Cognitive Development 24 (2009) 293–309

Contents lists available at ScienceDirect

Cognitive Development

The development of possibility judgment within and across domains

Andrew Shtulman*

Occidental College, Department of Psychology, 1600 Campus Road, Los Angeles, CA 90041, United States

ARTICLE INFO

Keywords:
Possibility judgment
Modal reasoning
Causal reasoning
Conceptual development
Domain-specificity

ABSTRACT

The ability to differentiate possible events from impossible ones is an invaluable skill when reasoning about claims that transcend the perceptual evidence at hand, yet preschool-aged children do not readily make this differentiation when reasoning about physically extraordinary events [Shtulman, A., & Carey, S. (2007). Improbable or impossible? How children reason about the possibility of extraordinary claims. *Child Development*, *78*, 1015–1032]. The present study sought to determine whether this failure stems from deficits in domain-specific knowledge or deficits in the domain-general procedure by which possibility judgments are made. Participants (48 children aged 4-9 years olds and 16 adults) were asked to judge the possibility of three types of extraordinary events – physical, psychological, and biological – and to justify their judgments of impossibility. Developmental changes in the ability to differentiate improbable events from impossible events were qualitatively similar across domains. Moreover, participants' propensity to judge improbable events possible was significantly correlated with the quality of their justifications, both within and across domains. These findings suggest that modal development involves a domain-general change in how modal claims are evaluated.

The fictional detective Sherlock Holmes was fond of a particular saying: "When you have eliminated the impossible, whatever remains, however improbable, must be the truth" (Doyle, 1860/2001). Sherlock often used this maxim to solve crimes, as when he deduced, in *The Sign of Four*, that a murderer must have entered his victim's room through a hole in the roof (an improbable event) given that the only other way to enter the room would have been to walk through a locked door (an impossible

* Tel.: +1 323 259 2633.
E-mail address: shtulman@oxy.edu.

0885-2014/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.cogdev.2008.12.006

event). Although Sherlock did not actually witness the murderer enter the room, he was able to deduce the location of the murderer's entrance from his prior knowledge of physical principles – namely, that one solid object cannot pass through another.

The differentiation of improbable events from impossible events is a form of modal reasoning, or reasoning about whether a particular state of affairs is possible and/or necessary in light of some predefined set of constraints (Hughes & Cresswell, 1968). Like Sherlock, we engage in modal reasoning when attempting to evaluate claims (or hypotheses) that transcend the perceptual evidence at hand and must instead be evaluated on the basis of their consistency with what we know to be true about the world in general. Consider, for instance, the claim that human beings evolved from nonhuman organisms. Although perceptual evidence can be brought to bear on the truth of this claim (e.g., fossil evidence, anatomical homologies, genetic sequencing), all such evidence is indirect; none of us has actually witnessed human evolution firsthand. Thus, in order for this evidence to be deemed plausible, one must first grant that human evolution is, in fact, possible. Similar problems arise when evaluating claims about events that have yet to occur. The claim that human beings will someday colonize Mars, for instance, is amenable to factual considerations only if one grants that extraterrestrial habitation is, in fact, possible.

How do individuals use their prior knowledge to differentiate possible events from impossible ones? Research on this topic has been conducted mainly from a developmental perspective (Browne & Woolley, 2004; Chandler & Lalonde, 1994; Phelps & Woolley, 1994; Rosengren & Hickling, 1994; Sobel, 2004; Subbotsky, 1994; Subbostky, 2004; Woolley & Cox, 2007). It has shown that children differentiate impossible events from ordinary events as early as they can be asked to make such differentiations. For instance, 3-year-old children label impossible physical events, like making an object disappear, as magic but do not label ordinary physical events, like painting an object a different color, as magic (Johnson & Harris, 1994). Likewise, 4-year-old children label impossible biological events, like a dog turning into a puppy, as magic but do not label ordinary biological events, like a puppy turning into a dog, as magic (Rosengren, Kalish, Hickling, & Gelman, 1994).

Many authors have interpreted these findings as evidence that children represent the causal constraints that prevent impossible events from occurring and are explicitly aware of the modal implications of such constraints. As Rosengren and Hickling (1994) put it:

> "The ability to classify events as possible or impossible indicates that children have a relatively good understanding of the causal principles which apply to events in the world . . . [and] realize that causal principles provide one with the knowledge to predict what is possible and also to recognize those events that are not possible" (p. 1625).

Rosengren and Hickling's claim, although consistent with the findings described above, is not the only explanation for those findings. An alternative explanation is that young children deny the possibility of any event that defies intuition, regardless of whether that intuition is grounded in nomologically relevant causal principles (i.e., the kinds of principles that purportedly underlie adults' modal judgments). For instance, a child who claims that people cannot walk through walls may do so either on the basis of knowing that one object cannot pass through another or on the basis of *not knowing* how such an event might transpire. In other words, young children may deny an event's possibility not because they can think of a causal principle that would prevent the event from occurring but because they are unable to think of circumstances that would allow the event to occur.

To test this hypothesis, Shtulman and Carey (2007) asked 4-, 6-, and 8-year olds to judge the possibility of two types of events: those that adults judge impossible, such as walking through a wall or walking on water, and those that adults judge improbable but not impossible, such as finding an alligator under the bed or making pickle-flavored ice cream. Consistent with previous findings (Johnson & Harris, 1994; Rosengren et al., 1994), children of all ages reliably denied the possibility of physically impossible events. They did not, however, reliably *affirm* the possibility of physically improbable events. Instead, they treated improbable events like impossible events, denying the possibility of both. This effect, though most pronounced among 4-year olds, was observed even among 8-year olds, suggesting that children's modal judgments are initially based on intuitions more superficial than the identification of a nomologically relevant causal principle.

Additional evidence in support of this claim can be found in the difference between children's and adults' justifications for their judgments of impossibility. Whereas adults typically appealed to facts about the world that would preclude an event's occurrence (e.g., walking through a wall is impossible because "walls are solid"), children rarely did so. Instead, they either restated their initial judgment (e.g., walking through a wall is impossible because "you can't walk through walls") or appealed to hypothetical events that could occur, or would occur, in place of the target event (e.g., "you couldn't walk through a wall, but you could walk through a door in the wall" or "if you tried to walk through a wall, you would hit your head"). From age four to age eight, children's propensity to provide factual justifications increased as their propensity to provide hypothetical justifications decreased, but even 8-year olds provided significantly fewer factual justifications (and significantly more hypothetical justifications) than did adults.

From these findings, Shtulman and Carey (2007) argued that children and adults use different standards of evidence for judging an event impossible. Whereas adults deny an event's possibility only if they can identify a fact about the world that would preclude the event's occurrence, children deny an event's possibility whenever they fail to imagine circumstances that would allow the event to occur. Put differently, children commit the fallacy described by Dennett (1991) as "mistaking a failure of imagination for an insight into necessity" (p. 401). This proposal is similar to a proposal made by Piaget (1987) for why children do not initially differentiate empirical truths, like the fact that all crows are black, from necessary truths, like the fact that all triangles are three-sided (see also Miller, 1986; Miller, Custer, & Nassau, 2000). According to Piaget, children initially assume that anything that is actual is necessary and must learn to differentiate positive evidence of necessity (i.e., a scientific law, logical proof, or theoretical model that mandates an event's occurrence) from ambiguous evidence of necessity (i.e., repeated observations of the same event).

The hypothesized difference between children's and adults' possibility-judgment strategies is perhaps best understood in terms of the roles that *intuition* and *reflection* play in the process of deriving a modal judgment. This difference is illustrated in Fig. 1, where Model 1 represents the typical child strategy and Model 2 represents the typical adult strategy. Both models assume that the reasoner's final modal judgment is informed by an initial modal intuition (which, in the case of a novel expectation-defying event, would be "impossible"). However, the pathway between modal intuition and modal judgment is mediated by reflection on the legitimacy of one's intuition in Model 2 but not in Model 1. Indeed, in Model 1, modal intuition leads directly to modal judgment, and reflection occurs only as an

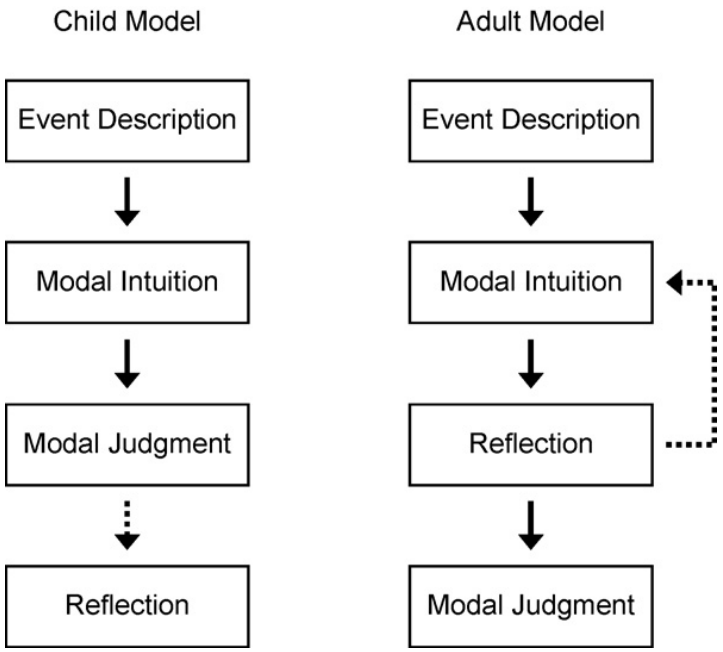

Fig. 1. Models of children's and adults' possibility-judgment strategies. Solid arrows depict obligatory connections and dashed arrows depict optional ones. In the child model, reflection takes the form of post-hoc rationalization, whereas, in the adult model, reflection mediates the relationship between modal intuition and modal judgment.

afterthought, if at all. In short, children are assumed to differ from adults in the extent to which they check their modal intuitions against explicitly known causal principles before pronouncing an event "impossible."

An important assumption underlying this proposal is that children's initial failure to differentiate improbable events from impossible events (and probable events from necessary events) is a *procedural* error, not an *informational* error. In other words, children are purported to differ from adults not in the information they possess but in how they apply that information to modal claims. Clearly, children are less knowledgeable about the world than adults are, yet such differences cannot account for the full pattern of results obtained by Shtulman and Carey (2007), among others (e.g., Piaget, 1987; Phelps & Woolley, 1994). In particular, they cannot account for (1) the fact that children deny the possibility of improbable events involving forms of causation they likely know and understand (e.g., painting polka dots on an airplane); (2) the fact that children justify their judgments of impossibility in ways that adults rarely do (e.g., by appealing to hypothetical events that could occur, or would occur, in place of the target event under consideration); and (3) the fact that children err on the side of judging possible events impossible rather than the side of judging impossible events possible.

Stronger evidence in support of a processing account of children's modal development would be to show that developmental differences in children's modal judgments and modal justifications are consistent across different domains of knowledge, even those that are impoverished compared to physics (the domain of choice in past studies of modal development). In other words, if children exhibit similar patterns of modal judgment and modal justification in domains for which they have varying amounts knowledge, then the claim that modal development can be reduced to knowledge acquisition would be undermined. The present study sought to address this issue by extending Shtulman and Carey's (2007) research on how children reason about extraordinary events from the domain of physics to the domains of psychology and biology as well. Children show an appreciation of core physical concepts, such as *solidity* and *support*, in infancy (Baillargeon, 2004; Spelke, 1990), but they do not show an appreciation of core psychological concepts, such as *knowledge* and *belief*, until around the age of four (Bartsch & Wellman, 1995; Perner, 1993), and they do not show an appreciation of core biological concepts, such as *growth* and *death*, until around the age of six (Carey, 1985; Hatano & Inagaki, 1994). Thus, it is quite plausible that children are able to differentiate improbable events from impossible events in the domain of physics before they are able to do so in the domains of psychology and biology. Similarly, children's justifications for their judgments regarding physical possibility may bear little resemblance to those regarding psychological or biological possibility.

One reason to think that modal development is *not* uniform across domains is that previous research by Kalish (1998) and Browne and Woolley (2004) has shown that young children reason about violations of physical laws differently than they reason about violations of social conventions. For instance, preschoolers appeal to causes and constraints when explaining why people conform to physical laws but appeal to permissions and obligations when explaining why people conform to social conventions (Kalish, 1998). Moreover, when preschoolers are presented with pairs of events – one violating a physical law (e.g., floating in the air, walking on the ceiling) and one violating a social convention (e.g., wearing shoes in the bathtub, singing "Jingle Bells" at a birthday party) – they tend to associate the former, but not the latter, with magic (Browne & Woolley, 2004). These findings are problematic for the claim that children's modal development is domain-general in nature insofar as children's reasoning about the necessity of different empirical regularities is not constant across domains. That said, the studies by Kalish (1998) and by Browne and Woolley (2004) were designed not to assess children's understanding of the difference between possible and impossible events but to assess their understanding of the difference between voluntary and automatic conformity. As a result, they involved events that differed not only by domain (physics vs. psychology) but also by modal status (impossible vs. improbable). Whether children differentiate impossible events from improbable events in some domains earlier than others is thus an open question.

In short, the present study sought to determine whether the development of adult-like modal judgments is domain-specific or domain-general in nature. If, as suggested above, differences between children's and adults' modal judgments are due mainly to procedural differences in how those judgments are made, then children should exhibit similar patterns of judgment (and similar patterns of justification) in all domains tested. If, on the other hand, such differences are due mainly to differences

in content knowledge, then children should exhibit different patterns of judgment in different domains of knowledge – e.g., differentiating improbable events from impossible events in some domains earlier than others or failing to differentiate improbable from impossible events in some domains altogether.

Eighteen extraordinary events were described to participants in the form of a story, six involving physical concepts, six psychological concepts, and six biological concepts. Because children acquire a rudimentary knowledge of physics before they acquire a rudimentary knowledge of psychology and acquire a rudimentary knowledge of psychology before they acquire a rudimentary knowledge of biology, it was predicted that, if children differentiated improbable events from impossible events at different rates in different domains, they would do so first in the domain of physics, second in the domain of psychology, and third in the domain of biology. Although the specificity of this prediction may have been unwarranted given that children acquire different concepts within each domain at different times, it served as the basis for selecting these particular domains.

1. Method

1.1. Participants

Participants were 48 children and 16 adults. The adults were recruited from introductory psychology classes at a large, northeastern US university. The children were recruited from various daycare centers in the greater Boston area. For data analysis, children were divided into three groups of 16: 4–5-year olds (M = 4–11, range 4-0 to 5-11), 6–7-year olds (M = 6–9, range 6-1 to 7-6), and 8–9-year olds (M = 8–6, range 7-7 to 9-8). These participants were predominantly white and predominately middleclass, although a range of ethnic and socioeconomic backgrounds were represented. Each age group included an approximately equal number of males and females.

1.2. Materials

The six concepts used to represent the domain of physics were *conservation*, *contact*, *entropy*, *mass*, *support*, and *solidity*; the six used to represent the domain of psychology were *attention*, *audition*, *communication*, *knowledge*, *speech*, and *vision*; and the six used to represent the domain of biology were *growth*, *nourishment*, *lifespan*, *reproduction*, *sleep*, and *vitality*. One improbable event and one impossible event were constructed to represent each concept. For example, the improbable event of reading someone's lips and the impossible event of reading someone's thoughts were constructed to represent the concept *communication*. The improbable event of living for 120 years and the impossible event of living for a thousand years were constructed to represent the concept *lifespan*. And the improbable event of walking on a telephone wire and the impossible event of walking on water were constructed to represent the concept *support*.

At two events per concept and six concepts per domain, a total of 36 events were constructed. These appear in Table A1 through Table A3, where they are organized by domain, concept, and event type. These events were integrated into one of two stories ("Story 1" or "Story 2") such that each story contained three improbable events and three impossible events from each domain, for a total of 18. Both stories, which were each entitled "Detective Blunder and the Case of the Missing Poodle," followed the adventures of a pet detective and his assistant. Improbable events and impossible events involving the same concept did not appear in the same story so as to minimize demand characteristics associated with asking participants to make multiple judgments on the same topic.

To prevent response biases, the 18 extraordinary events were intermixed with six ordinary events: owning a dog, smoking a pipe, losing one's keys, working in an office, wearing a hooded sweatshirt, and getting a tattoo. Most participants claimed to have experienced (or witnessed) all six ordinary events, and those who claimed not to have experienced one or more of the events almost always judged them possible. Thus, participants' possibility judgments for the ordinary events are not discussed further.

The storybooks themselves consisted of 21 double-sided pages, 8.5 inches long and 11 inches wide. Accompanying each page of text was a photograph illustrating the events described on that page. None of the extraordinary events were explicitly depicted in the accompanying illustrations. Instead, these illustrations depicted either an ordinary event (e.g., a man with a tattoo) or the outcome of an

extraordinary event (e.g., a man holding an umbrella described as having "appeared out of thin air"). The purpose of the photographs was not to convey additional information about these events but to hold participants' interest. In fact, the photographs were designed to be as uninformative as possible, with most (65%) depicting strictly ordinary situations.

1.3. Procedure

Half of the children in each age group were read Story 1 and half were read Story 2. Children were read the entire story at the beginning of each interview. They were then asked whether or not they had experienced the story's six ordinary events and 18 extraordinary events in the order they appeared in the story (e.g., "Have you ever seen a person catch a fly with chopsticks?"). 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 catch a fly with chopsticks in real life?"). Whenever they denied that an event was possible, they were asked to provide a justification for their judgment (e.g., "Why couldn't a person catch a fly with chopsticks in real life?"). Adult participants were asked the same questions as children but in the form of a questionnaire rather than an interview.

1.4. Coding

Participants provided a total of 736 justifications for why certain events should be considered impossible. Following Shtulman and Carey (2007), these justifications were sorted into three categories: factual justifications, hypothetical justifications, and redundant justifications. Factual justifications referenced facts about the world that would prevent the target event from actually occurring (e.g., walking on water is impossible because "water is a liquid, not a solid"). Hypothetical justifications referenced hypothetical events that could occur, or would occur, in place of the target event (e.g., walking on water is impossible because "they could walk on a bridge" or because "they would sink"). Lastly, redundant justifications provided no information beyond what was already discernable from the participant's initial judgment (e.g., "it's not possible," "it's not real," "it can only happen in stories," "I don't know"). Justifications that referenced magic were also included in this category, as these justifications comprised less than 3% of all justifications in total.

The distinction between factual and hypothetical justifications is further illustrated in Table 1, which provides an example of each type of justification for three improbable events (catching a fly with chopsticks, reading someone's lips, living without a nose) and three impossible events (catching a shadow, reading someone's thoughts, living without a heart). Although factual and hypothetical justifications share certain similarities in content, they were separated for two reasons. First, hypothetical justifications differed linguistically from factual justifications in that only hypothetical justifications included

Table 1
Sample factual and hypothetical justifications given in response to the same event.

Event	Type	Example
Catching a fly with chopsticks	Factual	"Flies are too fast to catch"
	Hypothetical	"You could catch it with your hand"
Catching a shadow	Factual	"Shadows don't have any mass"
	Hypothetical	"You could catch what made the shadow"
Reading someone's lips	Factual	"People talk too fast to keep track"
	Hypothetical	"If they had writing on their lips"
Reading someone's thoughts	Factual	"You can't look into a person's mind"
	Hypothetical	"You could guess what they're thinking"
Living without a nose	Factual	"Noses do not just stop working"
	Hypothetical	"You wouldn't be able to smell"
Living without a heart	Factual	"You need a heart to pump blood"
	Hypothetical	"You would die in five minutes"

conditional verbs, like *would*, *could*, or *should*. Second, hypothetical justifications differed conceptually from factual justifications in that hypothetical justifications did not actually answer the question at hand – namely, why is the target event impossible? Rather, hypothetical justifications 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 river without attempting to walk across its surface) or how a different outcome could be achieved by the same means (e.g., how a person could attempt to walk across the surface of a river without getting to the other side). For these reasons, hypothetical justifications were considered less epistemologically sound than factual justifications and were thus segregated from them.

The reliability of the justification coding scheme was assessed by comparing the exhaustive classifications of two independent coders, each of whom was blind to the age of the participants who had provided the justifications. Overall agreement between coders was 88%, and all disagreements were resolved via discussion.

2. Results

2.1. Experience judgments

As expected, participants of all ages claimed to have experienced few to none of the extraordinary events. On average, 4–5-year olds claimed to have experienced 0.7 improbable events (S.D. = 0.9) and 0.3 impossible events (S.D. = 0.8); 6–7-year olds claimed to have experienced 0.6 improbable events (S.D. = 0.6) and 0.3 impossible events (S.D. = 0.4); 8–9-year olds claimed to have experienced 1.1 improbable events (S.D. = 1.3) and 0.3 impossible events (S.D. = 0.8); and adults claimed to have experienced 1.9 improbable events (S.D. = 1.0) and 0.2 impossible events (S.D. = 0.4). The improbable events that participants claimed to have experienced most often were reading someone's lips (which 45% of the "Story 1" participants claimed to have experienced/witnessed) and reading a book that's upside down (which 39% of the "Story 2" participants claimed to have experienced/witnessed).

A 4×2 repeated-measures analysis of variance (ANOVA) was used to test the effects of age and event type on experience judgments. This analysis revealed no effect of age but a significant effect of event type, $F(1, 60) = 58.61$, $p < 0.001$, and a significant interaction between age and event type, $F(3, 60) = 8.56$, $p < 0.001$. Contrast analyses revealed that experience with the improbable events increased linearly with age, $F(1, 60) = 5.79$, $p < 0.01$, but that experience with the impossible events did not. Apparently, some of the variation in children's ability to judge improbable events possible was due to differences in their experience with those events. Nevertheless, the overall magnitude of that difference was small (i.e., less than a single event).

2.2. Possibility judgments

The mean number of extraordinary events judged possible by each age group is displayed in Fig. 2 as a function of domain and event type. Participants of all ages consistently denied the possibility of impossible events. Only adults, however, consistently *affirmed* the possibility of improbable events. This pattern of results was obtained not only in the domain of physics but also in the domains of psychology and biology, indicating that children's initial reluctance to judge improbable events possible is domain-general. Indeed, the rate at which children began affirming the possibility of improbable events was highly similar across domains.

A $4 \times 3 \times 2$ repeated-measures ANOVA was used to test the effects of age group, domain (physics, psychology, biology), and event type (improbable, impossible) on possibility judgments, the latter two factors analyzed within participants. This analysis revealed significant effects of age, $F(3, 60) = 14.43$, $p < 0.001$, domain, $F(2, 120) = 13.57$, $p < 0.001$, and event type, $F(1, 60) = 379.96$, $p < 0.001$, as well as a significant interaction between age and event type $F(3, 60) = 38.60$, $p < 0.001$. No significant interactions were found between domain and event type or between domain and age, implying that children's ability to differentiate improbable events from impossible events developed at similar rates across the three domains.

The interaction between age and event type was explored with univariate tests of the effect of age on possibility judgments for each type of event. These analyses revealed that possibility judgments varied

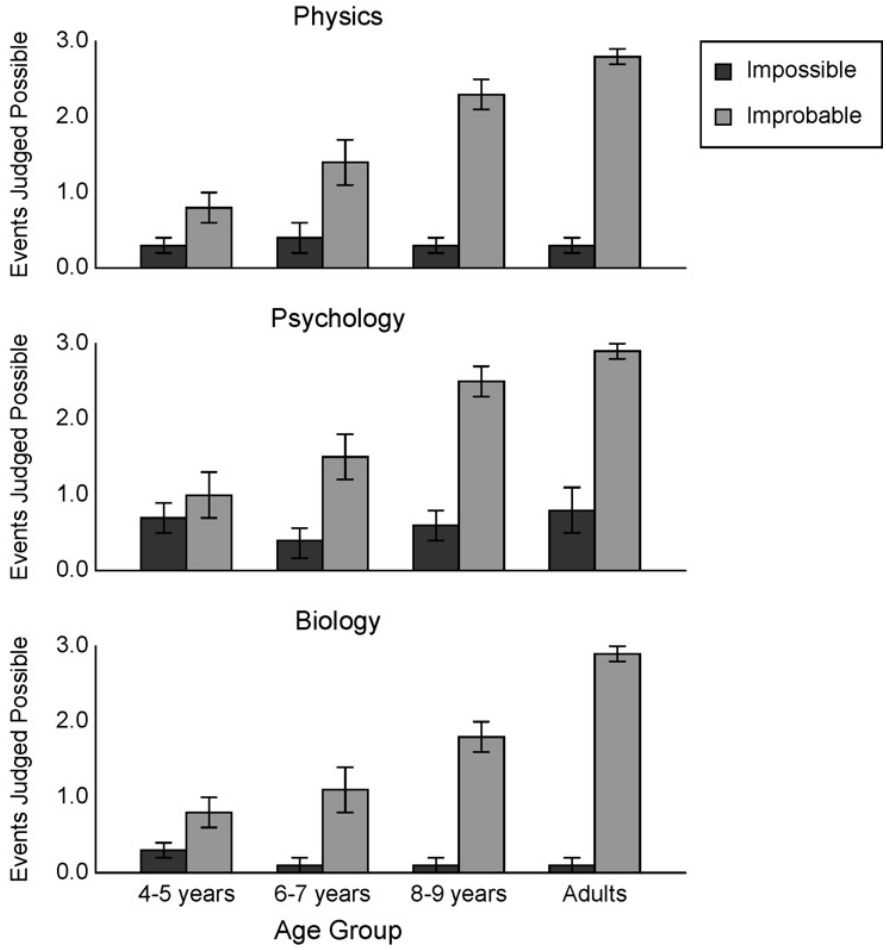

Fig. 2. The mean number of events judged possible by each age group (+S.E.) as a function of domain (psychology, biology, physics) and event type (impossible, improbable).

by age for the improbable events, $F(3, 60) = 29.85$, $p < 0.001$, but not the impossible events. Contrast analyses further revealed that propensity to judge improbable events possible increased linearly with age for all three domains – physics, $F(1, 60) = 39.90$, $p < 0.001$; psychology, $F(1, 60) = 35.11$, $p < 0.001$; and biology, $F(1, 60) = 39.90$, $p < 0.001$. These analyses remained significant when adults were excluded from the dataset, indicating that older children were more likely than younger children to judge improbable events possible in all three domains.

To what extent are the group means displayed in Fig. 2 representative of the judgments of individual participants? This question was addressed by calculating the total number of improbable events judged possible across domains and comparing those totals across age groups (see Table 2). Similar to the findings of Shtulman and Carey (2007), many of the youngest children (38%) judged improbable events possible significantly less often than predicted by chance. That is, many 4–5-year olds judged none (or almost none) of the improbable events possible. Adult participants, in contrast, reliably judged all (or

Table 2
The number of participants in each age group who judged between 0 and 9 improbable events possible.

Age group	Number of improbable events judged possible				
	0–1*	2–3	4–5	6–7	8–9*
4- and 5-year olds	6	5	4	1	0
6- and 7-year olds	3	5	3	3	2
8- and 9-year olds	0	1	2	8	5
Adults	0	0	0	0	16

* Frequencies with a binomial probability of $p < 0.05$.

almost all) of the improbable events possible. Older children tended to demonstrate response patterns intermediate between these two extremes, with only 9% of 6–9-year olds judging none (or almost none) of the improbable events possible and only 19% judging all (or almost all) of such events possible.

On average, participants judged 1.8 of the six biological events possible (S.D. = 1.2), 2.1 of the six physical events possible (S.D. = 1.4), and 2.6 of the six psychological events possible (S.D. = 1.6) – a difference of less than one event but a reliable difference nonetheless. Children did not, however, differentiate improbable events from impossible events in one domain earlier than another, as evidenced by the lack of interaction between domain and event type in the repeated-measures ANOVA reported above. Apparently, the psychological events (as a whole) were more plausible than the physical events (as a whole), which, in turn, were more plausible than the biological events (as a whole). This pattern held for both adults and children, which suggests that participants of all ages may have interpreted the psychological events more metaphorically than they interpreted either the physical events or the biological events. For instance, events like "reading someone's thoughts" or "speaking two languages at the same time" may have been more amenable to metaphoric interpretation than events like "unscrambling a scrambled egg" or "not eating for ten months."

2.3. Justifications for judgments of impossibility

Participants provided anywhere between 5 and 18 justifications for events judged impossible. Justifications were coded as factual, hypothetical, or redundant, as described earlier. Because participants provided different numbers of justifications, absolute frequencies had to be converted to relative frequencies (the number of times a participant provided a type of justification divided by the total number of justifications provided). The mean proportion of factual, hypothetical, and redundant justifications provided by each age group is displayed in Fig. 3 by domain. Because relative frequencies were computed separately for each domain, the means displayed in Fig. 3 constitute three independent measures of propensity to provide each type of justification. For each domain, propensity to provide factual justifications increased with age as propensity to provide redundant and hypothetical justifications decreased with age.

The relative frequencies displayed in Fig. 3 were analyzed for effects of age and justification type with three repeated-measures ANOVAs, one for each domain. These analyses revealed significant interactions between age and justification type for all three domains – physics, $F(6, 120) = 5.35$, $p < 0.001$; psychology, $F(6, 120) = 2.61$, $p < 0.05$; and biology, $F(6, 120) = 6.28$, $p < 0.001$. These interactions were explored with contrast analyses in which age was the independent variable and justification frequencies were the dependent variables. The results of these analyses are displayed in Table 3 as a function of justification type and domain. For all three domains, factual justifications increased linearly with age and redundant justifications decreased linearly with age. Hypothetical justifications also decreased with age, but this difference was statistically significant only for the domain of physics.

Table 3
Linear contrasts of justification scores for each type of justification and each domain of knowledge (d.f. = 60 for all participants, 45 for children only).

Justification type	Domain	*F*-value	
		All participants	Children only
Factual	Physics	23.69**	4.11*
	Psychology	15.11**	5.80*
	Biology	24.39**	5.46*
Hypothetical	Physics	−6.20*	−0.01
	Psychology	−1.86	−0.02
	Biology	−0.47	−1.41
Redundant	Physics	−8.17**	−4.55*
	Psychology	−4.63*	−2.23
	Biology	−26.63**	−14.49**

* $p < 0.05$.
** $p < 0.01$.

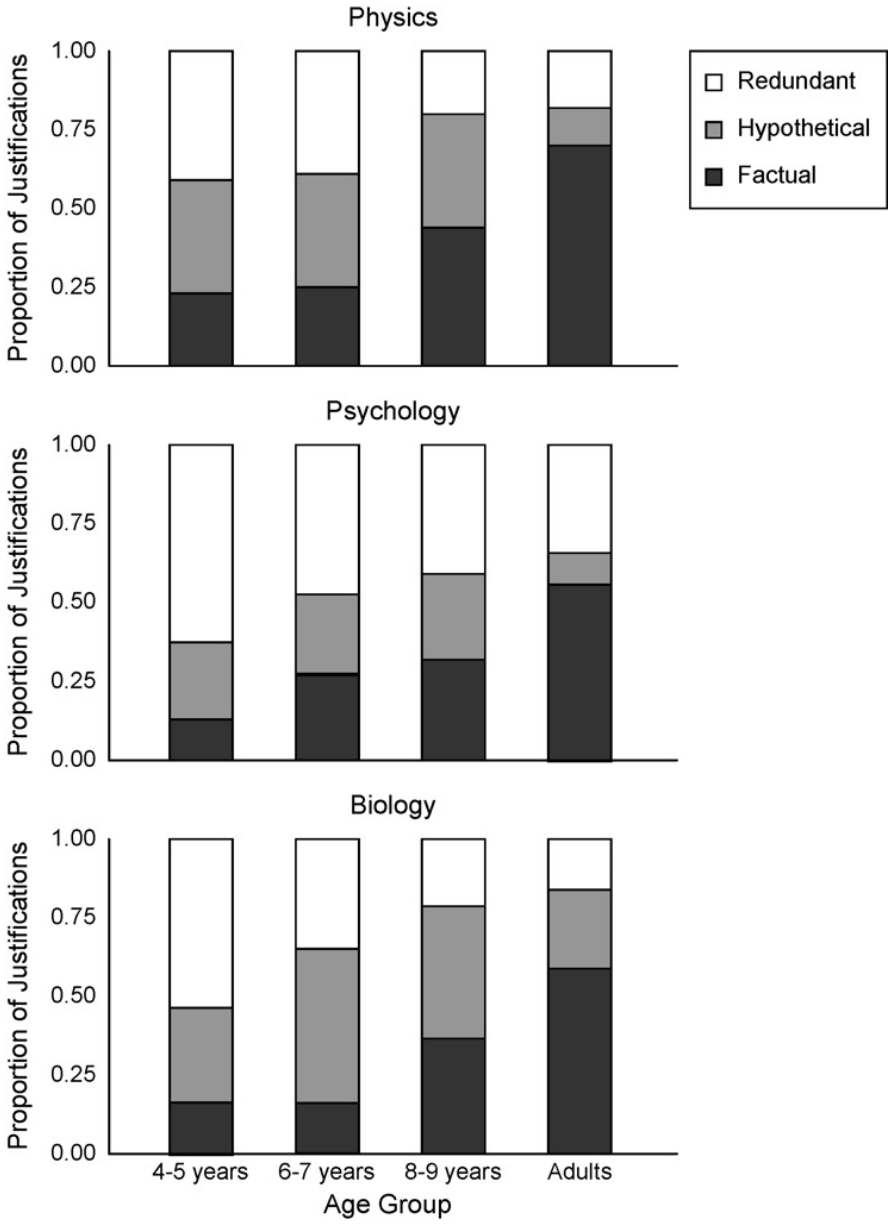

Fig. 3. The mean proportion of factual, hypothetical, and redundant justifications provided by each age group as a function of domain (psychology, biology, physics).

To determine whether these findings were driven mainly by differences between children and adults (as opposed to differences between all four age groups), adults were excluded from the dataset, and the analyses were repeated. This second set of analyses closely mirrored the first, indicating that older children were generally more likely to provide factual justifications – and generally less likely to provide redundant justifications – than younger children were. (Why older children were no less likely than younger children to provide hypothetical justifications is an issue addressed below.) In sum, the domain-general increase in children's propensity to judge improbable events possible was accompanied by a domain-general increase in children's ability to provide non-redundant, factual justifications for those judgments.

One noticeable difference in justification frequencies across domains is that participants provided more hypothetical justifications for biological events (*M* proportion = 0.37) than for physical events (*M* proportion = 0.29) or psychological events (*M* proportion = 0.22). This difference is due to the fact that participants frequently cited the same justification for why many biological events were impossible: "you would die." This justification was classified as hypothetical, as it describes a hypothetical event

Table 4
Correlations between propensity to judge improbable events possible and propensity to provide factual justifications for judgments of impossibility, both within and across domains.

Measure	1	2	3	4	5	6
All participants						
1. Physical judgments	1.0	0.66**	0.71**	0.49**	0.34**	0.49**
2. Psychological judgments		1.0	0.59**	0.50**	0.34**	0.49**
3. Biological judgments			1.0	0.59**	0.28*	0.54**
4. Physical justifications				1.0	0.26*	0.64**
5. Psychological justifications					1.0	0.34**
6. Biological justifications						1.0
Children only						
1. Physical judgments	1.0	0.58**	0.59**	0.41**	0.26	0.42**
2. Psychological judgments		1.0	0.44**	0.38**	0.30*	0.41**
3. Biological judgments			1.0	0.45**	0.09	0.43**
4. Physical justifications				1.0	0.29*	0.59**
5. Psychological justifications					1.0	0.25
6. Biological justifications						1.0

* $p < 0.05$.
** $p < 0.01$.

that would occur in place of the target event rather than a fact about the world that would preclude the target event from occurring (e.g., not eating for 10 months is impossible because "food is needed to provide the body with energy"). Nevertheless, the "you would die" justification was popular even among adults, suggesting that this justification may have been better coded as "factual." Indeed, doing so would reduce the average proportion of hypothetical justifications to total justifications in the domain of biology from 0.37 to 0.25 – a proportion more consistent with those obtained in the domains of physics and psychology – without changing the overall trends displayed in Fig. 3.

2.4. Correlations between judgments and justifications

Were participants' justifications predictive of their judgments? To address this question, the number of improbable events a participant judged possible in each domain was entered into a correlation matrix alongside the proportion of factual justifications he or she provided for each domain. (Note that these two measures are independent, as children were asked to justify only those events they judged impossible.) These correlations are displayed in Table 4. The top half of Table 4 displays correlations between adult-like judgments and adult-like justifications for all participants (adults included); the bottom half displays these same correlations for children only. As can be seen from both sets of correlations, participants' justifications were highly predictive of their judgments. Within domains, the mean correlation between number of improbable events judged possible and the proportion of factual justifications to total justifications was 0.46 for all participants and 0.38 for children alone. Across domains, these same means were 0.45 for all participants and 0.34 for children alone.

To determine whether the intercorrelations displayed in Table 4 reflect a single dimension or multiple dimensions underlying performance, the six intercorrelated variables were entered into a factor analysis (with all participants included). This analysis yielded one, and only one, factor with an Eigenvalue greater than 1.0 – i.e., one, and only one, factor capable of explaining more variance in participants' responses than any single variable on its own. That factor, which explained a total of 58% of the variance in participants' responses, is perhaps best interpreted as "strategic orientation." Participants on one end of the "strategic orientation" dimension denied the possibility of only those events for which they could identify a fact about the world that would preclude the event's occurrence; participants on the other end denied the possibility of any event that defied intuition, regardless of whether that event defied a nomologically relevant causal principle. It should be noted that a one-factor solution was obtained even when adults were excluded from the dataset, indicating that older children differed from younger children along the same dimension that children, as a whole, differed from adults.

A final set of analyses addressed whether children's propensity to judge improbable events possible was correlated with their propensity to provide different types of justifications independent of age. First-order correlations between possibility judgments and justification frequencies revealed that, across domains, children's propensity to judge improbable events possible was positively correlated with their propensity to provide factual justifications ($r = 0.64$, $p < 0.001$), negatively correlated with their propensity to provide hypothetical justifications ($r = -0.35$, $p < 0.01$) and negatively correlated with their propensity to provide redundant justifications ($r = -0.40$, $p < 0.01$). Two of these correlations remained significant even after controlling for age in months: factual justifications ($r = 0.33$, $p < 0.05$) and hypothetical justifications ($r = -0.32$, $p < 0.05$). Thus, consistent with the claim that differences in possibility judgment reflect differences in how those judgments are made, children's ability to judge improbable events possible was correlated with their ability to provide epistemologically sound justifications (i.e., factual justifications) in lieu of epistemologically *unsound* justifications (i.e., hypothetical justifications) independent of age.

This last set of findings hints at an explanation for the absence of an age effect in children's production of hypothetical justifications. Whereas redundant justifications may have been the justification of choice for younger children who failed to differentiate improbable events from impossible ones, hypothetical justifications may have been the justification of choice for older children who failed to differentiate improbable events from impossible ones. To test this hypothesis, the eight children in each age group who judged the fewest improbable events possible were compared to the remaining eight in terms of their propensity to provide (a) redundant justifications and (b) hypothetical justifications, averaged across domains. This analysis revealed that, in the youngest age group, those who judged fewest improbable events possible were marginally more likely than their peers to provide redundant justifications, $M = 0.63$ vs. $M = 0.41$, $t(14) = 1.87$, $p = .08$, but no more likely than their peers to provide hypothetical justifications ($M = 0.29$ vs. $M = 0.32$, $t(14) = -0.32$, *ns*). In the oldest age group, in contrast, those who judged fewest improbable events possible were significantly more likely than their peers to provide hypothetical justifications, $M = 0.45$ vs. $M = 0.27$, $t(14) = 2.35$, $p < 0.05$, but no more likely than their peers to provide redundant justifications $M = 0.20$ vs. $M = 0.34$, $t(14) = -1.44$, *ns* (children in the middle age group exhibited a pattern intermediate to these two extremes). Thus, older children who denied the possibility of many, or most, improbable events tended to appeal to new (yet irrelevant) information rather than merely restate their original judgment.

3. Discussion

As producers and consumers of testimony, we regularly encounter claims that cannot be verified on the basis of personal observation (Harris and Koenig, 2006). Our decision whether or not to accept such claims as true rests, in part, on our decision whether or not to accept the events they describe as possible. Previous research on the development of possibility-judgment strategies (Shtulman & Carey, 2007) showed that preschool-aged children readily differentiate ordinary events from impossible events but do not readily differentiate *improbable* events from impossible events. The present study sought to determine whether this developmental change in modal reasoning is domain-specific or domain-general in nature, as this distinction bears on the question of whether children's modal judgments differ from those of adults because children possess less content knowledge than adults or because children evaluate modal claims in a different manner than adults.

Replicating previous research, many of the youngest children in the present study failed to differentiate between improbable events, like walking on a telephone wire or living without a nose, from impossible events, like walking on water or living without a heart, claiming that both types of events are impossible. Older children were more likely than younger children to accept the possibility of improbable events, but even 8- and 9-year olds tended to deny the possibility of at least some improbable events. Extending previous research, this same pattern of results was obtained in three domains of knowledge – physics, psychology, and biology. Moreover, children's judgments were correlated with their justifications such that children who consistently denied the possibility of improbable events provided less factual support for their judgments than those who consistently *affirmed* the possibility of improbable events. In other words, children who justified their judgments of impossibility

with facts about the world that would preclude an event's occurrence (e.g., unscrambling a scrambled egg is impossible because "cooking something changes it for good") tended to deny the possibility of *fewer* improbable events than children who justified their judgments of impossibility redundantly (e.g., "that's impossible") or hypothetically (e.g., "you could get a new egg").

Taken together, these findings suggest that *how* domain-specific knowledge is recruited in the service of evaluating modal claims is as important, if not more important, than *what* that knowledge consists of. Across three domains of knowledge as diverse in content as possible, younger children consistently differed from older children in (a) their ability to discriminate between violations of necessary truths and violations of empirical truths, and (b) their ability to justify their discriminations in an epistemologically sound manner. The generality of these findings suggests that they are best explained in terms of a procedural account of modal development, like that depicted in Fig. 1. Although the account depicted in Fig. 1 is not the only account consistent with these findings, it is perhaps the simplest account one could offer. Additional considerations in favor of this account are discussed below, following discussion of how the present findings relate to previous findings on the domain-specificity of conceptual development.

3.1. Relation to previous findings

Noted earlier, there are at least two reasons to have expected children's modal judgments and modal justifications to have differed across domains: (1) previous research on the acquisition of domain-specific concepts has revealed robust developmental lags in the acquisition of biological and psychological concepts relative to physical concepts (Carey, 2009; Wellman & Gelman, 1992); and (2) previous research on children's understanding of the distinction between physical laws and social conventions has found evidence of such understanding in children as young as 3 years of age (Browne & Woolley, 2004; Kalish, 1998). How can the apparent discrepancy between these findings and the present findings be reconciled?

One could attempt to reconcile the first discrepancy by arguing that the absence of a domain effect in the present study should be viewed as evidence that the events used as stimuli were not actually representative of the domains they were intended to represent. In other words, the data displayed in Figs. 2 and 3 should be viewed as a null result, not a positive finding. Although this hypothesis cannot be dismissed on the basis of the data itself, a proponent of this hypothesis would need to give an account of why the events listed in Table A1 through Table A3 constitute a biased sample of all possible events from the domains of interest, as well as an account of why children's justifications bore a meaningful (and statistically significant) relationship to their judgments, both within and across domains.

An alternative explanation for the absence of a domain effect on children's propensity to judge improbable events possible is that explicit knowledge of domain-specific causal principles is initially irrelevant to the task of possibility judgment. That is, if young children base their possibility judgments on superficial considerations, such as an event's ease of imaginability or degree of peculiarity, whether or not those considerations are grounded in sound causal principles is of little consequence. However, by the time children stop relying on superficial considerations and begin relying on sound causal principles (i.e., around age six or seven), they will have acquired enough biological and psychological knowledge to obscure any differences in possibility judgment that would have been evident at an earlier age had they then relied on causal principles to make their judgments. In other words, discrepancies in the acquisition of domain-specific knowledge are likely to be resolved prior to when those discrepancies might influence children's possibility judgments.

Turning to the second issue, the present finding that children reason about physically extraordinary events similarly to how they reason about psychologically extraordinary events is not necessarily inconsistent with the finding that preschoolers understand that social conventions, but not physical laws, can be violated. Studies that have demonstrated the latter tested preschoolers on their ability to differentiate violations of physical law (e.g., walking through a wall) from violations of social convention (e.g., wearing pajamas to school) but did not test preschoolers on their ability to differentiate laws from conventions in general. In other words, the distinction between laws, or necessary truths, and conventions, or empirical truths, was confounded with the distinction between physical causation and mental causation.

When these two distinctions were teased apart – as was done in the present study (and was partially done in Experiment 2 of Browne & Woolley, 2004) – children's ability to differentiate violations of physical regularities from violations of psychological regularities was tied more closely to the violations' *modal status* (impossible vs. improbable) than to the violations' *content* (physics vs. psychology). This finding was true of even the youngest age group in the present study. Although many of these children failed to differentiate improbable events from impossible events as individuals (Table 2), they did manage to affirm the possibility of more improbable events than impossible events as a group ($M = 2.6$ vs. $M = 1.3$, $t(15) = 2.84$, $p < 0.05$). Thus, preschoolers' sensitivity to the distinction between physical laws and social conventions may be symptomatic of a more general sensitivity to the difference between necessary truths and empirical truths, albeit a weak one.

3.2. Potential sources of modal development

Whether or not the youngest children in the present study reliably differentiated improbable events from impossible events as a group, their overall performance was still profoundly different from that of older children and adults. What factors might be responsible for this difference? Previous research on the development of strategic knowledge (Crowley, Shrager, & Siegler, 1997; Siegler, 1996) points to the possibility that mature modal reasoning requires some degree of metacognitive awareness. In other words, the ability to make sound modal judgments may require awareness of the task of modal judgment itself. A precedent for this kind of development can be found in the work of Kuhn and Pearsall (1998), who showed that children's mastery of a multivariable causal induction task was correlated with their metastrategic understanding of that task, and in the work of Rittle-Johnson, Siegler, and Alibali (2001), who showed that children's procedural knowledge of fractional division was predicted by, and predictive of, their conceptual knowledge of the same domain. Precedents aside, the task of possibility judgment is less procedurally complicated than either causal induction or fractional division, so it is unclear what would constitute metastrategic knowledge of that task or even whether adults possess such knowledge.

A second, more promising explanation for the development of adult-like modal judgment is that, as children develop, they become more sensitive to the pragmatics of social discourse and thus the need to justify one's assertions. This sensitivity, when applied to modal judgment, would lead to greater internal reflection on the legitimacy of one's modal intuitions prior to the outward assertion of a modal judgment. Internalizing the need to provide a justification for one's assertions would thus change the nature of modal judgment from a gut reaction to an "internal argument" (Kuhn, 1991).

Research on the development of argumentive reasoning (Felton, 2004; Kuhn, Shaw, & Felton, 1997; Kuhn & Udell, 2003) has demonstrated the importance of social discourse in the development of argument skills, both social and internal. Whether these findings are relevant to modal development depends on whether mature modal reasoning does, in fact, constitute a form of internal argument. That said, there are at least two findings from the present study suggestive of the importance of social justification in older children's judgments. One finding, already detailed above, is that older children were significantly less likely than younger children to provide redundant justifications, presumably because they were aware of the inadequacy of such justifications. Even older children who lagged behind their peers in judging improbable events possible refrained from providing redundant justifications. Instead, they provided hypothetical justifications, which, though irrelevant to the exact question at hand, were at least relevant to the content of the events under consideration.

A second, less systematic finding is that older children (8–9-year olds) occasionally changed their mind about an event's modal status when attempting to provide a justification for their judgment, whereas younger children (4–5-year olds) almost never did so. In other words, older children were much more likely than younger children to change their modal judgment from "impossible" to "possible" after failing to identify a principled reason for why the event under consideration could not actually occur. This finding is admittedly anecdotal; changes in judgment were not systematically recorded, as we had no a priori reason to expect age differences in the frequency of this behavior. But, if we assume this observation is reliable, it suggests that older children see their justifications as being interconnected to their judgments in a way that younger children do not.

On the view that modal development entails the internalization (and prioritization) of the "justification process," two predictions follow. First, older children should take longer to make modal judgments than younger children, as they would require additional time to evaluate their modal intuitions against explicitly known causal principles. A stronger formulation of this prediction is that, regardless of age, the amount of time it takes children to make a modal judgment should be positively correlated with (a) their propensity to judge improbable events possible and (b) their propensity to provide factual justifications for those judgments.

Second, adults with less education than those in the present study should exhibit "less sophisticated" patterns of modal judgment and modal justification. As Kuhn (1991) has shown in her study of argumentive reasoning across the lifespan, adults without a college education are often incapable of providing solid justifications for their beliefs, mistaking explanation for evidence and restatements for rebuttals. Given that adults in the present study were all college undergraduates, it is quite possible that adults with other educational backgrounds would provide fewer factual justifications for their judgments of impossibility and may even judge fewer improbable events possible. Although adults in the present sample almost never judged improbable events impossible (Table 2), they did vary in their propensity to provide non-factual (i.e., hypothetical or redundant) justifications, with some providing as few as one and others providing as many as seven. Additional research with adults would not only help clarify the endpoint of modal development but would also help identify dispositional and/or situational factors underlying variation in "mature" modal judgment.

4. Conclusion

Modal claims – i.e., claims about possibility and necessity – are pervasive in everyday discourse and central to everyday reasoning. The present study used the ability to distinguish improbable events from impossible events as a basis for evaluating the nature and scope of children's modal development. Across three domains of knowledge, children exhibited similar patterns of modal judgment and modal justification, as well as meaningful correspondences between the two such that adult-like justifications were predictive of adult-like judgments independent of age. These findings militate against the idea that developmental differences in possibility judgment can be reduced to developmental differences in domain-specific knowledge and instead point toward the importance of studying procedural differences in modal reasoning, both across development and across education, in future research.

Appendix A.

See Tables A1–A3.

Table A1
The physical events, organized by concept and event type.

Concept	Event type	Event	Story
Conservation	Improbable	Making an umbrella out of glass	1
	Impossible	Making an umbrella out of air	2
Mass	Improbable	Catching a fly with chopsticks	1
	Impossible	Catching a shadow	2
Support	Improbable	Walking on a telephone wire	1
	Impossible	Walking on water	2
Contact	Improbable	Unlocking a door with a paperclip	2
	Impossible	Unlocking a door with one's mind	1
Entropy	Improbable	Gluing a broken eggshell back into an egg	2
	Impossible	Unscrambling a scrambled egg	1
Solidity	Improbable	Walking through a fire	2
	Impossible	Walking through a wall	1

Table A2
The psychological events, organized by concept and event type.

Concept	Event type	Event	Story
Attention	Improbable	Counting all the hairs on a dog's tail	1
	Impossible	Counting stars on an overcast night	2
Communication	Improbable	Reading someone's lips	1
	Impossible	Reading someone's thoughts	2
Knowledge	Improbable	Never forgetting anyone's name	1
	Impossible	Knowing someone's name just by sight	2
Audition	Improbable	Identifying a dog's breed by its bark	2
	Impossible	Hearing a sound before its made	1
Speech	Improbable	Speaking without moving one's lips	2
	Impossible	Speaking two languages simultaneously	1
Vision	Improbable	Reading a book that's upside down	2
	Impossible	Reading a book without opening its cover	1

Table A3
The biological events, organized by concept and event type.

Concept	Event type	Event	Story
Growth	Improbable	Growing a beard to one's toes	1
	Impossible	Growing from an adult back into an infant	2
Nourishment	Improbable	Not eating for 10 days	1
	Impossible	Not eating for 10 months	2
Vitality	Improbable	Living without a functional nose	1
	Impossible	Living without a functional heart	2
Lifespan	Improbable	Living for 120 years	2
	Impossible	Living for a thousand years	1
Reproduction	Improbable	Giving birth to 20 children in one's lifetime	2
	Impossible	Giving birth to a kangaroo	1
Sleep	Improbable	Staying awake for 5 days	2
	Impossible	Staying awake for 5 months	1

References

Baillargeon, R. (2004). Infants' physical world. *Current Directions in Psychological Science*, *13*, 89–94.
Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. Oxford: Oxford University Press.
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.
Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: The MIT Press.
Carey, S. (2009). *The origin of concepts.* New York: Oxford University 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.
Crowley, K., Shrager, J., & Siegler, R. S. (1997). Strategy discovery as a competitive negotiation between metacognitive and associative mechanisms. *Developmental Review*, *17*, 462–489.
Dennett, D. C. (1991). *Consciousness explained*. New York: Little, Brown and Company.
Doyle, A. C. (2001). *The sign of four*. New York: Penguin Classics. (Original work published in 1860.).
Felton, M. K. (2004). The development of discourse strategies in adolescent argumentation. *Cognitive Development*, *19*, 35–52.
Harris, P. L., & Koenig, M. A. (2006). Trust in testimony: How children learn about science and religion. *Child Development*, *77*, 505–524.
Hatano, G., & Inagaki, K. (1994). Young children's naïve theory of biology. *Cognition*, *50*, 171–188.
Hughes, G., & Cresswell, M. (1968). *An introduction to modal logic*. London: Methuen.
Johnson, C. N., & Harris, P. L. (1994). Magic: Special but not excluded. *British Journal of Developmental Psychology*, *12*, 35–51.
Kalish, C. (1998). Reasons and causes: Children's understanding of conformity to social rules and physical laws. *Child Development*, *69*(3), 706–720.
Kuhn, D. (1991). *The skills of argument*. New York: Cambridge University Press.
Kuhn, D., & Pearsall, S. (1998). Relations between metastrategic knowledge and strategic performance. *Cognitive Development*, *13*, 227–247.

Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentive reasoning. *Cognition and Instruction*, *15*, 287–315.
Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development*, *74*, 1245–1260.
Miller, S. A. (1986). Certainty and necessity in the understanding of Piagetian concepts. *Developmental Psychology*, *22*, 3–18.
Miller, S. A., Custer, W. L., & Nassau, G. (2000). Children's understanding of the necessity of logically necessary truths. *Cognitive Development*, *15*, 383–403.
Perner, J. (1993). *Understanding the representational mind*. Cambridge, MA: The MIT Press.
Piaget, J. (1987). *Possibility and necessity*. Minneapolis, MN: University of Minnesota Press.
Phelps, K. E., & Woolley, J. D. (1994). The form and function of young children's magical beliefs. *Developmental Psychology*, *30*, 385–394.
Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*, 346–362.
Rosengren, K. S., & Hickling, A. K. (1994). Seeing is believing: Children's explanations of commonplace, magical, and extraordinary transformations. *Child Development*, *65*, 1605–1626.
Rosengren, K., 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.
Shtulman, A., & Carey, S. (2007). Improbable or impossible? How children reason about the possibility of extraordinary claims. *Child Development*, *78*, 1015–1032.
Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.
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.
Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, *14*, 29–56.
Subbotsky, E. (1994). Early rationality and magical thinking in preschoolers: Space and time. *British Journal of Developmental Psychology*, *12*, 97–108.
Subbostky, 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.
Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology*, *43*, 337–375.
Woolley, J., & Cox, V. (2007). Development of beliefs about storybook reality. *Developmental Science*, *10*, 681–693.