

An alternative explanation for why we hold graded notions of impossibility lies in the fact that our causal knowledge is highly *interconnected*. Much of our causal knowledge is organized in abstract, coherent networks of directed causal relations, both for children (Gopnik, Glymour, Sobel, Schulz, Kushnir, & Danks, 2004; Sobel, Tenenbaum, & Gopnik, 2004) and for adults (Sloman, 2005; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003). These networks are derived from patterns of covariation between putative causes and putative effects and can be used to generate explanations, predictions, interventions, and counterfactuals. The property of these networks most relevant to graded notions of impossibility is that they involve multiple relations among multiple concepts. A network that represents our knowledge of physical objects, for instance, might have links between *mass*, *weight*, *volume*, *density*, *buoyancy*, *solidity*, *cohesion*, *contact*, and *support*. Given a representation of this nature, we suspect that severing one connection in the network would leave other connections intact. Severing the connection between *contact* and *support*, for instance, would leave the connection between *weight* and *support* intact, thus yielding the intuition that levitating a starship would be more difficult than levitating a stone.

In other words, what makes an impossible event *plausibly* impossible may be its conformity to a larger network of causal expectations of which the violated expectation is just one constituent. No impossible events are plausible in the sense of being probable, but they may be plausible in the sense of corresponding to existing explanatory structures. Our findings are consistent with this proposal, but further studies are needed to test it directly—studies that measure (or manipulate) the causal knowledge relevant to a particular causal violation and assess whether that knowledge is critical to establishing graded notions of impossibility. For instance, a child who has yet to discern the relation between density and buoyancy should have no expectations regarding the role of density in magical events pertaining to buoyancy (e.g., a spell for making clay float in water vs. a spell for making lead float in water), whereas a child who has discerned the relation between density and buoyancy should hold such expectations.

Implications for modal cognition

Our findings also highlight a quirk in how we reason about physical possibility. Reasoning about physical possibility is a form of *modal* cognition. Reasoning about moral permissibility is also a form of modal cognition, and the two forms of reasoning share a number of similarities (Perkins, 1983; Sinnott-Armstrong, Raffman, & Asher, 1995). Both are concerned with what is normatively true about the world rather than what is descriptively true; both entail the application of preexisting commitments (moral rules or physical laws) to hypothetical situations; and both can be expressed with the same modal verbs (“can,” “could,” “might,” “must,” “should”). Indeed, studies that have explored the two forms of reasoning in conjunction have found parallels between them—developmental parallels in the types of events judged impossible or impermissible (Browne & Woolley, 2004; Chernyak, Kushnir, Sullivan, & Wang, 2013; Kalish, 1998; Phillips & Bloom, 2014) and cognitive parallels in the considerations underlying those judgments (Shtulman & Tong, 2013).

These similarities withstanding, there is a key difference between the two forms of reasoning: impermissibility comes in degrees but impossibility does not, at least not on the surface. We regularly rank some impermissible actions (e.g., murder) as “more wrong” than others (e.g., stealing), and we describe impermissible events with hedges like “a little wrong,” “sorta wrong,” or “kinda wrong.” But we rarely acknowledge that impossible events could be rank ordered, and we rarely use hedges when describing them. Indeed, the phrases “a little impossible,” “sorta impossible,” or “kinda impossible” generate a quarter as many Google hits as generated by their moral counterparts. Thus, what is odd about graded notions of impossibility is not just that we apply causal expectations to causally anomalous events but that we do so without any corresponding recognition in our language or our everyday cognition.

Conclusion

In five studies, we have shown that children and adults have consistent expectations about the plausibility of magical events. Suspending disbelief in one causal violation does not lead to widespread suspension of disbelief; other causal expectations are maintained, even if those expectations are no longer appropriate. When Walt Disney specified that his feature-length films had to be plausibly impossible, he pinpointed an intuition that is consistently honored in fiction but not well understood from a psychological point of view, even today. Future research on graded notions of impossibility promises to shed light on the structure of causal cognition and the structure of imagination more generally.

Author Note This research was supported by National Science Foundation grant DRL-0953384 and a James S. McDonnell Foundation Understanding Human Cognition Scholar Award. The authors would like to thank Max Rattner and Rachel Yoo for their assistance with data analysis.

References

- Atran, S., & Norenzayan, A. (2004). Religion's evolutionary landscape: Counterintuition, commitment, compassion, communion. *Behavioral and Brain Sciences*, *27*, 713–730.
- Banerjee, K., Haque, O. S., & Spelke, E. S. (2013). Melting lizards and crying mailboxes: Children's preferential recall of minimally counterintuitive concepts. *Cognitive Science*, *37*, 1251–1289.
- Barnes, J., & Black, J. (2016, in press). Impossible or improbable: The difficulty of imagining morally deviant worlds. *Imagination, Cognition, and Personality*.
- Barrett, J. L., & Nyhof, M. A. (2001). Spreading non-natural concepts: The role of intuitive conceptual structures in memory and transmission of cultural materials. *Journal of Cognition and Culture*, *1*, 69–100.
- Boyer, P. (1999). Cognitive tracks of cultural inheritance: How evolved intuitive ontology governs cultural transmission. *American Anthropologist*, *100*, 876–889.
- Boyer, P., & Ramble, C. (2001). Cognitive templates for religious concepts: Cross-cultural evidence for recall of counter-intuitive representations. *Cognitive Science*, *25*, 535–564.
- 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.
- Buchsbaum, D., Bridgers, S., Weisberg, D. S., & Gopnik, A. (2012). The power of possibility: Causal learning, counterfactual reasoning, and pretend play. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *367*, 2202–2212.
- Chandler, M. J., & Lalonde, C. E. (1994). Surprising, magical, and miraculous turns of events: Children's reactions to violations of their early theories of mind and matter. *British Journal of Developmental Psychology*, *12*, 83–95.
- 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.
- Cook, C., & Sobel, D. M. (2011). Children's beliefs about the fantasy/reality status of hypothesized machines. *Developmental Science*, *14*, 1–8.
- Gendler, T. S. (2000). The puzzle of imaginative resistance. *The Journal of Philosophy*, *97*, 55–81.
- Gervais, W. M., & Henrich, J. (2010). The Zeus problem: Why representational content biases cannot explain faith in gods. *Journal of Cognition and Culture*, *10*, 383–389.
- Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*, 3–32.
- Griffiths, T. L. (2015). Revealing ontological commitments by magic. *Cognition*, *136*, 43–48.
- Harris, P. L. (2000). *The work of the imagination*. Malden, MA: Blackwell Publishing.
- Harris, P. L., & Kavanaugh, R. D. (1993). Young children's understanding of pretense. *Monographs of the Society for Research in Child Development*, *58*, serial no. 231.
- Harris, P. L., Kavanaugh, R. D., & Meredith, M. (1994). Young children's comprehension of pretend episodes: The integration of successive actions. *Child Development*, *65*, 16–30.
- Heiphetz, L., Lane, J. D., Waytz, A., & Young, L. L. (2015). How children and adults represent God's mind. *Cognitive Science*, *39*, 1–24.
- Kalish, C. (1998). Reasons and causes: Children's understanding of conformity to social rules and physical laws. *Child Development*, *69*, 706–720.
- Keil, F. C. (1979). *Semantic and conceptual development: An ontological perspective*. Cambridge, MA: Harvard University Press.
- Kelly, M. H., & Keil, F. C. (1985). The more things change...: Metamorphoses and conceptual structure. *Cognitive Science*, *9*, 403–416.
- Lane, A. (2006, Dec. 11). Wonderful world: What Walt Disney made. *The New Yorker*.
- Lane, J. D., Evans, E. M., Brink, K. A., & Wellman, H. M. (2016). Developing concepts of ordinary and extraordinary communication. *Developmental Psychology*, *52*, 19–30.
- Legare, C. H., & Souza, A. L. (2012). Evaluating ritual efficacy: Evidence from the supernatural. *Cognition*, *124*, 1–15.
- Lewis, D. (1978). Truth in fiction. *American Philosophical Quarterly*, *15*, 37–46.
- Lindeman, M., & Aarnio, K. (2007). Superstitious, magical, and paranormal beliefs: An integrative model. *Journal of Research in Personality*, *41*, 731–744.
- Nichols, S. (2006). Imaginative blocks and impossibility: An essay in modal psychology. In S. Nichols (Ed.), *The architecture of the imagination* (pp. 237–255). Oxford, UK: Oxford University Press.
- Norenzayan, A., Atran, S., Faulkner, J., & Schaller, M. (2006). Memory and mystery: The cultural selection of minimally counterintuitive narratives. *Cognitive Science*, *30*, 531–553.
- Perkins, M. (1983). *Modal expressions in English*. London, UK: Pinter.
- Phelps, K. E., & Woolley, J. D. (1994). The form and function of young children's magical beliefs. *Developmental Psychology*, *30*, 385–394.
- Phillips, J., & Bloom, P. (2014). *Deontic magic: Evidence for similarity in the representations of impossible and immoral events in children and adults*. Paper presented at the 40th meeting of the Society for Philosophy and Psychology, Vancouver, Canada.
- Purzycki, B. G. (2013). The minds of gods: A comparative study of supernatural agency. *Cognition*, *129*, 163–179.
- Purzycki, B. G., & Willard, A. K. (2016, in press). MCI theory: A critical discussion. *Religion, Brain, & Behavior*.
- Richert, R. A., & Smith, E. I. (2011). Preschoolers' quarantining of fantasy stories. *Child Development*, *82*, 1106–1119.
- Rosengren, K. S., & Hickling, A. K. (1994). Seeing is believing: Children's explanations of commonplace, magical, and extraordinary transformations. *Child Development*, *65*, 1605–1626.
- 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.
- Ryan, M. L. (1980). Fiction, non-factuals, and the principle of minimal departure. *Poetics*, *9*, 403–422.
- Shtulman, A. (2008). Variation in the anthropomorphization of supernatural beings and its implications for cognitive theories of religion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1123–1138.
- 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., & Lindeman, M. (2016, in press). Attributes of God: Conceptual foundations of a foundational belief. *Cognitive Science*.
- Shtulman, A., & Tong, L. (2013). Cognitive parallels between moral judgment and modal judgment. *Psychonomic Bulletin & Review*, *20*, 1327–1335.
- Sinnott-Armstrong, W., Raffman, D., & Asher, N. (1995). *Modality, morality, and belief*. Cambridge, UK: Cambridge University Press.
- Skolnick Weisberg, D., & Bloom, P. (2009). Young children separate multiple pretend worlds. *Developmental Science*, *12*, 699–705.

- Skolnick, D., & Bloom, P. (2006). What does Batman think about SpongeBob? Children's understanding of the fantasy/fantasy distinction. *Cognition, 101*, B9–B18.
- Sloman, S. (2005). *Causal models: How people think about the world and its alternatives*. Oxford, UK: Oxford University Press.
- Smith, S. M., Ward, T. B., & Schumacher, J. S. (1993). Constraining effects of examples in a creative generation task. *Memory & Cognition, 21*, 837–845.
- Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science, 28*, 303–333.
- Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science, 27*, 453–489.
- Subbotsky, E. (2004). Magical thinking in judgments of causation: Can anomalous phenomena affect ontological causal beliefs in children and adults? *British Journal of Developmental Psychology, 22*, 123–152.
- Taylor, M. (Ed.). (2013). *The Oxford handbook of the development of imagination*. Oxford, UK: Oxford University Press.
- Upal, M. A. (2011). Memory, mystery and coherence: Does the presence of 2-3 counterintuitive concepts predict cultural success of a narrative? *Journal of Cognition and Culture, 11*, 23–48.
- Walker, C. M., Gopnik, A., & Ganea, P. A. (2015). Learning to learn from stories: Children's developing sensitivity to the causal structure of fictional worlds. *Child Development, 86*, 310–318.
- Ward, T. B. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology, 27*, 1–40.
- Weisberg, D. S., & Goodstein, J. (2009). What belongs in a fictional world? *Journal of Cognition and Culture, 9*, 69–78.
- Weisberg, D. S., & Gopnik, A. (2013). Pretense, counterfactuals, and Bayesian causal models: Why what is not real really matters. *Cognitive Science, 37*, 1368–1381.