Improbable or Impossible? How Children Reason About the Possibility of Extraordinary Events

Andrew Shtulman and Susan Carey

Harvard University

The present study investigated the development of possibility-judgment strategies between the ages of 4 and 8. In Experiment 1, 48 children and 16 adults were asked whether a variety of extraordinary events could or could not occur in real life. Although children of all ages denied the possibility of events that adults also judged impossible, children frequently denied the possibility of events that adults judged improbable but not impossible. Three additional experiments varied the manner in which possibility judgments were elicited and confirmed the robustness of preschoolers' tendency to judge improbable events impossible. Overall, it is argued that children initially mistake their inability to imagine circumstances that would allow an event to occur for evidence that no such circumstances exist.

One of the traits that differentiates human beings from other animals is our ability to learn about entities and events that we have not personally observed. From the testimony of other individuals, we regularly learn about people we have never met (e.g., Beethoven, Rembrandt, Einstein), places we have never been (e.g., Pompeii, Antarctica, Mars), and objects we have never seen (e.g., genes, electrons, radio waves). We can even conceive of entities and events that *no one* has ever observed (e.g., antigravity machines, time travel, human cloning). Of course, not everything that is conceivable is possible. The ability to differentiate possible things from impossible things is therefore an invaluable skill when reasoning about the unobserved and the unobservable.

The development of such an ability is of interest to cognitive psychologists for at least two reasons. First, much of the knowledge we acquire in childhood is learned from the testimony of other individuals, and the extent to which children are able to differentiate possible events from impossible events may determine the extent to which they are able to differentiate fact from fiction, truth from falsehood. Second, many researchers have likened conceptual development to the construction and revision of domain-specific theories, or conceptual structures that embody one's ontological commitments and causal-explanatory knowledge (Carey, 1985; Gopnik & Meltzoff, 1997;

Keil, 1989; Wellman & Gelman, 1992), and yet it is unclear when the causal constraints represented *by* the child become causal constraints *for* the child. In other words, when are children able to reflect upon the causal constraints implicit in their theories to make explicit judgments of what is possible and what is not?

Previous research on children's understanding of magic (Chandler & Lalonde, 1994; Johnson & Harris, 1994; Phelps & Woolley, 1994; Rosengren & Hickling, 1994; Rosengren, Kalish, Hickling, & Gelman, 1994; Subbotsky, 1994, 2004) suggests that the ability to differentiate possible events from impossible events develops early in life. For instance, Johnson and Harris (1994) presented 3- and 4-year-olds with pairs of events in which one event violated a physical principle (e.g., moving a marble with one's mind) and one event did not (e.g., moving a marble with one's hand) and asked the children to decide which event had been performed by an ordinary person and which event had been performed by a magic fairy. Children of all ages tended to claim that the possible event in each pair was performed by an ordinary person and the impossible event was performed by a magic fairy, and they did so for events that violated four different physical principles (i.e., inertia, object permanence, object continuity, and the conservation of matter). Likewise, Rosengren et al. (1994) found that 4- and 5-year-olds consistently denied the possibility of events that violated common biological principles, like growing younger or growing smaller, but did not deny the possibility of events that conformed to those principles.

In reviewing these studies, Rosengren and Hickling (1994) drew the conclusion that "young children make a sharp distinction between possible and impossible events" (p. 1606), and Sharon and Woolley (2004) drew the conclusion that "young children have clear ideas about the kinds of things real entities can and cannot do" (p. 294). However, several other findings call these strong conclusions into question. First, children believe in the existence of fantasy characters whose magical properties (e.g., the ability to fly, the ability to visit millions of homes in a single night) are inconsistent with the causal principles thought to underlie children's differentiation of magical events from ordinary events in the aforementioned experiments. For instance, Prentice, Manosevitz, and Hubbs (1978) found that most children under the age of 6 believe in the existence of Santa Claus, the Tooth Fairy, and the Easter Bunny. Likewise, Woolley, Boerger, and Markman (2004) found that most preschoolers were willing to believe in the existence of a novel fantasy character—the "Candy Witch," or a flying witch that visits children on Halloween and gives them a toy in exchange for their unwanted candy—after only one or two exposures to this counterintuitive concept.

Second, preschoolers will accept the possibility of impossible transformations, like making an object shrink or making an object disappear, if they witness visual illusions in which such transformations appear to have occurred before their very eyes (Chandler & Lalonde, 1994; Rosengren & Hickling, 1994; Subbotsky, 2004). For instance, Chandler and Lalonde (1994) asked children between the ages of 3 and 4 whether one object (a screen) could pass through another object (a box). Although children of all ages denied the possibility of this event when it was described to them verbally, 67% changed their mind after "witnessing" the event first hand. That is, they insisted that the event they had witnessed was not a trick but was actually "real magic."

Third, research on children's understanding of environmental regularities (Komatsu & Galotti, 1986; Lockhart, Abrahams, & Osherson, 1977; Miller, Custer, & Nassau, 2000; Nicholls & Thorkildsen, 1988) has shown that children not only deny the possibility of events that violate physical laws but also deny the possibility of events that violate social laws. For example, Komatsu and Galotti (1986) asked children between the ages of 6 and 10 whether dogs could be called "wugs" and red traffic lights could be switched with purple traffic lights if everyone agreed to the change and found that most 6-year-olds, many 8-year-olds, and even a few 10-year-olds denied the possibility of altering these conventions. Likewise,

Miller et al. (2000) found that many 7-year-olds believe that social laws, like physical laws, are universally true and impervious to change. Although young children explain why people conform to social laws differently than they explain why people conform to physical laws—appealing to mental states in the first case and physical forces in the second (Kalish, 1998; Schult & Wellman, 1997; Sobel, 2004)—the fact that they deny the possibility of both types of events calls into question children's possibility judgments for physically impossible events.

Despite these inconsistencies in how children reason about physical possibility, children appear to understand the concepts *possible* and *impossible* in general. For instance, research on modal language acquisition (Byrnes & Duff, 1989; Hirst & Weil, 1982; Noveck, Ho, & Sera, 1996) has found that children master linguistic distinctions corresponding to the conceptual distinction between possibility and impossibility by the age of 3 (although this may not indicate a full understanding of modal logic; see Coates, 1988). Furthermore, research on children's inferential search behavior (Fabricius, Sophian, & Wellman, 1987; Gonsalves, 1999; Sommerville & Capuani-Shumaker, 1984; Sophian & Somerville, 1988) has found that most 3-year-olds make a clear distinction between possible object locations and impossible object locations, searching for hidden objects in all possible hiding locations but not in any impossible hiding locations.

Perhaps children understand the concepts of possibility and impossibility but have difficulty applying these concepts to a domain as large and as ill-defined as the "real world." After all, the evidence that children understand modal concepts, in general, has come from studies involving small, well-defined domains in which all possibilities can be individuated and enumerated. Obviously, one cannot enumerate all events possible within the real world to determine whether a particular event (e.g., calling dogs "wugs") is among them. Instead, one must draw upon one's knowledge of causal processes and causal constraints to accomplish this task. Accordingly, children who deny the possibility of events that are actually possible (e.g., altering a social convention) may be unaware of what counts as evidence of possibility and what counts as evidence of impossibility.

As an illustration, consider one of the many events that children reliably judge impossible: floating in the air (Browne & Woolley, 2004; Kalish, 1998; Schult & Wellman, 1997; Sobel, 2004). On the one hand, children may deny the possibility of this event because they recognize that the event violates

an *explicitly known* causal principle—for example, "unsupported objects fall," "gravity pulls us downward," "people are denser than air," "everything that goes up must come down." On the other hand, children may deny the possibility of this event simply because they are ignorant of how the event could occur. In other words, children may attempt to imagine circumstances that would allow a person to float in the air, fail to identify these circumstances, and conclude that the event is impossible without ever entertaining a causal principle like "unsupported objects fall." On this view, children who deny the possibility of calling dogs "wugs" or switching red traffic lights with purple traffic lights do so not because they believe that social conventions are unalterable but because they are unable to think of any circumstances that would allow those particular conventions to be altered.

Put differently, children may initially mistake their inability to imagine circumstances that would allow an event to occur for evidence that no such circumstances exist. This claim is reminiscent of Piaget's (1987) claim that children initially confuse necessity with actuality and must learn to differentiate true necessity from "pseudo-necessity," or the impression of necessity based on a superficial analysis of the relevant domain. Piaget supported this claim with the finding that children under the age of 10, when presented a finite set of objects and asked to order the objects in as many ways as possible, often fail to discover the full range of permutations.

In the present study, we explored the applicability of Piaget's ideas about logical possibility to children's reasoning about physical possibility. In particular, we explored the hypothesis that children initially fail to differentiate physical impossibilities, like walking through a wall or walking on water, from "pseudo-impossibilities," like finding an alligator under the bed or growing a beard to one's toes. Although both types of events violate known causal principles, only the former violates physical laws and are therefore impossible. By measuring the extent to which children differentiate these two types of events—that is, "improbable events" and "impossible events"—we hoped to shed light both on what children *mean* when they say that something is impossible and on how they arrive at this conclusion.

Experiment 1

One of the most frequent ways children learn about the occurrence of unobserved events is in the context of stories. We therefore chose a storybook task as an ecologically valid means of probing children's intuitions about physical possibility. A 1,500-word story containing a variety of extraordinary events—both improbable and impossible—was written and illustrated specifically for this purpose.

Method

Participants. Sixteen 4-year-olds ($M = 4$ years 6 months, range = 4 years 0 month to 5 years 0 month), sixteen 6-year-olds ($M = 6$ years 4 months, range = 5 years 11 months to 6 years 10 months), sixteen 8-year-olds ($n = 17$, $M = 8$ years 6 months, range = 8 years 2 months to 9 years 0 month), and 16 adults participated in Experiment 1. The adult participants, mostly Harvard undergraduates, were recruited from the lobby of the psychology department and received monetary compensation for their participation. The 6- and 8-year-olds were recruited from a suburban elementary school and tested at the school itself. The 4-year-olds were recruited from the greater Boston area and tested at the Harvard Laboratory for Developmental Studies, as were those in the other three experiments. Participants in all four experiments were predominantly White and from middle-class families, although a range of ethnic and socioeconomic backgrounds were represented. Approximately equal numbers of boys and girls were included in each experiment.

Materials. The story contained eight ordinary events, eight improbable events, and eight impossible events in a random order. The eight ordinary events (in order of presentation) were wearing a baseball cap, eating an apple, washing a car, losing money, meeting a clown, building a house out of bricks, cleaning a closet, and winning a game. The eight improbable events were finding an alligator under the bed, drinking onion juice, growing a beard to one's toes, owning a lion for a pet, eating pickle-flavored ice cream, getting struck by lightning, making a mug-shaped building, and painting polka dots on an airplane. And the eight impossible events were turning applesauce back into an apple, growing money on a tree, walking through a brick wall, traveling back in time, making a car vanish into thin air, walking on water, opening a window with one's mind, and eating lightning for dinner.

The eight impossible events were designed to violate physical laws known at least implicitly by most children (see Spelke, 1990), and the eight improbable events were designed to violate empirical regularities not typically thought of as "laws," either social or physical. Note that the improbable events were intended to be improbable from a conceptual point of view rather than merely a statistical point of view.

Our objective, after all, was to investigate how children reason about the possibility of events that violate their causal expectations rather than low-frequency events that conform to such expectations (see Hoemann & Ross, 1971, for an example of the latter). Consequently, our selection criterion for the improbable events was that they violate "contingent truths" (e.g., that alligators are found in swamps, that buildings are rectangular, that people dislike the taste of onion juice), and our selection criterion for the impossible events was that they violate "necessary truths" (e.g., that objects are solid, that objects are permanent, that objects require support). Our intuition that the improbable events violated contingent truths and the impossible events violated necessary truths was validated by the adult participants in Experiment 1, who claimed that the improbable events could occur in real life but that the impossible events could not.

The storybook itself contained twenty 8.5″ × 11″ double-sided pages. Accompanying each page of text was a photograph illustrating the events described on that page. Although the ordinary events were easily illustrated with a single, unaltered photograph, the improbable events and the impossible events were illustrated with "before" and "after" photographs or with a single photograph doctored in Adobe Photoshop. To ensure that the inclusion of illustrations did not sway participants' possibility judgments, all 16 extraordinary events were illustrated in some fashion. Figure 1 depicts two sample illustrations: one of an improbable event (finding an alligator under the bed) and one of an impossible event (eating lightning for dinner).

Procedure. Children were read the entire story at the beginning of each interview. Children were then asked whether or not they had experienced each of the events listed above in the order they appeared in the story (e.g., "Have you ever seen a person walk through a wall?"). Whenever children denied having experienced an event, they were asked whether or not the event could occur in real life (e.g., "Could a person walk through a wall in real life?"). Whenever participants denied that an event could occur in real life, they were asked to provide a justification for their judgment (e.g., "Why couldn't a person walk through a wall in real life?"). Note that participants were asked whether the event could be performed by "a person," rather than by the participants themselves, as many of the improbable events could be performed only by individuals with specialized training (e.g., owning a lion for a pet, making a mug-shaped building). Also note that we accepted children's first justification without furthering probing, both to reduce the length of the task and to maintain a standard protocol across children of different ages and different verbal abilities. Adults were asked the same questions as children but in the form of a questionnaire rather than an interview.

Coding justifications. Participants provided a total of 671 justifications for why certain events are impossible. These justifications were sorted into two categories: informative justifications and redundant justifications. Informative justifications provided information about the target event that was not already mentioned in the story. Redundant justifications, on the other hand, provided no information beyond what was already mentioned in the story or what was already discernable from a participant's initial judgment (e.g., "that's not possible," "that's not real," "no one can do it," "it can only happen in stories," "I don't know"). Justifications that referenced magic (e.g.,
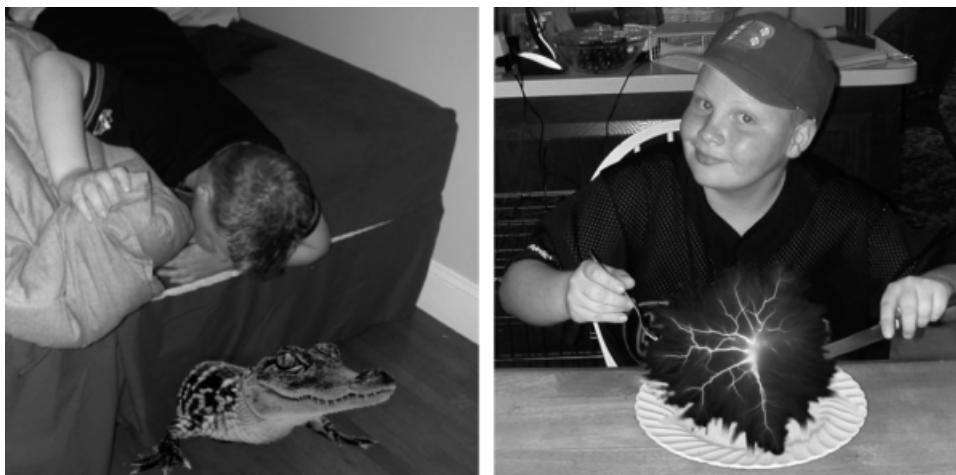

Figure 1. Illustrations of an improbable event (finding an alligator under the bed) and an impossible event (eating lightning for dinner).

Table 1

Sample Factual and Hypothetical Justifications Given in Response to Each Impossible Event

Event	Justification type	Example
Turning applesauce into an apple	Factual	"Applesauce is all mushed"
	Hypothetical	"You could get a new apple"
Growing money on a tree	Factual	"Money is man-made"
	Hypothetical	"It should have leaves on it"
Walking through a brick wall	Factual	"Walls are solid"
	Hypothetical	"You would bonk your head"
Traveling back in time	Factual	"Times goes forward"
	Hypothetical	"A dinosaur would eat you"
Making a car vanish into thin air	Factual	"Nothing just disappears"
	Hypothetical	"You could drive it away"
Walking on water	Factual	"Water is a liquid"
	Hypothetical	"Your feet would sink"
Opening a window with one's mind	Factual	"Minds do not have that power"
	Hypothetical	"You could pretend to open it"
Eating lightning for dinner	Factual	"You can't catch lighting"
	Hypothetical	"It would strike your tummy"

"you would need magic to do that") were also included in this category, as these justifications comprised <6% of all justifications and <7% of the justifications provided by any particular age group.

Informative justifications were sorted into two subcategories: factual justifications and hypothetical justifications. Factual justifications referenced facts about the world that would preclude an event's occurrence (e.g., walking through a wall is impossible because "walls are solid"). Hypothetical justifications, on the other hand, referenced hypothetical events that could occur, or would occur, in place of the actual event under consideration (e.g., walking through a wall is impossible because "you could walk through a door" or because "you would hit your head"). Examples of both types of justifications are displayed in Table 1.

As can be seen from this table, hypothetical justifications were linguistically distinct from factual justifications in that only hypothetical justifications included conditional verbs like "would," "could," or "should." Moreover, hypothetical justifications were *conceptually* distinct from factual justifications in that hypothetical justifications did not actually answer the question at hand—that is, the question of why an event is impossible. Rather, they answered the question of how the same outcome could be achieved by different means (e.g., how a person could get to the other side of a wall without attempting to walk through it) or how a different outcome could be achieved by the same means (e.g., how a person could attempt to walk through a wall without getting to the other side). For this reason, hypothetical justifications could be considered evidence of a failed attempt to identify circumstances that would allow an extraordinary event to occur. In other words, individuals may have provided hypothetical justifications whenever they attempted to think of way the target event could be actualized but fell short of achieving this goal.

The reliability of our coding scheme was assessed by comparing the exhaustive classifications of two independent coders, each blind to the age of the participants providing the justifications. Overall agreement between the two coders was 95%, and all disagreements were resolved through discussion.

Results

Experience judgments. As expected, participants of all ages claimed to have experienced the ordinary events far more often than they claimed to have experienced either the improbable events or the impossible events. On average, children claimed to have experienced 7.0 of the eight ordinary events, 0.1 of the eight improbable events, and 0.1 of the eight impossible events. Likewise, adults claimed to have experienced 7.9 of the eight ordinary events, 0.3 of the eight improbable events, and none of the eight impossible events. One-way analyses of variance (ANOVAs) were used to test the effect of age on participants' experience with each type of event. These analyses revealed a significant effect of age for the ordinary events, $F(3, 63) = 8.03$, $p < .001$, but not the improbable events, $F(3, 63) = 1.30$, *ns*, or the impossible events, $F(3, 63) = 1.40$, *ns*.

On average, 4-year-olds claimed to have experienced 6.6 ordinary events, 6-year-olds claimed to have experienced 6.9 ordinary events, 8-year-olds claimed to have experienced 7.6 ordinary events, and adults claimed to have experienced 7.9 ordinary events. A contrast analysis confirmed that participants' experience with the ordinary events increased linearly with age, $F(1, 63) = 16.27$, $p < .001$. The three ordinary events that children claimed to have experienced least often were wearing a baseball cap (which 25% of children claimed not to have experienced), cleaning a closet (which 21% claimed not to have experienced), and winning a game (which 19% claimed not to have experienced).

Possibility judgments. The average number of ordinary, improbable, and impossible events judged possible by each age group is displayed in Figure 2. As can be seen from this figure, participants of all ages reliably affirmed the possibility of ordinary events, like washing a car or eating an apple, and reliably denied the possibility of impossible events, like turning applesauce into an apple or growing money on a tree. These results confirm those of Johnson and Harris (1994), Rosengren et al. (1994), and Subbotsky (2004) in that even the 4-year-olds distinguished ordinary events from impossible events. However, the meaning of this discrimination was clearly different for children and adults, as only adults reliably affirmed the possibility of improbable events, like owning a lion for a pet or eating pickle-flavored ice cream. Whereas adults judged 99% of the improbable events possible, 4-year-olds judged 22% of the improbable events possible, 6-year-olds judged 50% of the improbable events possible, and
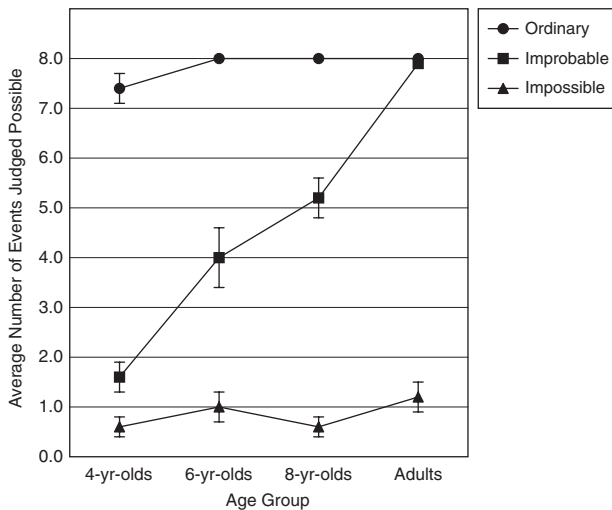
8-year-olds judged 65% of the improbable events possible. The three improbable events that children judged possible least often were finding an alligator under the bed (judged possible by 6% of children), owning a lion for a pet (judged possible by 38% of children), and making a mug-shaped building (judged possible by 42% of children).

A repeated measures ANOVA was used to test the effect of age group (4-, 6-, 8-year-olds, adults) and event type (ordinary, improbable, impossible) on participants' possibility judgments. Age was analyzed between participants and event type was analyzed within participants. As expected, this analysis revealed a significant main effect of age, $F(3, 63) = 32.02$, $p < .001$; a significant main effect of event type, $F(3, 63) = 775.23$, $p < .001$; and a significant interaction between the two factors, $F(3, 63) = 31.82$, $p < .001$. This interaction was explored with contrast analyses of the number of ordinary events, improbable events, and impossible events judged possible at each age. These analyses confirmed that participants' possibility judgments increased linearly with age for only one type of event: the improbable events, $F(1, 63) = 120.29$, $p < .001$.

To determine the extent to which participants in each age group differentiated the three types of events, we performed separate repeated measures ANOVAs for each age group. These analyses revealed a significant effect of event type for all four age groups: 4-year-olds, $F(2, 32) = 323.05$, $p < .001$; 6-year-olds, $F(2, 28) = 99.78$, $p < .001$, 8-year-olds, $F(2, 32) = 241.40$, $p < .001$; and adults, $F(2, 34) = 497.45$, $p < .001$. Bonferroni comparisons of the three event types revealed that all four age groups judged improbable events possible significantly more often than they judged impossible events possible. Only adults, however, judged improbable events possible as often as they judged ordinary events possible. In other words, children of all ages judged improbable events possible significantly less often than they judged ordinary events possible.

Four-year-olds, as a group, judged improbable events possible significantly more often than they judged impossible events possible, but half of the individuals within this group judged fewer than two improbable events possible, as shown in Table 2. This table, which displays the number of participants in each age group who judged few (0–1), some (2–3), many (4–6), or most (7–8) improbable events possible, highlights two aspects of the data not obvious from group means. First, 50% of the 4-year-olds and 25% of the 6-year-olds judged improbable events possible significantly less often than predicted by

Figure 2. Average number of ordinary, improbable, and impossible events judged possible by each age group (out of eight).

Table 2

Number of Participants in Each Age Group Who Judged Few (0 – 1), Some (2 – 4), Many (5 – 6), or Most (7 – 8) Improbable Events Possible

	No. of improbable events judged possible			
Age group	1 – 2	3 – 4	5 – 6	7 – 8
Four-year-olds	8	8	0	0
Six-year-olds	4	4	5	3
Eight-year-olds	0	7	4	5
Adults	0	0	0	16

Note. The first and last columns represent frequencies with a binomial probability < .05.

Figure 3. Average proportion of redundant, hypothetical, and factual justifications provided by each age group.

chance (binomial probability, $p < .05$). Second, only 31% of the 8-year-olds judged improbable events possible significantly more often than predicted by chance (in contrast to 100% of the adults). Thus, most 4-year-olds effectively denied the possibility of all improbable events, and only a few of the 8-year-olds effectively affirmed the possibility of all improbable events.

Justifications. Participants were asked to provide a justification for every event they judged impossible, thereby providing anywhere from 4 justifications to 16 justifications. Justifications for ordinary events were excluded from further analysis due to their infrequency. On average, 4-year-olds provided 13.9 justifications, 6-year-olds provided 11.0 justifications, 8-year-olds provided 10.3 justifications, and adults provided 6.8 justifications. Justifications were coded as factual, hypothetical, or redundant in accordance with the criteria discussed above. Because participants provided different numbers of justifications, absolute frequencies were converted to relative frequencies (by dividing the number of times a participant provided a particular type of justification by the total number of justifications he or she provided) in order to compare justifications across individuals. These data were normally distributed and did not require a logarithmic transformation. The average proportion of factual, hypothetical, and redundant justifications provided by each age group is displayed in Figure 3. As can be seen from this figure, the tendency to provide hypothetical justifications and redundant justifications decreased with age, but the tendency to provide factual justifications increased with age.

One-way ANOVAs were used to test the effect of age on participants' tendency to provide each type of justification. As expected, the effect of age was significant for all three types: redundant justifications, $F(3, 63) =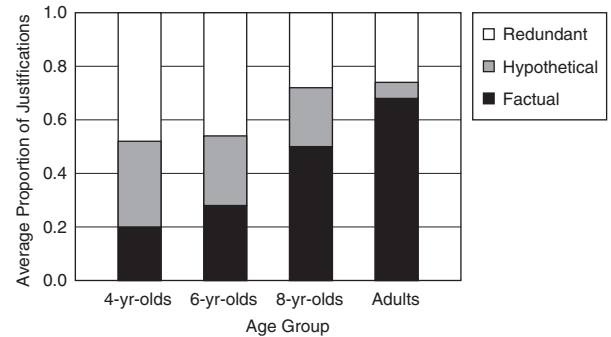 6.43$, $p < .01$; hypothetical justifications, $F(3, 63) = 12.63$, $p < .001$; and factual justifications, $F(3, 63) = 27.94$, $p < .001$. Moreover, contrast analyses confirmed that the tendency to provide factual justifications increased linearly with age, $F(1, 63) = 70.68$, $p < .001$, but the tendency to provide hypothetical justifications decreased linearly with age, $F(1, 63) = 11.54$, $p < .01$, and the tendency to provide redundant justifications decreased linearly with age, $F(1, 63) = 36.96$, $p < .001$. Bonferroni comparisons of the four age groups revealed that children of all ages provided significantly more hypothetical justifications than adults but that 4-year-olds did not provide significantly more hypothetical justifications than either 6- or 8-year-olds. Apparently, children of all ages thought that hypothetical justifications were pragmatically acceptable.

Pearsons correlations were used to determine whether children's tendency to provide a particular type of justification was related to their tendency to judge improbable events possible. Whereas children's tendency to provide factual justifications was positively correlated with their tendency to judge improbable events possible, $r(48) = .35$, $p < .001$, their tendency to provide redundant justifications was negatively correlated with this same tendency, $r(48) = -.31$, $p < .001$. Neither correlation, however, remained significant after controlling for the children's age (in months), possibly because age was not a continuous variable.

To determine whether children provided different types of justifications for the different types of events, we divided children's justifications into two groups—those provided for improbable events and those provided for impossible events—and calculated the proportion of factual, hypothetical, and redundant justifications within each group. These proportions were highly similar for all justification types and for all age groups. In fact, none of the nine comparisons were statistically significant, indicating

that children justified their judgments of impossibility for the improbable events no differently than they justified their judgments of impossibility for the impossible events. Even factual justifications were provided equally often for both types of events, as some children claimed that drinking onion juice is impossible because "onions don't have juice" or that getting struck by lightning is impossible because "lightning is too short." From an adult's perspective, these "facts" are either untrue (as in the case of onions) or only occasionally true (as in the case of lightning). From a child's perspective, on the other hand, these facts may be as universally true as the kinds of facts they provided in response to the impossible events.

In sum, not only were children markedly more likely than adults to judge improbable events impossible, they also failed to differentiate improbable events from impossible events in how they justified those judgments.

Discussion

When judging the possibility of events in a domain as large and as ill-defined as the "real world," individuals must draw upon their prior knowledge of causal processes and causal constraints to make such judgments. Experiment 1 investigated the development of this ability by asking children between the ages of 4 and 8 to decide whether a variety of extraordinary events could or could not occur in the real world. Like adults, children of all ages denied the possibility of events that violated physical laws. However, unlike adults, they also denied the possibility of events that did not violate *any* kind of law, like finding an alligator under the bed or making a mug-shaped building, and they rarely justified their judgments of impossibility with facts about the world that would preclude an event's occurrence.

One interpretation of these results is that children and adults rely on different standards of evidence for deciding that an event is impossible. Whereas adults rely on identifying facts about the world that would preclude an event's occurrence, children rely on their ability to identify circumstances that would allow an event to occur. Failing to identify such circumstances, they deny the possibility of many events that adults judge to be possible. Granted, even adults may base their initial impressions of an event's possibility on their ability to imagine circumstances that would allow the event to occur, but the adult participants in Experiment 1 rarely based their final judgments on such considerations. Rather, they sought out principled reasons (i.e., factual justifica-

tions) for why an event that *seemed* impossible *was* impossible.

Of course, part of the reason why children showed little evidence of using explicit causal principles to guide their possibility judgments may be that children are less capable than adults at articulating the reasoning behind such judgments. Indeed, children of all ages might have provided more factual justifications for their judgments of impossibility if we had probed their reasoning further. They might even have changed their minds about the possibility of improbable events initially judged impossible. Children's justifications should not therefore be taken as evidence that children cannot differentiate improbable events from impossible events *in principle*. Still, the fact that 4-year-olds rarely appealed to causal principles in their spontaneous justifications suggests that these principles did not play a large role —or at least not an explicit role—in their initial judgments.

In truth, there are at least two ways children could have arrived at the conclusion that an event is impossible without ever having referenced an explicit causal principle. First, they may have attempted to identify circumstances that would allow the event to occur, failed in their attempt to do so, and concluded that the event must be impossible. Second, they may not have attempted to "model" the event per se but simply based their judgment on the phenomenological experience of having an expectation violated—that is, feelings of surprise, confusion, or disbelief. Evidence of both processing accounts can be found in the way in which children most typically justified their judgments of impossibility. Whereas children's hypothetical justifications (e.g., "he couldn't find a real alligator under the bed but he could find a toy alligator") are consistent with the first account, children's redundant justifications (e.g., "it can't happen") are consistent with the second. Note that this explanation, if correct, implies that children who provided mainly redundant justifications (i.e., 4- and 6-year-olds) based most of their judgments on pure phenomenological experience.

Two relatively trivial explanations for children's tendency to deny the possibility of improbable events can be ruled out immediately. First, it was not the case that children denied the possibility of any event they had not personally experienced. Children denied the possibility of 55% of the improbable events they had not personally experienced, but only 15% of the ordinary events they had not personally experienced. Moreover, 3 of the 5 children who occasionally denied the possibility of ordinary events

affirmed the possibility of *other* ordinary events that they had not personally experienced.

Second, it was not the case that children denied the possibility of any event with a negative emotional valence. Although half of the improbable events had a negative emotional valence (i.e., finding an alligator under the bed, drinking onion juice, eating pickle-flavored ice cream, and getting struck by lightning), the other half did not (i.e., growing a beard to one's toes, owning a lion for a pet, making a mug-shaped building, and painting polka dots on an airplane), and children denied the possibility of both types of events equally often (i.e., 57% of the time vs. 54% of the time). Furthermore, children consistently denied the possibility of impossible events with *positive* emotional valences, like growing money on a tree or walking on water, even though many participants may have wished that these events were, in fact, possible. Clearly, the events' emotional valence cannot explain children's overall patterns of judgments.

Two other explanations for children's failure to judge improbable events possible are not as easily dismissed. First, children may have interpreted the task as a test of their factual knowledge rather than a test of their modal intuitions. In other words, children may have interpreted the question *"Could* [the event] happen in real life?" as *"Does* [the event] happen in real life?" This explanation is unlikely to be true given that children as young as 3 differentiate factual verbs from modal verbs (as noted previously). Nevertheless, we tested this possibility directly in Experiment 2 by assessing children's understanding of the modal verb "could" before administering the storybook task. Second, children may have denied the possibility of improbable events because they systematically underestimated the scope of the possibility judgment they were being asked to make. In other words, children may have interpreted the question "Could [the event] happen *in real life?"* as "Could [the event] happen *under normal circumstances?"* Experiment 3 tests this possibility by asking children to make magic judgments rather than possibility judgments.

Experiment 2

In English, one can express the likelihood that a statement is true using modal verbs (e.g., must, might, may, can, could, should), modal adjectives (e.g., possible, probable, likely, certain), modal adverbs (e.g., definitely, maybe, perhaps), and modal inflections (e.g., imperative tense, subjunctive tense, conditional tense; See Perkins, 1983). From the per-spective of a mature modal language user, the linguistic distinction between "could" and "could not" maps onto the conceptual distinction between possible and impossible states of affairs, not the distinction between actual and hypothetical states of affairs. Might deficits in modal language (or modal logic) have led the youngest participants in Experiment 1 to interpret modal questions like *"Could* a person find an alligator under the bed in real life?" as factual questions like *"Do* people find alligators under the bed in real life?"

This explanation would be consistent with the claim made by Pieraut-LeBonniec (1980) that young children systematically misunderstand modal language. In one of several studies, Pieraut-LeBonniec showed children a stick, a marble, and a box with two holes cut into its top. Although the stick fit through both holes, the marble fit through only the larger of the two holes. After children were familiarized with these materials, the experimenter put one of the objects into the large hole and asked the child "Can you tell without opening the [box's] door what it is?" None of the 4-year-olds, 25% of the 6-year-olds, 40% of the 8-year-olds, and 85% of the 10-year-olds in her study were able to answer the question correctly (i.e., "no"). The remaining children claimed that they knew for sure which object had been put into the box: the marble.

Pieraut-LeBonniec interpreted her findings as evidence that children under the age of 12 use a modal logic that is qualitatively different from that used by adults, which, if true, might explain the developmental change documented in Experiment 1. There are, however, at least two reasons to doubt Pieraut-LeBonniec's general claim in favor of a more specific claim—namely, that many children under the age of 12 fail to understand the distinction between guessing and knowing (see Fay & Klahr, 1996; Klahr & Chen, 2003).

First, even though none of the 4-year-olds in Pieraut-LeBonniec's original experiments could answer questions about epistemic certainty, virtually all of them could answer questions about physical possibility, like "Can the marble fit through the small hole?" or "Can the stick fit through the large hole?" In other words, children who failed to acknowledge the uncertainty of which possibility was an actuality still acknowledged the existence of multiple possibilities (see also Horobin & Acredolo, 1989). Second, multiple researchers (e.g., Byrnes & Duff, 1989; Hirst & Weil, 1982; Noveck et al., 1996) have found that children as young as 3 understand that statements expressed with a factive verb (e.g., "the peanut *is* under the box") are more likely to be true than

statements expressed with certain modal verbs (e.g., "the peanut *may be* under the box"), further strengthening the claim that young children understand modal language as applied to the world even if they fail to understand modal language as applied to states of knowledge.

Given that the participants in Experiment 1 were asked modal questions about the world, it is unlikely that their failure to judge improbable events possible was due to deficits in modal language or modal logic. In Experiment 2, we attempted to confirm this assumption by asking 4-year-olds (the youngest participants in Experiment 1) to judge the possibility of probable, improbable, and impossible events in two well-defined domains before asking them to judge the possibility of ordinary, improbable, and impossible events with respect to the world at large. If preschoolers' failure to judge improbable events possible is due to general deficits in modal language or modal logic, then they should deny the possibility of improbable events in both tasks. If, on the other hand, preschoolers' failure to judge improbable events possible is specific to reasoning about physical possibility, then they should affirm the possibility of improbable events in the first task but continue to deny the possibility of improbable events in the second. Furthermore, if young children's performance in Experiment 1 reflects a shaky command of modal language, then practice answering modal questions may improve their performance.

Method

Participants. Twelve 4-year-olds ($M = 4$ years 3 months, range = 4 years 0 month to 4 years 9 months) similar to those in Experiment 1 participated in Experiment 2. One additional participant was replaced after his mother intervened during the storybook task.

Marbles task. The purpose of this task was to assess the children's ability to make a frequency-based modal inference. Each child was shown a container containing 10 blue marbles and one red marble. After the child had counted both kinds of marbles, the experimenter poured the marbles from the container into an empty bag. The experimenter then retrieved one marble from inside the bag and, without revealing the color of the marble in his hand, asked the child three questions: (1) "Could the marble in my hand be blue?" (a probable event), (2) "Could the marble in my hand be red?" (an improbable event), and (3) "Could the marble in my hand be yellow?" (an impossible event). If 4-year-olds correctly inter-

pret "could" as a modal verb, then they should respond affirmatively to the first two questions and negatively to the last. If, on the other hand, 4-year-olds incorrectly interpret "could" as a factive verb, then they should respond affirmatively to one, and only one, of the first two questions (as the marble in the experimenter's hand could not be both blue and red).

Mouse task. This task, which was modeled after a task used by Gonsalves (1999), was used to assess children's ability to make a rule-based modal inference. Each child was shown an array of six cups— four red cups and two blue cups—and told that two toy mice—"Mickey" and "Minnie"—were hiding under separate cups of the same color. Once the child could successfully repeat this rule, the experimenter asked the child two questions: (1) "Could Mickey and Minnie be hiding under the red cups?" (a probable event), and (2) "Could Mickey and Minnie be hiding under the blue cups?" (an improbable event). The experimenter then lifted one of the blue cups, revealing Mickey, and asked one final question: "Could Minnie be hiding under a red cup?" (an impossible event, at least within the context of the task). Again, if 4-year-olds correctly interpret "could" as a modal verb, then they should respond affirmatively to the first two questions and negatively to the last. If, on the other hand, 4-year-olds incorrectly interpret "could" as a factive verb, then they should respond affirmatively to one, and only one, of the first two questions about Mickey's location.

Storybook task. Following the modal language tasks, the experimenter administered the storybook task as described in the methods section of Experiment 1.

Results and Discussion

As expected, children performed quite accurately in both the marble task and the mouse task. Ten of the 12 children affirmed the possibility of both probable events, 10 affirmed the possibility of both improbable events, and 9 denied the possibility of both impossible events. Thus, as a group, children affirmed the possibility of 92% of the probable events, 92% of the improbable events, and 21% of the impossible events. Moreover, only 1 child failed in *both* tasks to affirm the possibility of both the probable event and the improbable event, which is what all children should have done if they interpreted "could" as a factive verb.

Although children rarely denied the possibility of improbable events in the objects tasks, they

Table 3

Percentage of 4-Year-Olds in Experiment 2 Who Affirmed the Possibility of Each Improbable Event in the Objects Task and the Storybook Task

Task	Event	Judged possible
Object	Pulling a red marble from the bag	100
	Finding both mice under the blue cups	83
Storybook	Eating pickle-flavored ice cream	42
	Getting struck by lightning	42
	Painting polka dots on an airplane	33
	Drinking onion juice	33
	Growing a beard to one's toes	17
	Making a mug-shaped building	8
	Owning a lion for a pet	0
	Finding an alligator under the bed	0

frequently denied the possibility of improbable events in the storybook task. A comparison of the percentage of children who judged each improbable event possible in the objects tasks and the percentage of children who judged each improbable event possible in the storybook task is displayed in Table 3. As can be seen from this table, children denied the possibility of each improbable event in the storybook task at least twice as often as they denied the possibility of each improbable event in the objects tasks. Paired t tests were used to test the effect of task type (objects tasks, storybook task) on the proportion of ordinary/probable, improbable, and impossible events judged possible by each participant. These analyses revealed a significant effect of task for the improbable events, $t(1, 11) = 10.31$, $p < .001$, but not the ordinary/probable events, $t(1, 11) = 0$, ns, or the impossible events, $t(1, 11) = 1.15$, ns. Apparently, children's tendency to deny the possibility of improbable events in the storybook task is not due to deficits in modal language or modal logic. Rather, this tendency appears to be specific to reasoning about possibility within a domain as large and as ill-defined as the "real world."

Answering modal questions in the objects tasks did not appear to improve 4-year-olds' performance in the storybook task relative to the performance of the 4-year-olds in Experiment 1. On average, the 4-year-olds in Experiment 2 judged 7.3 of the eight ordinary events possible, 2.1 of the eight improbable events possible, and 0.7 of the eight impossible events possible. The comparable numbers in Experiment 1 were 7.4, 1.6, and 0.6. A repeated measures ANOVA confirmed that children's possibility judgments in Experiment 2 varied significantly by event type, $F(2, 22) = 143.68$, $p < .001$. Bonferroni comparisons of the different types of events revealed that children judged improbable events possible signifi-

cantly less often than they judged ordinary events possible but significantly more often than they judged impossible events possible. That said, 5 children judged improbable events possible significantly less often than would be predicted by chance (binomial probability, $p < .05$). In short, the 4-year-olds in Experiment 2 denied the possibility of improbable storybook events as often as the 4-year-olds in Experiment 1—$M = 2.1$ versus $M = 1.6$, $t(26) = 0.98$, ns—despite having just succeeded on tasks that required (and potentially primed) accurate comprehension of the modal verb "could."

As in Experiment 1, children were asked to justify their judgments of impossibility. The number of justifications provided by individual children ranged from 9 to 16. The average proportion of children's justifications that were factual, hypothetical, and redundant were .33, .21, and .46. These proportions were statistically similar to the proportions of factual, hypothetical, and redundant justifications provided by the 4-year-olds in Experiment 1, and children provided approximately the same proportion of factual, hypothetical, and redundant justifications for the improbable events as they provided for the impossible events. Thus, once again, 4-year-olds failed to differentiate improbable events from impossible events both in their judgments of impossibility and in their justifications for those judgments.

In summary, these data corroborate the claims made by other authors (e.g., Hirst & Weil, 1982) that young children distinguish between possibility and actuality and have learned various linguistic markers for this distinction. Why then did children perform comparably to adults in the objects tasks but not comparably to adults in the storybook task? Both the objects tasks and the storybook task required children to answer modal questions about the world. However, only the storybook task required children to draw upon their prior knowledge of causal processes and causal constraints to generate their own evidence of possibility or impossibility. In contrast, the space of possible events in the both objects tasks was finite and well specified, allowing the child to discern that the improbable events (i.e., pulling a red marble from the bag, finding both mice under the blue cups) were among those events but the impossible events (i.e., pulling a yellow marble from the bag, finding one mouse under a red cup and one mouse under a blue cup) were not. In other words, when children discovered a contradiction between the target event and the set of events deducible from the information they had been given, that contradiction normatively specified an impossibility. This was not true for the storybook task, as con-

tradictions between the target event and all "imaginable" events may or may not have specified an impossibility.

Having established that the failure to judge improbable events possible documented in Experiment 1 and replicated in Experiment 2 cannot be attributed to language deficits, we now turn to the question of whether children systematically underestimated the scope of the possibility judgment they were asked to make.

Experiment 3

Philosophers interested in modal logic have often pointed out that modal language is context-dependent (see Johnson-Laird, 1978). The sentence "John *may* leave," for example, can be interpreted metaphysically, as in "it is logically/physically possible for John to leave"; epistemically, as in "I am uncertain whether John will leave"; deontically, as in "John is allowed to leave"; or dynamically, as in "John has the ability to leave." Because of this ambiguity, we specified the scope of the possibility judgments that participants were asked to make (possible in real life) with every question. Still, children may have interpreted questions like "Could a person drink onion juice *in real life*?" as "Could a person drink onion juice *under normal circumstances*?"

One way to address this concern would be to "contextualize" the improbable events—that is, to provide a context in which the events might actually occur—and compare children's judgments for the contextualized events with their judgments for the "decontextualized" events in the original story. If this manipulation improved children's performance, one could conclude that children do, in fact, understand the modal question posed to them but are impaired at identifying circumstances that might allow an improbable event to occur. If, on the other hand, this manipulation did not improve children's performance, one would be unable to determine whether children failed to understand the modal question posed to them or failed to find the contextualized events more plausible than the decontextualized events.

Another way of addressing this same concern would be to circumvent the ambiguities inherent in modal language altogether and ask children to make "magic" judgments rather than possibility judgments. We decided to adopt this approach, rather than the previous approach, for two reasons. First, many authors (e.g., Johnson & Harris, 1994; Phelps & Woolley, 1994; Rosengren & Hickling, 1994) have shown that children appeal to magic when asked to explain the occurrence of physically impossible events, implying that children interpret the word "magic" as synonymous with "impossible." Second, Browne and Woolley (2004) have shown that children as young as 3 are better able to differentiate events that violate social laws from events that violate physical laws when asked to make magic judgments than when asked to make possibility judgments, although the magnitude of this differentiation was small (i.e., a 31% difference). Posing modal questions in terms of magic may therefore make the scope of such questions clearer, thereby allowing children as young as 4 to treat improbable events like ordinary events if their failure to do so in Experiments 1 and 2 was due mainly to a misinterpretation of the question.

Method

Participants. Twelve 4-year-olds ($M = 4$ years 5 months, range = 4 years 1 month to 4 years 10 months) similar to those in previous experiments participated in Experiment 3. Two children objected to the premises of the task, insisting that magic is not real, and had to be replaced.

Procedure. The storybook task was administered in the same manner as it was administered in Experiments 1 and 2 with the exception that whenever children denied having experienced an event, they were asked whether or not the event required magic to occur in real life (e.g., "Would it take magic to walk through a wall in real life?") as opposed to whether or not the event was possible. Thus, the measure of interest is the number of improbable events that children claimed could occur in real life *without magic*.

Results and Discussion

Overall, the magic judgments of the 4-year-olds in Experiment 3 were quite similar to the possibility judgments of the 4-year-olds in Experiments 1 and 2. On average, children claimed that 7.7 of the eight ordinary events could occur without magic, 2.6 of the eight improbable events could occur without magic, and 0.5 of the eight impossible events could occur without magic. In comparison, the 4-year-olds in Experiment 1 claimed that 7.4 of the eight ordinary events could occur in real life, 1.6 of the eight improbable events could occur in real life, and 0.6 of the eight impossible events could occur in real life (on average).

A repeated measures ANOVA confirmed that children's magic judgments varied significantly by event type, $F(2, 22) = 180.90$, $p < .001$. Bonferroni comparisons of the different events revealed that children associated improbable events with magic significantly more often than they associated ordi-

nary events with magic but significantly less often than they associated impossible events with magic. Still, 3 of the 12 children claimed that fewer than two improbable events could occur without magic and only 1 child claimed that more than half of the improbable events could occur without magic.

Overall, the number of improbable events that the 4-year-olds in Experiment 3 claimed could occur without magic was greater than the number of improbable events that the 4-year-olds in Experiment 1 claimed could occur in real life, $t(27) = 2.19$, $p < .05$. However, the magnitude of this difference was 1.0 out of eight events (or 12.5%), a very modest difference. Indeed, similar to 4-year-olds in Experiment 1, 92% of the 4-year-olds in Experiment 2 claimed that it would take magic to make a mug-shaped building, 83% claimed it would take magic to find an alligator under the bed, and 75% claimed it would take magic to grow a beard to one's toes.

In short, children's ability to differentiate improbable events from impossible events was not substantially improved when children were asked to make magic judgments rather than possibility judgments. Children claimed that it would take magic not only to turn applesauce into an apple or to make a car vanish into thin air but also to paint polka dots on an airplane or to own a lion for a pet. Although different children may have interpreted the word "magic" in different ways, children's overall usage was consistent with Phelps and Woolley's (1994) claim that children initially appeal to magic to explain events that "both violate their expectations and elude adequate physical explanation" (p. 385). In support of this claim, they show that children between the ages of 4 and 8 use the word "magic" to describe extraordinary physical transformations that they are unable to explain in terms of known causal processes (e.g., thermodynamically induced color changes, magnetically induced motion changes) but refrain from using the word "magic" to describe extraordinary transformations that they *are* able to explain. Presumably, the act of making a mug-shaped building is as unexplainable to most 4-year-olds as the act of changing an object's color with heat, but is the act of making a mug-shaped building no less mysterious than the act of walking through a wall? This question is addressed in Experiment 4 with the use of a forced-choice paradigm.

Experiment 4

Some of the aforementioned studies that purported to reveal an early-developing understanding of physical impossibility used a forced-choice paradigm in which children were asked to identify which of two events was possible and which was impossible (e.g., Browne & Woolley, 2004; Johnson & Harris, 1994). Under these conditions, even 3-year-olds are able to differentiate impossible events from possible events (albeit, ordinary events). Would 4-year-olds, who fail to judge improbable events possible when presented in isolation, be able to distinguish improbable events from impossible events when these two types of events are presented in a pair?

One reason to believe they could is that the impossible events in the story violate expectations present in infancy—for example, expectations regarding object permanence (Wynn, 1992), object solidity (Baillargeon, Spelke, & Wasserman, 1985), object continuity (Spelke, Kestenbaum, Simons, & Wein, 1995), object support (Needham & Baillargeon, 1993), and contact causality (Leslie & Keeble, 1987)—but the improbable events in the story violate expectations acquired later in life. Even though such differences in knowledge entrenchment do not appear to affect children's possibility judgments, they may affect the certainty with which those judgments are held. Accordingly, we should expect at least some improvement in 4-year-olds' ability to judge improbable events possible if children are forced to choose between an improbable event and an impossible event. At issue, however, is whether this improvement is categorical—that is, whether giving children some training on the contrast of interest and highlighting this contrast in a forced-choice paradigm induces adult-like performance.

To investigate this issue, we paired each improbable event with an impossible event and asked 4-year-olds to decide which event could occur in real life and which event could not. If 4-year-olds are equally confident that both types of events cannot occur in real life, then they should claim that the improbable event is possible on half the trials and that the impossible event is possible on the other half. If, on the other hand, 4-year-olds are more confident that the impossible events cannot occur in real life than they are that the improbable events cannot occur in real life, then they should claim that the improbable event is possible on significantly more trials than they claim that the impossible event is possible.

Method

Participants. Twelve 4-year-olds ($M = 4$ years 7 months, range = 4 years 0 month to 4 years 11 months) similar to those in previous experiments

participated in Experiment 4. One child refused to complete the task and was replaced.

Materials. The story used in Experiments 1–3 was used in Experiment 4 as well. Each improbable event in the story was paired with each impossible event in the story to create 64 unique pairs of events. These pairs were divided into eight groups of eight such that each group contained one instance of every event. The order of the events within each pair was counterbalanced such that the improbable event preceded the impossible event in half of the pairs and the impossible event preceded the improbable event in the other half. Illustrations from the story were then reproduced on $3'' \times 3''$ cards. Each card corresponded to a unique event. Six additional cards were created for the purpose of training children on how to complete the task. These cards depicted three ordinary events (a chicken laying eggs, two children talking, and a flying bird) and three impossible events (a chicken laying tennis balls, two dogs talking, and a flying pig).

Procedure. Children were shown two containers—one labeled with a ''thumbs-up'' sticker and one labeled with a ''thumbs-down'' sticker—and told that the thumbs-up container was for pictures of things that can happen in real life and that the thumbs-down container was for pictures of things that cannot happen in real life. The experimenter demonstrated the function of each container by sorting one pair of training cards (the chickens) into the appropriate containers. Each child was then encouraged to sort the remaining pairs of training cards on their own. All 12 children sorted the cards correctly on their first attempt. Of course, these training trials contrasted impossible events with ordinary events, a differentiation that 4-year-olds easily made in Experiments 1–3. Following the training procedure, the experimenter read the story in its entirety. The children were then asked to sort the eight pairs of pictures taken from the story itself. Because we tested 12 children but constructed only eight groups of event pairings, half of the event pairings were used twice.

Results and Discussion

Overall, the forced-choice paradigm improved children's ability to differentiate improbable events from impossible events, relative to that of Experiments 1–3, but did not induce adult-like performance. On average, each child placed 6.0 improbable-event cards in the thumbs-up container and 2.0 improbable-event cards in the thumbs-down container. The frequency of placing improbable-event cards in the correct container (i.e., the thumbs-up container) was significantly >4.0, the frequency predicted by chance, $t(11) = 5.75$, $p < .001$. In addition, all improbable-event cards were placed in the thumbs-up container more often than they were placed in the thumbs-down container, and all impossible-event cards were placed in the thumbs-down container more often than they were placed in the thumbs-up container, as shown in Table 4. In other words, children judged each improbable event possible more often than they judged each impossible event possible (as a group).

Despite the 4-year-olds' successful performance as a group, only 4 children placed more improbable-event cards in the thumbs-up container than the number predicted by chance (i.e., 7 or 8). Indeed, a quarter of the children placed fewer than six improbable-event cards in the thumbs-up container, and half of the children tried to place *both* cards in the thumbs-down container at some point in the experiment—a situation in which the experimenter intervened and asked the child to place one, and only one, card in each container.

A comparison of the performance of the 4-year-olds in Experiment 4 and the performance of the 4-year-olds in Experiments 1–3 is displayed in Figure 4. Because the 4-year-olds in Experiment 4 could be expected to judge four improbable events possible by chance alone, we could not use the number of improbable events judged possible as a metric of

Table 4

Percentage of 4-Year-Olds in Experiment 4 Who Placed Each Extraordinary Event in the Thumbs-Up Container, Thereby Judging the Event Possible

| Event type	Event description	Judged possible
Improbable	Drinking onion juice	100
	Eating pickle-flavored ice cream	92
	Owning a lion for a pet	83
	Growing a beard to one's toes	67
	Painting polka dots on an airplane	67
	Finding an alligator under the bed	58
	Getting struck by lightning	58
	Making a mug-shaped building	58
Impossible	Opening a window with one's mind	42
	Eating lightning for dinner	42
	Traveling back in time	33
	Walking on water	33
	Walking through a brick wall	17
	Making a car vanish into thin air	17
	Turning applesauce back into an apple	8
	Growing money on a tree	8
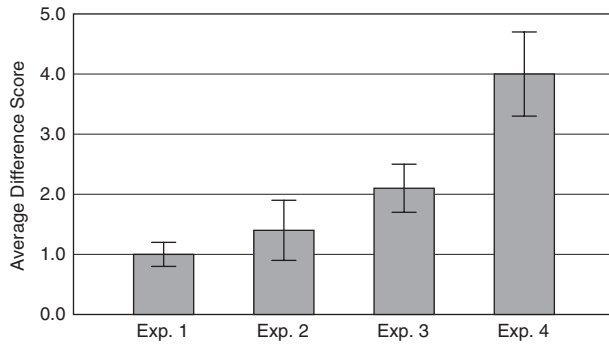
Figure 4. Average difference between the number of improbable events judged possible (or nonmagical) and the number of impossible events judged possible (or nonmagical) by each 4-year-old in Experiments 1–4.

comparison. Instead, we used the difference between the number of improbable events judged possible (or nonmagical, in the case of Experiment 3) and the number of impossible events judged possible (or nonmagical) as a metric of comparison.

As can be seen from this figure, the 4-year-olds in Experiment 4 differentiated improbable events from impossible events to a much greater extent than the 4-years-olds in Experiments 1–3 did. Indeed, a one-way ANOVA confirmed that participants' difference scores varied significantly by experiment, $F(3, 51) = 8.80$, $p < .001$. Moreover, Bonferroni comparisons of the four experiments revealed that the average difference scores in Experiments 1–3 were significantly smaller than the average difference score in Experiment 4 but did not differ from one another. Thus, even though 4-year-olds deny the possibility of improbable events when evaluating these events in isolation, they recognize that improbable events are more likely to occur in the real world than impossible events when evaluating these two types of events in conjunction.

That said, 4-year-olds' success in the forced-choice paradigm was still quite moderate. Rather than judging all eight improbable events possible under these constraints, most children judged only six improbable events possible under these constraints. Furthermore, the paradigm used in Experiment 4 was highly artificial, for the kinds of possibility judgments that individuals need to make outside the psychology laboratory are almost always single-event judgments. Although Experiment 4 confirms the fact that the causal constraints violated by the impossible events (i.e., physical laws) are more deeply entrenched than those violated by the improbable events (i.e., empirical regularities), it does not elucidate how children would treat these events

in any real-world situation. Thus, it may be inappropriate to attribute to children an understanding of the distinction between possible events and impossible events on the basis of such evidence, as others (e.g., Browne & Woolley, 2004; Johnson & Harris, 1994) have done.

General Discussion

Previous research has shown that, from a very early age, children deny the possibility of many impossible events, like floating in the air, walking through a wall, moving a marble with one's mind, or hanging on a tree branch forever. The present study sheds new light on how children make these judgments. When adults reason about the possibility of expectation-defying events, they consistently differentiate events that violate physical laws (impossible events) from those that do not (improbable events), and they tend to justify their judgments of impossibility by appealing to those laws. Children, on the other hand, deny the possibility of both impossible events *and* improbable events, and they typically justify their judgments in redundant or pragmatically inappropriate ways. Moreover, children's justifications for why impossible events could not occur in real life do not differ from their justifications for why improbable events could not occur in real life.

This pattern of results, which was replicated in three experiments, cannot be due to general deficits in modal language or modal logic given that the children in Experiment 2 drew accurate modal inferences when operating within small, well-defined domains. Why, then, might children differ from adults in how they reason about physical possibility? Clearly, part of the answer is that children possess less domain-specific knowledge than adults do. One cannot reflect upon the laws of nature to decide whether an event is possible or impossible unless one knows those laws. Likewise, one cannot identify the circumstances under which an improbable event could occur if those circumstances involve unfamiliar causal mechanisms. Still, there are at least three reasons to doubt that differences in domain-specific knowledge are the only relevant differences between children and adults.

First, it is difficult to specify exactly what knowledge is needed to affirm the possibility of improbable events like growing a beard to one's toes or painting polka dots on an airplane that 4-year-olds do not already possess. Second, children frequently justified their judgments of impossibility in a way that adults almost never did—that is, by referencing hypothetical events that could occur, or would occur,

in place of the actual event under consideration—which suggests that children may often rely on a different form of reasoning than adults do. Third, it is unclear why lacking domain-specific knowledge would lead children to judge possible events impossible but would not also lead them to judge impossible events possible, and yet children systematically exhibited one bias but not the other.

How, then, might children have approached the task differently than adults? As mentioned previously, there are at least two possibilities: (1) children denied the possibility of any event they found bizarre or surprising and (2) children denied the possibility of any event for which they were unable to identify circumstances that would allow the event to occur. Note that these accounts are not mutually exclusive. Rather, both accounts assume that children's judgments of impossibility are based on *ignorance* of how an event could occur rather than *knowledge* of why the event could not occur. If either account is accurate, then young children's possibility judgments cannot be taken as evidence of an explicit appreciation of the causal principles that constrain real-world events, as many authors have done in the past (e.g., Johnson & Harris, 1994; Rosengren & Hickling, 1994).

Admittedly, the above hypotheses are supported only indirectly by the findings of Experiments 1–3. Direct support for these hypotheses would entail prior measurement of children's causal expectations or the intentional manipulation of children's imaginative activities. For instance, one could vary the context in which an improbable event is presented and compare children's willingness to affirm the possibility of this event in different contexts (e.g., drinking onion juice at home vs. drinking onion juice at an onion-themed festival). Alternatively, one could vary the content of an improbable event and compare children's willingness to affirm the possibility of different versions of the same event (e.g., finding an alligator under the bed vs. finding a rabbit under the bed).

On the whole, the present study helps to resolve several inconsistencies highlighted in the introduction regarding children's understanding of magic, fantasy, and physical necessity. First, our findings help to explain why young children not only deny the possibility of events that violate physical laws but also deny the possibility of events that violate social laws, like going barefoot at school or calling dogs "wugs" (Kalish, 1998; Komatsu & Galotti, 1986; Miller et al., 2000; Nicholls & Thorkildsen, 1988). Although these findings have typically been interpreted as evidence that children initially conflate different types of necessities—that is, physical necessity and social necessity—our findings suggest that children may initially base their possibility judgments on information that does not allow them to distinguish between these two types of necessity. For instance, children may deny the possibility of calling dogs "wugs" not because they believe that social laws are unalterable but because they are unable to think of any circumstances under which all English speakers would change their lexicons accordingly. We believe that a process-based interpretation of children's failure to differentiate social laws from physical laws is more parsimonious than a knowledge-based interpretation of this failure in light of our finding that children deny the possibility of events that do not embody *any* type of necessity, physical or social.

Second, our findings help to explain why most 4-year-olds endorse the existence of "real magic" and will accept the possibility of impossible physical transformations if those transformations are "performed" before their very eyes (Chandler & Lalonde, 1994; Rosengren & Hickling, 1994; Subbotsky, 2004). These findings, although seemingly inconsistent with children's early-developing ability to differentiate impossible events from ordinary events, are consistent with our finding that children rarely identify *principled reasons* for why impossible events are, in fact, impossible. Thus, children's judgments of impossibility may be easily overturned in the face of conflicting evidence, particularly perceptual evidence. After all, what is actual must be possible. Interestingly, children begin to doubt the existence of real magic around the same time they begin to differentiate improbable events from impossible events (i.e., age 6), which suggests that both behaviors may be due to an initial absence of principle-based reasoning.

Third, our findings help to explain why children believe in the existence of fantasy characters, like Santa Claus, the Tooth Fairy, and the Easter Bunny (Prentice et al., 1978; Sharon & Woolley, 2004; Woolley et al., 2004). On its surface, children's skepticism toward the possibility of expectation-defying events, appears to be inconsistent with their belief in the existence of fantasy characters who possess numerous expectation-defying properties. Nevertheless, the fact that children rarely identify reasons as to why impossible events are impossible might leave them vulnerable to suggestion in the same way that it might leave them vulnerable to magic tricks and visual illusions. That is, serious testimony from a credible adult (e.g., "yes, Virginia, there is a Santa Claus") may be sufficient to convince children of the actuality of these fantasy characters if children have

not identified a reason as to why such characters could not exist in the real world.

To conclude, we would like to raise two questions in need of further research. First, what kind of knowledge is needed to judge improbable events possible? On the one hand, children may need to know more about the specific circumstances that would allow an improbable event to occur. For instances, children may need to know something about construction practices and construction materials in order to affirm the possibility of making a mug-shaped building. On the other hand, children may simply need additional experience reasoning about a particular type of causality (e.g., intentional causality) or a particular type of causal violation (e.g., the creation of unconventional artifacts, buildings or otherwise). Future research could address this question by systematically varying the type of causality violated by the improbable events (e.g., intentional causality, physical causality, biological causality) and comparing children's willingness to affirm the possibility of each class of events.

Second, how consistent are adults' possibility judgments across different individuals and different events? Little is known about how adults judge the possibility of events in domains other than those in which all possibilities can be defined and enumerated (see Bell & Johnson-Laird, 1998; Goldvarg & Johnson-Laird, 2000). Although the adults in Experiment 1 were quite consistent in their judgment of improbable events, they were somewhat inconsistent in their ability to provide factual justifications for their judgments of impossibility. Indeed, two adults provided redundant justifications as often as they provided factual ones. One explanation for this variation is that some adults were more engaged in the task than others. Another explanation is that some adults, like children, continue to mistake their inability to imagine circumstances that would allow an event to occur for evidence that no such circumstances exist. Distinguishing between these two possibilities would not only clarify the endpoint of the developmental trajectory outlined in this article but would also shed light on the question of why adults often disagree about what is possible and what is not.

References

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition, 20,* 191–208.

Bell, V. A., & Johnson-Laird, P. N. (1998). A model theory of modal reasoning. *Cognitive Science, 22,* 25–51.

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

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

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: The MIT Press.

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.

Coates, J. (1988). The acquisition of the meanings of modality in children aged eight and twelve. *Journal of Child Language, 15,* 425–434.

Fabricius, W. V., Sophian, C., & Wellman, H. M. (1987). Young children's sensitivity to logical necessity in their inferential search behavior. *Child Development, 58,* 409–423.

Fay, A. L., & Klahr, D. (1996). Knowing about guessing and guessing about knowing: Preschoolers' understanding of indeterminacy. *Child Development, 67,* 689–716.

Goldvarg, Y., & Johnson-Laird, P. N. (2000). Illusions in modal reasoning. *Memory and Cognition, 28,* 282–294.

Gonsalves, J. (1999). *Relations between conceptual and semantic development: Preschoolers' understanding of modality across linguistic and nonlinguistic domains.* Unpublished doctoral dissertation, Clark University, Worcester, MA.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories.* Cambridge, MA: The MIT Press.

Hirst, W., & Weil, J. (1982). Acquisition of epistemic and deontic meaning of modals. *Journal of Child Language, 9,* 659–666.

Hoemann, H. W., & Ross, B. M. (1971). Children's understanding of probability concepts. *Child Development, 42,* 221–236.

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

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

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

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

Keil, F. C. (1989). *Concepts, kinds, and cognitive development.* Cambridge, MA: The MIT Press.

Klahr, D., & Chen, Z. (2003). Overcoming the positive-capture strategy in young children: Learning about indeterminacy. *Child Development, 74,* 1275–1296.

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

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25*, 265–288.

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

Miller, S. A., Custer, W. L., & Nassau, G. (2000). Children's understanding of the necessity of logically necessary truths. *Cognitive Development, 15*, 383–403.

Needham, A., & Baillargeon, R. (1993). Intuitions about support in 4.5-month-old infants. *Cognition, 47*, 121–148.

Nicholls, J. G., & Thorkildsen, T. A. (1988). Children's distinctions among matters of intellectual convention, logic, fact, and personal preference. *Child Development, 59*, 939–949.

Noveck, I. A., Ho, S., & Sera, M. (1996). Children's understanding of epistemic modals. *Journal of Child Language, 23*, 621–643.

Perkins, M. (1983). *Modal expressions in English*. London: Pinter.

Phelps, K. E., & Woolley, J. D. (1994). The form and function of young children's magical beliefs. *Developmental Psychology, 30*, 385–394.

Piaget, J. (1987). *Possibility and necessity*. Minneapolis: University of Minnesota Press.

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

Prentice, N. M., Manosevitz, M., & Hubbs, L. (1978). Imaginary figures of early childhood: Santa Claus, Easter Bunny, and the Tooth Fairy. *American Journal of Orthopsychiatry, 48*, 618–628.

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

Rosengren, K. S., & Hickling, A. K. (1994). Seeing is believing: Children's explanations of commonplace, magical, and extraordinary transformations. *Child Development, 65*, 1605–1626.

Schult, C. A., & Wellman, H. M. (1997). Explaining human movements and actions: Children's understanding of the limits of psychological explanation. *Cognition, 62*, 291–324.

Sharon, T., & Woolley, J. D. (2004). Do monsters dream? Young children's understanding of the fantasy/reality distinction. *British Journal of Developmental Psychology, 22*, 293–310.

Sobel, D. M. (2004). Exploring the coherence of young children's explanatory abilities: Evidence from generating counterfactuals. *British Journal of Developmental Psychology, 22*, 37–58.

Sommerville, S. C., & Capuani-Shumaker, A. (1984). Logical searches of young children in hiding and finding tasks. *British Journal of Developmental Psychology, 2*, 315–328.

Sophian, C., & Somerville, S. C. (1988). Early developments in logical reasoning: Considering alternative possibilities. *Cognitive Development, 3*, 183–222.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science, 14*, 29–56.

Spelke, E. S., Kestenbaum, R., Simons, D., & Wein, D. (1995). Spatiotemporal continuity, smoothness of motion and object identity in infancy. *The British Journal of Developmental Psychology, 13*, 113–142.

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.

Subbotsky, E. V. (1994). Early rationality and magical thinking in preschoolers: Space and time. *British Journal of Developmental Psychology, 12*, 97–108.

Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology, 43*, 337–375.

Woolley, J. D., Boerger, E. A., & Markman, A. B. (2004). A visit from the Candy Witch: Factors influencing young children's belief in a novel fantastical being. *Developmental Science, 7*, 456–468.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358*, 749–750.