The Plausible Impossible: Causal Constraints on Magical Reasoning

Andrew Shtulman (shtulman@oxy.edu)
Caitlin Morgan (cmorgan@oxy.edu)
Department of Psychology, Occidental College
1600 Campus Road, Los Angeles, CA 90041

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 1-2 involved college undergraduates ($n = 192$), and Study 3 involved preschool-aged children ($n = 32$). In Study 1, participants saw pairs of magical events (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 2 replicated the effects of Study 1 with Likert-type ratings, and Study 3 replicated those effects in 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 inference, explanation, imagination

Introduction

The animator Walt Disney began his career making cartoons of a very different style than those he made at the end of his career. 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 one that violates an expectation about the world deemed true by necessity—an expectation that holds in all conceivable permutations of reality (Nichols, 2006; Shtulman & Carey, 2007; Shtulman, 2009). Violations of this nature are presumably black-and-white; an event either violates a "law of nature" and cannot occur in the real world or it violates no such law and can occur in the real world, 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, for instance, the Jedi master Yoda teaches the Jedi apprentice Luke Skywalker to levitate stones before he teaches him how to levitate an entire 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 is less attainable 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 hard work and attainability; 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.

What psychological principles might account for graded notions of impossibility, as exemplified by the popular fictional narratives above? Psychologist interested in the mental representation of impossible events have typically analyzed such events in terms of ontological violations. An ontology is a basic category of existence (e.g., "object," "animal," "number"), and an ontological violation is a violation of one of those categories' core properties (e.g., that objects are solid or that animals can die). Ontological violations have been shown to affect an event's memorability (Boyer, 1999). Concepts that violate one or two ontological commitments (e.g., a talking tree) are more memorable than those that violate several (e.g., a talking tree that floats in the air and never dies) or those that violate none (e.g., a flowering tree), which presumably affects how often, and how reliably, these concepts are passed from one person to another and from one generation to another.

Such memory effects have been documented both for artificial materials, constructed in the lab (Barrett & Nyhof, 2001), and for more authentic materials, like Aesop's fables (Upal, 2011) and Grimm's fairytales (Norenzayan, Atran, Faulkner, Schaller, 2006). These effects tell us that an event's memorability is determined by its conformity to our ontological commitments, but they tell us less about an event's plausibility. Indeed, memorability and plausibility

Table 1: The six pairs of spells in each domain, grouped by the (irrelevant) causal constraint they exemplify.

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

may be driven by independent factors. 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). 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 omnipotent, omniscient, omnipresent, omnibenevolent, invisible, and immortal being—violate several ontological commitments but are commonly believed to exist (Shtulman, 2008; Shtulman & Lindeman, 2016).

Here, we explore another property of conceptual representations that may account for graded notions of impossibility: the interconnectedness of causal knowledge. Recent research on the structure of causal knowledge suggests that much of this knowledge is organized in abstract, coherent networks of directed causal relations, both for children (Gopnik, Glymour, Sobel, Schulz, Kushnir, & Danks, 2004) and for adults (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.

Critically, with respect to the phenomena of interest (plausible impossibility), these networks involve multiple relations among multiple concepts. A network that represents our knowledge of physical objects, for instance, might have links between concepts as varied as 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, as done when contemplating magic, 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 the studies that follow, we test this idea across multiple tasks and multiple domains, on the prediction that what makes a magical event plausibly impossible is its conformity to a larger network of causal expectations of which the violated expectation is just one constituent. The cover story for all studies was that participants were reviewing the curriculum for Hogwarts School of Witchcraft and Wizardry, the fictional school in J. K. Rowling's "Harry Potter" series. Participants were asked to evaluate the difficulty of various spells on the Hogwarts curriculum. In all studies, we found that participants relied on explanatory considerations relevant to the spells' outcomes but precluded by the spells' causal structure nonetheless.

Study 1

Method

The participants in Studies 1 through 3 were undergraduates at Occidental College. They were recruited from introductory psychology courses and compensated for their participation with extra credit in those courses. 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. The pairs were constructed such that both spells violated the same deep-seated causal principle—and were thus impossible for the same reason—but differed from in surface-level properties relevant to a subsidiary principle. For instance, the pair "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 weight affects an object's liftability. Of interest was whether participants' responses would honor causal principles not explicitly suspended in the spell descriptions but dependent on the suspended principles nonetheless.

Six pairs of spells involved physical principles, six involved biological principles, and six involved psychological principles. 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 substantially in weight but do not differ substantially in size or shape (both of which might also influence an object's liftability). Also, it should be noted that the outcome of each spell was not always impossible, but it would be impossible to produce that outcome instantaneously or extemporaneously. For example, "making someone laugh" and "making someone smile" are not impossible in principle but are impossible to do instantaneously.

Participants received the 18 pairs of spells in one of two random orders. Half of the participants 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).

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. 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 spells (13 or more, binomial probability < .05), and all spells 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 in Study 1A, but participants were given three response options rather than two. Accordingly, all three percentages were significantly greater than expected by chance (33%): physics: $t(31) = 5.49$, $p < .001$; biology: $t(31) = 11.67$, $p < .001$; psychology: $t(31) = 9.38$, $p < .001$. Most participants (72%) demonstrated the anticipated effect for a significant number of spells (10 or more, binomial probability < .05), and most spells (78%) elicited the anticipated effect for a significant number of participants (15 or more, binomial probability < .05).

Discussion

Participants demonstrated a statistically reliable sensitivity to the spells' implicit causal ordering, judging spells that represented a greater departure from reality as "more difficult" in all three domains. Might these findings be an artifact of the task? The task was designed, after all, to highlight 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 to learn (Study 1B). Nevertheless, we attempted to provide stronger evidence of causality-based reasoning in Study 2, by asking participants to evaluate the 36 spells from Study 1 as individuals rather than as members of a pair.

Study 2

The goal of Study 2 was to replicate the findings of Study 1 using Likert-like ratings of spell difficulty 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 2A), by asking participants to rate the two versions of each spell intermixed with the other spells (Study 2B), and by asking participants to rate either the more-extreme version of each spell or the less-extreme version but not both (Study 2C).

Method

The participants in Study 2 were 128 undergraduates; 32 participated in Study 2A, 32 participated in Study 2B, and 64 participated in Study 2C. Each rated the difficulty of the spells presented in Table 1 on a seven-point scale, from "slightly difficult" to "extremely difficult." Studies 2A and 2B employed within-participants design, whereas Study 2C employed a between-participants design.

Results

In Study 2A, 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 2B), 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 2C), as assessed with independent-samples comparisons (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. In Study 2A, 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 2B and for 83% of pairs (15 of 18) in Study 2C. All three percentages were significantly greater than expected by chance (binomial probability < .05).

Discussion

Participants' intuitions about spell difficulty honored causal principles precluded by the spell's causal structure even when those intuitions were assessed spell-by-spell, rather than in comparison to its causal match. These effects have proven robust across tasks and across domains, but they have only been documented, to this point, among adults. It is possible that adults have come to perceive magic as causally-constrained through exposure to magical narratives in books, shows, or films. Study 3 explores this possibility by testing participants who have acquired significantly less exposure to such narratives: preschool-aged children.

Study 3

The question of interest in Study 3 was whether four- and five-year-old children view magical events through a causal lens, similar to adults. Children of this age are adept at distinguishing possible events from impossible events, labeling only impossible events as "magic" (see Woolley, 1997, for a review), but it is unclear how they reason about the content of such events. Do they, like adults, believe that some impossible events are more impossible than others?

Method

Thirty-two children between the ages of four to five (*M* age = 4 years, 8 months) participated in Study 3. Approximately half were male, and half were female. They were recruited from a preschool in Southern California and tested on site. Participants were shown the same 18 pairs of spells used in Studies 1 and 2. Each spell was visually depicted on an 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 on how 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 the easy spells and then point to the bucket for the 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 significantly more often than expected by chance ($M = 10.7$, $t(31) = 4.14$, $p < .001$). This effect was obtained for each domain (physics: $M = 3.7$, $t(31) = 2.96$, $p < .01$; biology, $M = 3.5$, $t(31) = 2.22$, $p < .05$; psychology, $M = 3.5$, $t(31) = 1.96$, $p < .05$) and for at least three of the six spells within those domains.

The effect was smaller for children than it was for adults (Cohen's $d = 0.74$ vs. 2.70), but it was present and increasing with the age. Indeed, 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 five, children appear to hold causality-based intuitions regarding 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 the genre of magical realism. Certainly, children of this age have had some exposure—they were, after all, 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.

That said, children's sensitivity to the causal structure of our stimuli increased over the ages sampled, from 3 years and 10 months to 5 years and 7 months. This is a period over which children are becoming increasingly sensitive to another aspect of magical phenomena: the difference between a magic trick and "real" magic (Chandler & Lalonde, 1994; Rosengren & Hickling, 1994). That is, five-year-olds tend to recognize that visual illusions, like the illusion of one object passing through another or the illusion of one object turning into another, are brought about by sleight of hand or trick apparatuses, whereas four-year-olds tend to claim that such illusions are genuine instances of magic. This developmental change, from identifying illusions with magic to identifying illusions with trickery, is presumably 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.

General Discussion

In three studies, we found that participants used real-world causal expectations to interpret events that presumably fall outside those expectations: magical events. We obtained these findings regardless of whether our task involved forced-choice comparisons, opt-out comparisons (in which ties were allowed), or Likert-type ratings; regardless of whether our stimuli involved physical principles, biological principles, or psychological principles; and regardless of whether our participants were preschoolers or college-educated adults.

As a whole, these finding imply that causal knowledge is not easily suspended. Events that explicitly violate everyday causation still elicit causal inferences, possibly because severing one link in a causal network still leaves the rest of the network intact. 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 to achieve than the former.

Our findings accord with a growing body of literature demonstrating that products of the imagination—fiction, fantasy, pretense, superstition, ritual—are structured by causal constraints on reality. For instance, it's been shown that we prefer stories in which humans transform into animals (a close ontological match) to stories in which humans transform into plants (Kelly & Keil, 1985), and we prefer stories in which animals transform into humans (an ontological promotion) to stories in which humans transform into animals (Griffiths, 2015). We prefer stories that violate contingent truths (e.g., that Washington DC is the capital of the US) to those that violate mathematical truths (e.g., that two plus two equals four; Weisberg & Goodstein, 2009). We prefer extraterrestrial creatures that honor the biological properties of terrestrial creatures—e.g., bilateral symmetry, cephalization—to those that do not (Ward, 1994). And we prefer divine agents (gods) 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 those that violate such expectations (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). For instance, two-year-old children 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 tea will spill out of the cup if the cup is overturned (Harris, Kavanaugh, & Meredith, 1994). Four-year-old children can distinguish the properties of one pretend world (the world of Batman) from another (the world of SpongeBob) and keep those properties separate when drawing inferences about what is likely to be true in those worlds (Skolnick & Bloom, 2006). And four-year-old children distinguish fictional stories that resemble reality (stories about finding ladybugs and climbing trees) to those that do not (stories about finding fairies and talking to trees), privileging the former as a more secure source of information than the latter (Richert & Smith, 2011; Walker, Gopnik, & Ganea, 2015).

Our findings extend this literature by highlighting a property of causal knowledge whose effects on imagination have yet to be explored: its interconnectedness. Causal knowledge is organized in abstract, coherent networks, and our findings suggest that this form of organization constrains our interpretation of events that violate such knowledge. Of course, our findings provide only indirect evidence of the influence of causal structure on the interpretation of causal violations. Future research could explore that influence more directly by measuring or manipulating the causal knowledge relevant to a particular causal violation. 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 aluminum 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.

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. 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, 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) and cognitive parallels in the considerations underlying those judgments (Shtulman & Tong, 2013).

That said, 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 when describing impermissible events, we use hedges like "a little wrong," "sorta wrong," or "mostly wrong." We tend not to use hedges when describing impossible events, though; the phrases "a little impossible," "sorta impossible," or "mostly impossible" generate a tenth as many Google hits as generated by their moral counterparts. Thus, what is odd about the phenomenon at hand is not just that we apply causal expectations to a-causal events but also that we treat the distinction between possibility and impossibility as binary rather than graded.

In conclusion, children and adults alike 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 both on the structure of causal cognition and the structure of imagination.

References

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.

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*, 367, 2202-2212.

Chandler, M. J. & Lalonde, C. E. (1994). Surprising, magical, and miraculous turns of events: Children's reactions to violations of 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.

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., Kavanaugh, R., & Meredith, M. (1994). Young children's comprehension of pretend episodes: Integration of successive actions. *Child Development, 65*, 16-30.

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.

Nichols, S. (2006). Imaginative blocks and impossibility: An essay in modal psychology. In S. Nichols (Ed.), *The architecture of the imagination* (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.

Pieraut-LeBonniec, G. (1980). *The development of modal reasoning: Genesis of necessity and possibility notions*. London, UK: Academic Press.

Purzycki, B. G. (2013). The minds of gods: A comparative study of supernatural agency. *Cognition, 129*, 163-179.

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.

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, D., & Bloom, P. (2006). What does Batman think about SpongeBob? Children's understanding of the fantasy/fantasy distinction. *Cognition, 101*, B9-B18.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science, 27*, 453-489.

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.

Woolley, J. D. (1997). Thinking about fantasy: Are children fundamentally different thinkers and believers from adults? *Child Development, 68*, 991-1011.