This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

2 The Evolution of Rational Constructivism

Theory Theory (or the Child-as-Scientist Metaphor)

Developmental psychologists, observing how rapidly children are able to learn about the world around them, have sometimes claimed that children are little scientists (e.g., Gopnik et al., 1999). But what does it mean to say that a child is (or is like) a scientist? As we noted in chapter 1, the relevant dimension of similarity between children and scientists is that both are faced with the task of figuring out how the world works. Scientists go about this task in a systematic fashion, by making observations, designing experiments, collecting data, and drawing conclusions. The metaphorical view of child-as-scientist argues that children do the same thing. In their own way, children also make observations, design experiments, collect data, and draw conclusions.

A classic example of this can be found in children's play. When Dave's son was a toddler, he had a set of stacking cups in the kitchen, which he liked to stack up and knock down, over and over. Such toys—and such stacking activities—are common in many households, at least in our culture. The child-as-scientist metaphor argues that toddlers are not doing this randomly. Rather, their behavior is allowing them to learn about the properties of objects: They fall down when they are unsupported. At early ages, such stacking activities provide children with first-person experiences of building and with information about objects' mass and center of gravity that is necessary for appreciating physical relations involving support. This particular behavior even has a two-for-one bonus. At later ages, once concepts of support might be mastered, such stacking activities often become more of psychological intervention on the part of the child: How many times can I make

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

this loud banging noise before dad stops cooking and plays with me (or takes these toys away)?[1]

The idea of children as little scientists has been formalized into one of the major modern theories of cognitive development, known as the *theory theory*. Theory theory states that babies have a set of initial theories about the world. These theories are not exactly the same as the ones that scientists use, but they have a similar structure. Like full-blown scientific theories, babies' theories are abstract, coherent representations of causal structure (Gopnik & Meltzoff, 1997; Gopnik & Wellman, 1994). For example, babies might have a theory of the physical world that includes a concept of gravity: Objects fall down when they are not fully supported by a surface. Although these theories are simple, they allow babies and children to make predictions about the world and to interpret the information they observe. As children's language capacities develop, they become able to use what other people are saying as another form of evidence for their theories. Eventually, they become able to generate their own explanations.

One of the most important features of children's theories is that, like scientific theories, they can change over time. This is good, because the initial theories that children hold usually are not fully accurate. For example, babies initially behave as though any amount of support under an object will prevent it from falling, no matter how tiny of a sliver of the object's bottom is resting on the surface underneath (Baillargeon et al., 1992). There is something correct about this theory—objects that are entirely unsupported do fall—but this theory also needs to change to account for different types of scenarios. For example, some objects are asymmetrical and will fall unless they are balanced properly on the supporting surface. As children observe or hear about new evidence, they are able to incorporate this new information into their theories and adjust these theories to be a better fit to how the world actually is (e.g., Karmiloff-Smith & Inhelder, 1974).

To take another example, children's initial observations generally lead them to believe—incorrectly, but understandably—that the world is flat. As they grow up, they hear testimony from trusted adults that the world is actually round, and they may observe pictures of the Earth from space or encounter globes. Through a slow process of weighing this new observation against their initial belief, children come to learn that the world is actually round; they change their theory (e.g., Vosniadou, 1994). This kind of theory change

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

can be described as directly analogous to scientific reasoning: Children's theories change through the acquisition of new evidence and through a process of integration of old with new information, similar to how scientific theories change as new evidence is generated or discovered.

This example raises an important question to consider before we more fully discuss the development of scientific thinking: If children engage in this process of theory change, why do some erroneous, pseudoscientific beliefs persist? Consider again the process of learning that the Earth is round. Centuries of scientific data back up this assertion, yet some individuals continue to espouse the belief that the Earth is flat. These Flat Earthers have societies and conventions. Although there are many fascinating aspects to this phenomenon, we want to highlight that Flat Earthers do not seem to hold other beliefs that are incommensurate with reality. They believe that unsupported objects fall and that solid objects cannot pass through one another. Neither do they deny that science exists, nor do they say that scientific thinking is invalid. Indeed, perhaps surprisingly, they often try to use scientific methods to support their claims. To quote a report on one of their conferences, "While flat earthers seem to trust and support scientific methods, what they don't trust is scientists" (Dyer, 2018). That is, what seems to promote Flat Earth beliefs (and, presumably, many of the other beliefs that adults hold that are incommensurate with scientific evidence) is a misunderstanding about the relation between truth and power. Demonstrating that something (particularly something nonobvious) is true has the potential to be misinterpreted as a form of scientific elitism, which seems to be a big part of the Flat Earth movement. Some Flat Earthers potentially dislike the idea that there are truths about the world they did not (and potentially cannot) discover through own direct observations, possibly because they perceive that they lose some power in this process.

We suspect (although to our knowledge this has never been investigated) that individuals who hold Flat Earth beliefs (or who deny climate change or the safety of vaccines, or who hold other erroneous, pseudoscientific beliefs) developed their scientific thinking skills in the same way as individuals who do not hold these beliefs as adults. One of the reasons we had for writing this book was to articulate the idea that scientific thinking is not exclusive to a particular group of individuals (i.e., scientists), and that the power of scientific thinking is available to everyone. All individuals

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

have the potential to engage in the kinds of causal reasoning and scientific thinking that leads to our understanding that the Earth is round, that climate change is real, and that vaccines are safe.

A Problem with the Child-as-Scientist Metaphor

Theory theory has a lot to offer. Most notably, it can explain both the origin and the development of children's knowledge. But this theory, and the child-as-scientist metaphor in general, nevertheless suffers from a serious problem: It's vague.

Theory theory can be stated in two sentences: *Children have theories about how the world works. Their theories change as they gain more information.* But how, exactly, are children's theories represented? And how do these theories change as children observe information in the world?

In fairness, vagueness is a challenge faced by all theories of development. For example, to explain how children advance from one development stage to the next, Jean Piaget posited two mechanisms of development: *assimilation* and *accommodation*. As any developmental textbook will describe, children move from one state of equilibrium to another via these processes. Information that fits with a child's existing concepts is assimilated to that concept, while information that does not fit with the concept can be used to transform (accommodate) the concept to a new one. To take the example used above, as children are told that the Earth is round instead of flat, they come to accommodate their ideas about the shape of the Earth to the standard concept. To use the terms of theory theory, children's astronomical theories undergo change. But both of these descriptions, while sensible, are too vague to really be of much use in describing children's conceptual development. How do children's concepts change on the basis of new information that they receive? Changes in children's responses and behaviors can be observed, but what is actually happening in children's minds as they revise their theories? And how can we access these cognitive processes?

Speaking about the development of children's knowledge as a process of theory change can function as a productive representation of development only if there is some way to cash out the details of how this happens. Otherwise, it is only a metaphor—one that some people might be tempted to take too literally. For example, Dave was once challenged by a well-known

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

researcher who argued that the child-as-scientist metaphor could not be right because children do not actually wear lab coats and carry around clipboards. We're pretty sure that this is an apt observation: Children do not wear lab coats or carry clipboards that often.[2] But the fact that children do not do these things does not invalidate the child-as-scientist metaphor. The broader point of the metaphor is that children are actively testing their beliefs about the world and changing their knowledge based on their observations and their interactions with the world, much like the process of science.

Similarly, there is a strong and important tradition in studies of scientific thinking to observe scientists in their laboratories or other natural environments to see how they approach problem-solving (e.g., Dunbar, 1995, 2000; see also Latour & Woolgar, 1979). We think that this is also a valid way to study how children might learn science, reason scientifically, and, perhaps most importantly, become engaged with science. Knowing how scientists do their work also allows educators to emulate these processes with children. But analyzing the practices of adult scientists as a way of telling us how children should think scientifically misses a crucial point. There are certain practices that adult scientists engage in that children would never do. For example, as practicing scientists, we keep notebooks detailing our findings and experimental ideas. We use those notes to help design new experiments or to integrate findings together (and, sometimes, to remember what groceries to get). No 4-year-old would ever make such notes or engage in such abstract reasoning practices. But the absence of this behavior, just like the absence of lab coats and clipboards, does not indicate an inability to engage in any kind of scientific thinking; it just suggests the possibility that scientific thinking looks different in young children than in adult scientists. If we take the approach that scientific thinking is only and exactly what adult scientists do, then we will miss situations in which children (and adults) engage in scientific thinking in the course of their everyday activities.

Nevertheless, obviously, to make the child-as-scientist approach do real work for developmental science, we must go beyond the metaphor. We must specify what theories are, how they are represented, and the process by which one representation of a theory changes to another. Luckily, a response to this challenge comes from the definition of theory theory laid out above: Theories are abstract, coherent representations of *causal structure*. To understand how theories are represented and how theories change, we have to understand

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

how causal structures are represented, how they change, and how they can allow for abstraction and coherence. This is something that researchers in psychology, philosophy, and computer science have been working on for a long time.

Causal Reasoning as Associative Learning

Following Piaget's (1929) description of young children as "precausal," many early theories of children's causal reasoning posited that they made inferences based on associations, hence took their cues from the long literature on associative learning in comparative psychology.[3] In this view, over the course of many observations, human beings (and nonhuman animals) build up stronger and stronger links between events that tend to co-occur, which are generally causes and effects. Nonhuman animals can learn various kinds of associative relations in order to make predictions about their environments. For example, nonhuman animals can learn over the course of many trials that a particular type of action on their part will tend to be accompanied by a reward. The more these animals perform this action, the stronger this link between action and reward becomes. This can be interpreted as understanding the relation between a cause and an effect—action causes reward. Based on these sorts of experiments, classic psychological theories of associative learning formalized how animals relate conditioned and unconditioned stimuli to each other in such a way as to make better predictions about their environments (e.g., Mackintosh, 1974; Pearce & Hall, 1980; Rescorla & Wagner, 1972).

These frameworks provide one way of thinking about causal reasoning in human beings, including in children, because human beings can use associative mechanisms similar to the ones that are available to nonhuman animals (see e.g., Cramer et al., 2002; Dickinson, 2001; Dickinson & Shanks, 1995). This process manifests in the ability to pick out associations among events in the world. Indeed, infants and young children appear to possess sophisticated capacities for noticing statistical regularities in the environment (e.g., Fiser & Aslin, 2002; Goldstein et al., 2009; Haith et al., 1993; Kirkham et al., 2002; Saffran et al., 1996), making such a hypothesis developmentally viable.

The ability to learn individual associations is a powerful mechanism for interpreting the world. However, this mechanism is limited. One limitation is the computational complexity that this system would require. For

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

example, most of the research cited above focuses on how children learn that events unfold over time (e.g., in a given language, syllable A is more likely to be followed by syllable B than by syllable C, and is never followed by syllable D) or how events are paired in space (e.g., members of category X have both features A and B, while nonmembers of category X lack both features A and B). Massive demands on memory and reasoning are needed in order to learn all the associations necessary to make such inferences. Given the sheer volume of data necessary to make these inferences, such reasoning requires a good amount of exposure to data. Moreover, many attentional resources are involved. The world has many statistical regularities, and many of those associations fail to indicate causal relations. To avoid being swamped by thousands of spurious associations, learning from association must also involve attentional or arousal systems to know what events to process.

Finally, for associative learning to be a good representation of causal structure, one also needs to be able to stack associations on top of one another to create higher-level correlations. Knowing that event A correlates with event B, and that event A correlates with event C, does not indicate that event B will correlate with event C without the knowledge that A is also present. For example, a simple associative learning system will associate objects that have hands together with objects that have legs (i.e., bodies). Infants can register these kinds of correlations among object features (Younger & Cohen, 1983). Infants also understand that hands (but not other objects like sticks) produce goal-directed actions (Sommerville et al., 2005; Woodward, 1998). Registering the kind of higher-order correlations we are describing means that infants will additionally be able to infer that objects with legs will engage in goal-directed actions, even if the infant never sees the object's hands.

Although it may seem implausible that children or babies could engage in such complicated inferences, several studies have demonstrated that young children can indeed detect such second-order correlations among static features of objects (Cuevas et al., 2006; Yermolayeva & Rakison, 2016) as well as among dynamic features of objects (Rakison & Benton, 2019). Some of our work demonstrated that children can use second-order correlations to make causal inferences about nonobvious object properties (Benton, Rakison & Sobel, 2021). We introduced 2- and 3-year-olds to two objects that differed in shape, color, and size (objects A and B). Object A had a unique feature (X),

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

while object B had a different unique feature (Y). The two objects just sat on a table in front of the children for about 10 seconds, so that children could see them but not explore them. Then the objects were put away.

Children were then introduced to a novel machine (a blicket detector, which we describe below in more detail, but for now, we can just describe it as a small box) and two new objects. One object was identical to object A, but did not have the unique X feature. The other object was identical to object B, but did not have the unique Y feature. These two objects were placed on the machine individually; one of them activated it (caused it to light up and play music when it came into contact with it) and the other did not.

Finally, children were introduced to a third pair of objects. These objects looked completely different from A and B, except that one of them had feature X and the other had feature Y from the original pair of objects that children saw. Children were asked to make the machine go with these new objects. Children tended to choose the object that had the second-order correlation with the machine's activation. That is, if the object that looked like object A had activated the machine, then children chose the object in the test pair with feature X, whereas if the object that looked like object B had activated the machine, they chose the object with feature Y. These data suggest that even 2-year-olds can register second-order causal relations. In turn, these results support the idea that children's (and adults') causal reasoning could be the result of layers of associative reasoning, which in turn keep track of different kinds of correlations.

Although children may have access to the computational abilities to represent a wide variety of complex associative relations, we believe that an associationist framework is not the best way to model children's causal reasoning, because this framework makes a series of predictions that are not borne out by other data. For example, in cases like the study described in the last paragraph, higher-order associative reasoning may lead children to simply put both objects on the machine. If they do not know which one makes it go, or if they have even the slightest uncertainty, putting both objects on the machine is a reasonable thing to do, because this would reflect the highest possible correlational structure. Critically, children never did this in our study, and they almost never do this in other studies where they could solve a problem by using this kind of associative reasoning (see e.g., Gopnik et al., 2001; Sobel, 2020, for similar results). We need to look

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

elsewhere for a full answer to the question of how children construct and represent theories.

Constraining Causal Inferences

Associative reasoning is a good first step, but learning causal structure involves more than just understanding that events are associated. It involves appreciating the hows and the whys of that association. The hows involve appreciating *mechanism*, that is, the ways that events relate to one another; this allows for *interventions,* that is, the ability to act on events in the world to produce outcomes. The whys involve appreciating the reasons behind those mechanisms—understanding when mechanisms generalize to novel events.

Empirical work demonstrates that children do appreciate more than just associations between events. For example, children can integrate various pieces of causal knowledge they possess with their associative reasoning and statistical learning capacities (e.g., Denison & Xu, 2010; Madole & Cohen, 1995). In one study of this ability, Madole and Cohen (1995) showed 14- and 18-month-olds a set of trucks. The trucks had either small black wheels or large yellow ones, and the top parts of the truck looked like either a small person or a green tree. The wheels could either roll or were fixed, and the top part could make a whistling sound. These researchers initially showed that infants could learn the relation between an object's parts and those parts' functions. Specifically, they habituated infants to trucks on which the black wheels rolled, but the yellow ones did not, regardless of what the top part of the truck was. At the same time, they habituated infants to trucks on which the top part that looked like a tree whistled, but the part that looked like a person did not. Infants in both age groups could learn these relations, understanding that the perceptual features of the parts (black or yellow wheels; top that looked like a tree or a person) predicted the function of those parts (rolling vs. stationary; whistling vs. not).The main idea of their experiment, however, was not to show that infants could learn these correlations, but to show that infants did not learn all kinds of correlations. There are a lot of correlations in the world, and most of them have little to do with the actual causal structure of the world. For instance, there is a strong inverse correlation between the number of pirates on Earth and the Earth's temperature, but it's hard to find a causal mechanism for

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

how fewer pirates directly causes an increase in the Earth's temperature. We naturally reject this correlation as indicating causal structure because there is no mechanism by which it could work. Nor do we think that popularizing "Talk Like a Pirate Day" (a real thing, apparently—it's September 19) will effectively combat climate change.

To show that infants understand this as well, Madole and Cohen (1995) presented infants with a different kind of correlation. Now, the perceptual appearance of one part predicted the function of the *other* part. So if a truck had black wheels, the top part made a whistling sound, regardless of whether it was shaped like a person or a tree. Similarly, if the top part was a tree, then the truck's wheels rolled, regardless of whether they were black or yellow. The nature of the correlations in this condition were the same as in the other one. The amount of data that the babies were given to learn these correlations was the same. But these correlations just do not make sense. Like in the case of pirates and climate change, there does not tend to be real-world situations in which the shape of one part of an object correlates with the function of another part of the object, which can differ in shape. Eighteen-month-olds did not learn this correlation. That is, they did not dishabituate to the trucks that violated the correlation. However, 14-month-olds did. We interpret these findings to show that the older babies did not pay attention to this correlation in the first place, or that they did not think it was worth learning. And why should they? By this age, they have other aspects of causal knowledge that allowed them to register this correlation as unimportant and to reject it, or to not attend to it at all in the first place.

This study shows that, by 18 months, children can use aspects of their real-world knowledge to constrain which correlations they will learn. This implies that human causal learning is not purely driven by correlations. Human beings have access to other information that guides their learning of particular types of correlations but not others.

We drew the same conclusion using a different approach (Tummeltshammer, Wu, Sobel & Kirkham, 2014). In this experiment, we were interested in understanding how well infants could track the reliability of others as a source of knowledge. We showed 8-month-olds videos of a person in the middle of a screen. The person on the screen got the baby's attention, and then turned to a location in space (one of the four corners of the screen). An interesting cartoon with a weird sound effect then appeared in one of the four corners of the screen. The trick of the experiment is that one set

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

of videos showed a person who was always accurate; the person looked to different locations four times, and, all four times, that's where the cartoon appeared. The other set of videos showed a different person who was accurate only 25% of the time, only looking at the corner where the video appeared one out of four times. We wanted to determine whether the infants would learn to follow these two people's gaze differently from this relatively brief exposure. And infants did. When the accurate informant looked to a novel location on the screen, 8-month-olds followed their gaze. But they did not do so when that same location was cued by the inaccurate speaker.

These two conditions of the study show that even 8-month-olds can track the accuracy with which other people generate information (a point that we raised in chapter 1 and that we return to in chapter 3). But the more important point for the present purposes is the control condition. In that condition, the faces were replaced by two different blobs that morphed into arrow-like shapes to cue the location of the cartoon. Eight-month-olds did not learn in this condition; they did not look in the direction that the blob was "pointing" at test. Like the Madole and Cohen (1995) study described above, this suggests that babies are not just always associating information together to make inferences. Instead, they are using higher-level information—that people can be sources of knowledge, but objects are less likely to be—to figure out when to track the accuracy of someone or something.

What is important about all of these studies is that infants can integrate their existing knowledge about how the world works into their ability to register and reason about statistical regularities they observe. Critically, there might not be a specific age when they can do this generally; the two studies we reviewed examined infants of different ages. This is because the specific pieces of causal knowledge that are necessary to constrain particular causal inferences develop at different times. The important point is that the causal knowledge children possess at any given time constrains how they process new information and has cascading effects on the development of subsequent knowledge. In turn, this strongly implies that children's reasoning abilities are supported by something more than mere association.

Bridging Statistical Learning to Causal Models

The fact that some associations are stronger or more salient than others is not a novel idea. This is even the case for nonhuman animals, and it is consistent

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

with classic models of associative reasoning.[4] For example, rodents learn taste aversion in a single trial, rather than needing several trials to build up the negative association (e.g., Garcia et al., 1955). This suggests that there are other forces (perhaps evolutionary ones) that prime or bias rodents' learning of the relation between taste and food avoidance, or that make this kind of relation particularly salient.

Based on results like this one, many models that started as purely associative began incorporating a way to calculate the causal strength of known causal relations between events or properties, allowing them to integrate top-down knowledge with associative learning principles (e.g., Krushke & Blair, 2000; McClelland & Thompson, 2007; Rogers & McClelland, 2004; Van Hamme & Wasserman, 1994; Wasserman & Berglan, 1998). A related approach relies on estimating causal parameters based on the frequency with which events co-occur, such as in the ΔP model (Allan, 1980; Jenkins & Ward, 1965; Shanks, 1995) and the Power PC model (Cheng, 1997; Novick & Cheng, 2004).

Although many published papers validate each of these models (and many more papers challenge them), a problem for our purposes is that these models were designed with adult reasoners in mind, and the evidence that has been generated in support of these models comes from adult participants. Empirically, children's causal reasoning does not match adults' in many cases, and there is evidence that children's inferences are not well-explained by any of these models (e.g., Griffiths et al., 2011; Lucas et al., 2014; Sobel et al., 2004).

There are also theoretical concerns with using these models in development. For example, in some of the models mentioned above, learners are assumed to have separate processes for identifying events in the world that could be causes and effects and for deciding how those events fit together in a causal structure. That is, they first identify the causal structure, and then separately determine the strength of individual causal relations. It seems reasonable to keep these processes separate; describing which events in the world might be potential causes and effects can be done independently of determining how these causes and effects might be related to one another. If I ask you to determine the relation between flipping a switch and a light turning on, for example, it is relatively straightforward to posit that the switch causes the light to turn on, and not the opposite. This process could be independent from determining the strength of the relation between the

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

switch and the light.[5] The trouble is that it's not clear whether these processes actually are separate for young children.

Given these complexities, researchers in cognitive science started to look for alternative computational frameworks to describe how children represent causal knowledge and engage in causal inference. One successful approach has relied on a computational framework called *causal graphical models* (Glymour, 2001; Pearl, 1988, 2000; Spirites et al., 1993; Woodward, 2003). These models came out of a research program at the intersection of philosophy and computer science and are now a popular tool in cognitive and developmental psychology. In particular, Gopnik et al. (2004; Gopnik & Wellman, 2012) argued that causal graphical models could solve the vagueness of the theory theory by specifying how causes and effects could be represented and at what level of abstraction. Moreover, well-established algorithms have been developed for these models to describe how relations between causes and effects could be learned, particularly from observed data. Further, these models provide a description of how inference and counterfactual reasoning occur, supporting the coherence of these representations.

Overall, then, causal graphical models seem like they could be an especially good fit for describing children's behavior. But are young children's inferences and representations of causal structure really consistent with this framework? In the next two sections, we describe a body of work that shows that they are, establishing the causal graphical model framework as a viable and productive way of understanding causal reasoning, hence theory representation and theory change, in development.

What Are Causal Graphical Models?

The first step is to explain what a causal graphical model is.[6] To explain that, we have to explain what a plain graphical model is. Briefly, a graphical model is a formal (mathematical) way to represent a joint probability distribution.

So what's a joint probability distribution? Imagine a list of all possible combinations of the events under consideration and the probability that each combination occurs—that's a joint probability distribution. This can be represented by a graph. In this formalism, the nodes in the graph are used to represent events, objects, or their properties. The vertices in the model are used to represent particular types of dependencies between such objects or

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

events, such as their relational structure. Figure 2.1 shows a simple example: Variable X is related to variable Y, and both X and Y are independent of Z.

Importantly, information about conditional probabilities (that is, which events are likely to occur given that other events have occurred) can be extracted from these structures. For the simple example in figure 2.1, if you know that X has occurred, you will think that it's likely that Y has occurred too (or, more precisely, it affects the probability that Y occurs). But if you know that Z has occurred, you do not have any information about either X or Y.

Given that graphical models can be used to represent events and the probabilities that they occurred, it starts to become clearer how these kinds of models can be used to talk about causal reasoning. However, before we can formally interpret these models as psychological representations of causal knowledge, we must make two further assumptions about the underlying structure of the connections between events (nodes) and the relations among them (vertices): the *faithfulness assumption* and the *Markov assumption*. Importantly, these are assumptions, not empirical claims. While there is some empirical evidence that suggests that these assumptions are being used, they are part of the necessary background for reasoning with graphical models.

Faithfulness

The faithfulness assumption is the assumption that the data we observe are indicative of the actual, ontological structure of the world. The broad assumption here is that we see the world the way it really is.[7] For our purposes, what faithfulness means is that there is never a situation in which there would be a direct contradiction between the causal data that we observe and the causal structure that actually exists in the world.

For example, suppose that we observe that two events are independent, like X and Z in the example described above (figure 2.1, or the top panel of

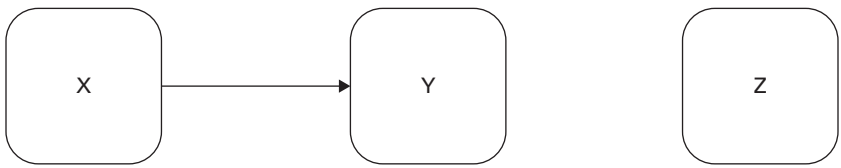


Figure 2.1
A representation of a (not very interesting) graphical model.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

figure 2.2). The faithfulness assumption says that there is never going to be another event (like the unobservable naughty monkey, shown in the bottom panel of figure 2.2) that inhibits the occurrence of Z with exactly the same causal efficacy as event X causes the presence of Z whenever X occurs, thus masking the presence of a causal relation between X and Z (represented by the dashed line in the bottom panel of figure 2.2). Put another way, in order to get causal reasoning off the ground, we have to assume that there are no unobservable naughty monkeys playing with our perceptions and inferences about the world.

Of course, no specific empirical evidence fully supports or invalidates the faithfulness assumption. How can we prove that there are no unobservable naughty monkeys? Wouldn't the unobservable naughty monkeys just

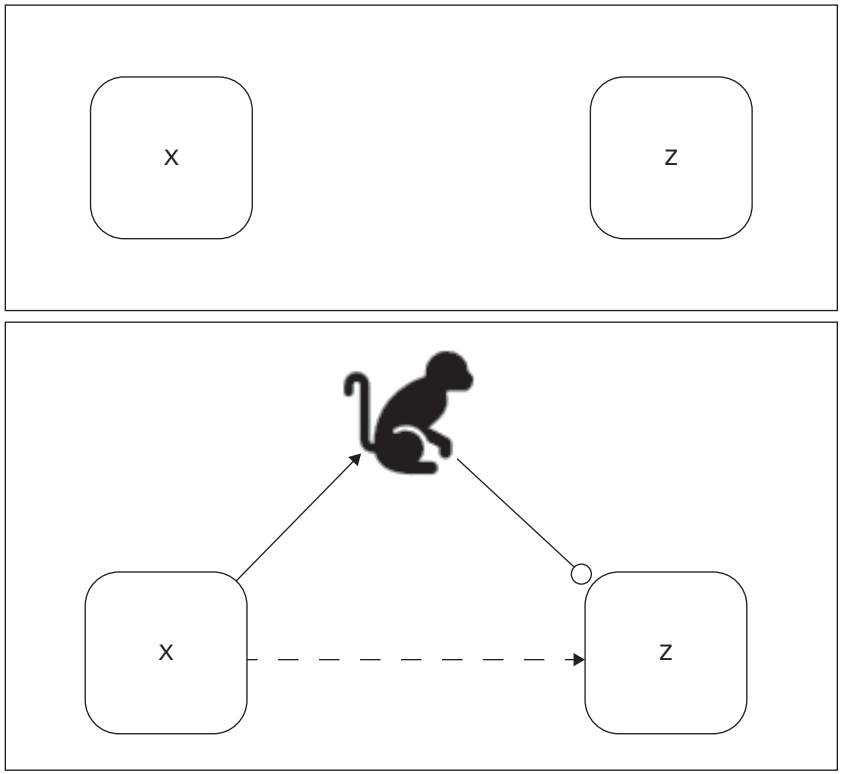


Figure 2.2
Illustrating the faithfulness assumption.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

manipulate any experiment we do to result in evidence that does not suggest their presence? This is correct; we can't prove that there are not unobservable naughty monkeys (as Descartes noted in his *Meditations*). Faithfulness is an assumption. But once we accept it, we can make enormous progress in conceptualizing how the world works.[8]

The Markov Assumption

The Markov assumption states that the value of an event (i.e., a node in the graph) is independent of all other events, with the exception of its children (i.e., its direct effects) and its parents (i.e., its direct causes). To take a concrete example, let's look at an incredibly simple model of the weather, shown in figure 2.3.

For the graph shown in figure 2.3, all else being equal, the Markov assumption states that the event of raining yesterday and the event of raining tomorrow are independent; there is no direct relation between them. Whether it rained yesterday and whether it rains today are related, as are whether it rains today and whether it will rain tomorrow. That is, the event of raining yesterday and the event of raining tomorrow are dependent on each other, because they are related through the event of raining today. However, the only influence that raining yesterday has on raining tomorrow depends on whether it rains today; these two events are otherwise independent. If /rɔθ/ ("roTh," the backwards doG of Thunder[9]) comes along and makes it rain today, it no longer matters whether it rained yesterday when predicting whether it will rain tomorrow. Knowing that it rained today is all you need.

Do children reason about the relations among events using the Markov assumption? And, if they do, how could we test this? To answer this question, our test would need to include a way to examine whether children can recognize that some events are independent of each other. But, as reviewed

Figure 2.3
A simple model of the weather.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

in chapter 1, even young children have a good deal of knowledge of how events are causally related to each other and could bring that prior knowledge to bear when responding in our studies. This means that any test of the Markov assumption in children would need to avoid using any causal relations about which children could have prior knowledge.

That's what blicket detectors are for.

Blicket Detectors

As a graduate student, Dave was part of a team[10] that developed a novel paradigm for testing children's causal knowledge: the *blicket detector*. An example is shown in figure 2.4. The blicket detector is a box that can light up or play music when certain kinds of objects are placed on top of it, making it look as though these objects have activated the machine. The machine actually works via an enabler switch, which is controlled by an experimenter and kept hidden

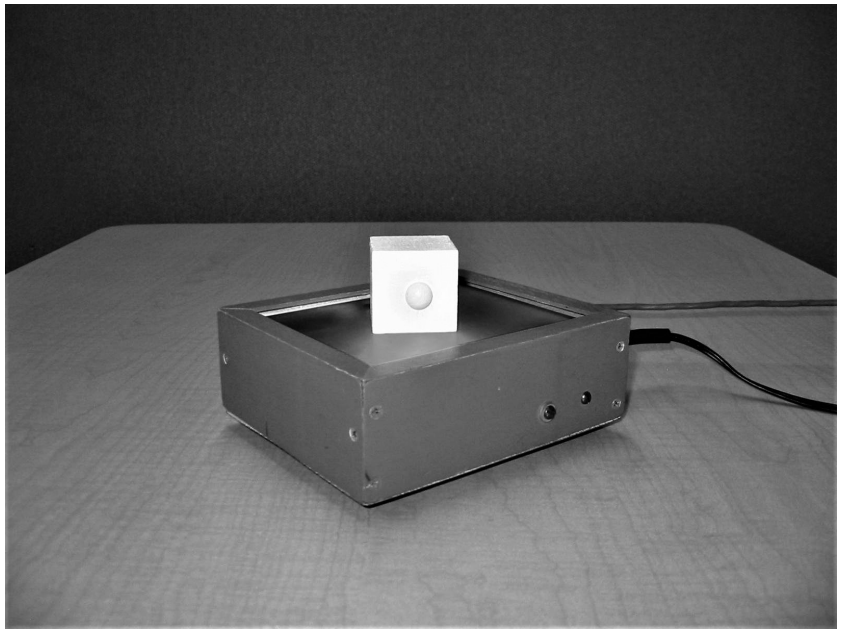


Figure 2.4
The original blicket detector with an object on it. This box lit up red and played a MIDI recording of *Fur Elise* when it was active.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

from the participant. When the switch is in the "on" position, anything that is placed on the box activates it; when the switch is in the "off" position, nothing activates the machine. The top of the box has a pressure-sensitive plate, which is connected to this enabling switch and which makes it appear as though the objects that are placed on top of the machine are making it turn on.[11] Later versions of the detector added more features, like turning different colors or being activated by remote control, which could allow researchers to present different examples of causality at a distance (e.g., Kushnir & Gopnik, 2007; Sobel & Buchanan, 2009). The crucial aspect of these machines, however, always stays the same: Some objects placed on them make them activate while others do not, controlled by the experimenter. This machine, though simple, is a powerful tool for studying children's causal reasoning abilities.

The first studies to use the blicket detector did not focus on children's causal reasoning, but rather on the extent to which causal features of objects were important for categorical inferences. We (Gopnik & Sobel, 2000) presented 2- to 4-year-olds with a set of nondescript objects (wooden blocks of different shapes and colors) and the machine. We showed children that some objects made the machine activate, while others did not. For example, we placed four objects on the table with the machine and then placed the objects on the machine one at a time. Two activated the machine and two did not. We then labeled one of the objects that made the machine go a "blicket" and asked children to show us the other "blicket." Children—certainly by the age of 4—picked the second object that had activated the machine (see also Nazzi & Gopnik, 2000).

Most of these studies also ran a condition in which the experimenter would hold each object over the machine in one hand and use their other hand to press the panel on the machine so that it activated. This way, the object was associated with the machine's activation, but there was clearly another cause for the machine's activation (i.e., the experimenter's hand). In this case, children responded at chance when asked to find the other blicket. That is, children were differentially sensitive to cases where an object seemed to cause the machine's activation and cases where an object was merely associated with the machine's activation.

The crucial point of connection between the blicket detector and the causal graphical model framework introduced above is that the detector presents a novel causal system. Although children may have some prior knowledge

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

about machines and some basic knowledge about causal relations (like the fact that causes temporally precede their effects), children have never seen this machine before and hence do not have any expectations about how it works. That is part of what makes the blicket detector such a powerful tool for testing children's causal reasoning abilities: It allows researchers to flexibly present new causal systems without having to worry about whether different children are approaching these systems with different levels of domain-specific causal knowledge. In this way, it allows for a good test of whether children are using the Markov assumption, because some events can be presented as dependent on each other, and some can be presented as independent. Children can then be tested to determine how they interpret these new events.

This is what we did in Gopnik, Sobel, Schulz, and Glymour (2001). We started by noting the importance of the control condition in Gopnik and Sobel (2000), in which the experimenter holds the object over the machine and presses the panel of the machine down with his hand to activate the machine. In this case, even though an object is associated with the machine's activation, there is another candidate cause (the experimenter's hand) that is a better explanation for the machine's activation. The assumption that children seemed to make is that the experimenter's hand explains the activation of the machine, and the presence of the object over the machine is independent of the machine's activation (or of the experimenter's decision to activate the machine). That is, children appeared to reason according to the Markov assumption.

To test this more directly, this study contrasted children's inferences about objects that activate the machine by themselves with objects that only activate the machine when another object is present. We told 3- and 4-year-olds that certain objects, called "blickets," would activate the detector. Then, children observed one of two types of trials. On the *one-cause* trials (figure 2.5), children were shown two objects (A and B). Each object was placed by itself onto the detector. One object (A) activated the detector by itself. The other object (B) did not. Children then saw objects A and B placed on the detector together (twice), and the detector activated (both times). This means that object B was associated with the activation of the detector two out of the three times it was placed on it. If children were keeping track of only associations between causes and effects, they might reasonably conclude that B was a blicket. But if children were reasoning about the events that they observed

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

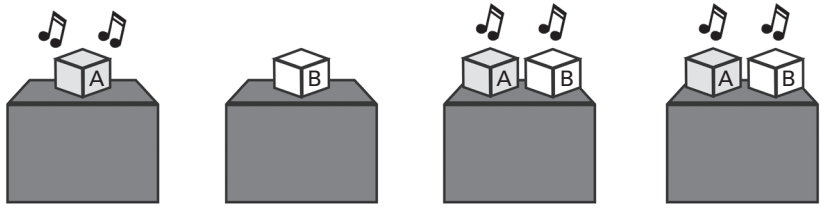


Figure 2.5
Illustration of the one-cause trial from Gopnik et al. (2001).

using the Markov assumption, they should realize that B only activated the detector dependent on the presence of object A. They should not use the positive association between object B and the machine's activation to infer that B alone could make the machine activate.

This is what children did. To test children's understanding of these events, at the end of the experiment, they were asked whether each object was a blicket. Children in the experiment generally labeled only object A as a blicket. That is, they recognized that B's association with the detector's activation (the effect) was dependent on the presence of object A. To put it the other way around, children understood that object B lacked causal efficacy when object A was not in the equation.

On the analogous *two-cause* trials (figure 2.6), everything was the same except that B activated the machine independently two of the three times it was placed on the machine. Even though B was associated with the effect (machine activating) with the same frequency as in the one-cause trials, children tended to label B as a blicket in this case. In these trials, object B activated the machine independently from object A, and children used this fact to conclude that it was also a blicket.

Various investigations since the publication of this paper have extended these findings to infants (e.g., Sobel & Kirkham, 2006), to other domains of knowledge (Schulz & Gopnik, 2004), and to other kinds of inferences supported by the Markov assumption (i.e., those involving causal chains or common causes, Sobel & Sommerville, 2009). These data all suggest that children adhere to the Markov assumption in their causal inferences. In turn, we have good reason to think that the causal graphical model framework provides a productive way to describe and explain young children's causal reasoning.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

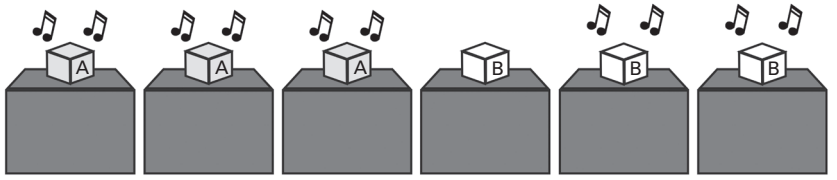


Figure 2.6
Illustration of the two-cause trial from Gopnik et al. (2001).

Back to Associative Reasoning

The Gopnik et al. (2001) findings suggest that children are reasoning according to the Markov assumption. However, what that work did not consider is *how* they might be doing that. One worrying possibility is that other reasoning mechanisms could simulate the Markov assumption. That is, it could appear as though children were reasoning in a manner consistent with the Markov assumption but were not representing causal knowledge via the causal graphical model framework; rather, a simpler mechanism could account for their inferences. A candidate for that simpler mechanism is associative reasoning. Indeed, the findings on second-order conditioning and the findings on statistical learning described earlier in this chapter are both consistent with the idea that children's causal reasoning is associations all the way down. Although some of these associative models of causal reasoning include top-down information, at base, they still calculate the associations between events to make causal judgments. This has led numerous researchers across different domains of psychology to argue for the importance of associative mechanisms in causal reasoning (e.g., Blaisdell, 2008; Cramer et al., 2002; Dickenson & Shanks, 1995; Heyes, 2012; see also Hanus, 2016, which provides a good review of the comparative literature but does not consider developmental work).

To illustrate how this might work, consider again the procedure from Gopnik et al. (2001). In the one-cause condition, object A activates the machine by itself (we'll represent this as A+), then B fails to activate the machine by itself (B–), then they both activate the machine together twice (AB+, AB+). In the two-cause condition, the associations are the same—it's just that the objects are never presented together (A+, A+, A+, B–, B+, B+). On a simple version of associative reasoning, in which you just count the

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

number of times the stimulus (the object A or B) is paired with the effect (+), these two conditions are equivalent. But children treated these two conditions differently: In the one-cause trials, B was not considered to be a blicket, but in the two-cause trials, it was. Because of this, we can conclude that children are not reasoning according to simple association; something more complicated is going on.

If you are familiar with associative reasoning, you should recognize that the pattern of data presented to children in the one-cause condition is similar to the pattern of data presented in nonhuman animals in studies on *blocking* (Kamin, 1969). In basic Pavlovian conditioning experiments, if a conditioned stimulus (A) is paired with an outcome, and then that stimulus A is presented in compound with a novel stimulus (AB) and also paired with that outcome, animals will learn the association between A and the outcome more than between B and the outcome. These data suggest that a simple model of associative reasoning cannot account for the animal's behavior. But more complex models of reinforcement learning could.

One such model is the Rescorla-Wagner (1972) model, which was an attempt to describe how associative learning worked; one of its earliest successes was that it provided an explanation of the phenomenon of blocking. We are going to step through the calculations of this model, because the associative challenge is significant not only here, but also to arguments we want to make in the chapter 3. So it's important to appreciate how it (and more contemporary models like it) describe the process of making inferences from data.[12]

The Rescorla-Wagner model is a way of updating the associative relation between conditioned and unconditioned stimuli, though for our purposes, we will just say a stimulus and an outcome. It is defined with the following formula:

$$\Delta V_{n+1} = K(\lambda - \Sigma V_n)$$

To unpack this, start with the first term, ΔV_{n+1}. This represents the change in the associative strength of a relation between a stimulus and an outcome on the next trial (i.e., exposure to the pairing). Here, V is a variable that represents associative strength, and the Greek letter Δ is a symbol that means "change in." So the change in the associative strength of a stimulus is a function of three things: (1) the salience of the stimulus and the outcome (represented here by K),[13] (2) the relation between the associative strength

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

of all the stimuli that are present on trial n (ΣV_n), and (3) how much can be learned on a given trial (λ). This last variable (λ) represents how much of an associative relation that stimulus can have with this outcome. It's the largest amount of associative strength that can be learned about this outcome. Finally, ΣV_n represents the sum (Σ) of the associative strength (V) of all stimuli that are present on this trial (trial number n). Taken together, the formula is a way to represent what an organism learns from a particular event, given a stimulus, an outcome, and the organism's prior knowledge about the strength of the association between and the stimulus and the outcome.

There is one more important point here. The units used in doing calculations with this model are arbitrary, because there is not a unit of measurement for the strength of an associative relation. Rather, this formula provides a mathematical way of contrasting differences among different learning scenarios. As long as you keep your numbers constant within a modeling framework, what matters is the difference in values, not the values themselves.

Let's apply this formula to the Gopnik et al. (2001) procedure, starting with the two-cause condition. To do so, we will make up some values. We have no reason to assume that object A's relation with the machine is more salient than object B's relation, so we can assign $K_A = K_B$. For this example, let's assign both of these variables the value of 0.2. Let's also posit that the most associative strength you can learn when the effect is present is 100 units. So $\lambda = 100$. Now, when you run the numbers:

Initially: $V_A = V_B = 0$
Trial 1 (A+): $\Delta V_A = .2(100 - 0) = 20$; $V_A = 20$; $V_B = 0$
Trial 2 (A+): $\Delta V_A = .2(100 - 20) = 16$; $V_A = 36$; $V_B = 0$
Trial 3 (A+): $\Delta V_A = .2(100 - 36) = 12.8$; $V_A = 48.8$; $V_B = 0$
Trial 4 (B–): $\Delta V_B = .2(0 - 0) = 0$; $V_A = 48.8$; $V_B = 0$
Trial 5 (B+): $\Delta V_B = .2(100 - 0) = 20$; $V_A = 48.8$; $V_B = 20$
Trial 6 (B+): $\Delta V_B = .2(100 - 20) = 16$; $V_A = 48.8$; $V_B = 36$

Let's unpack this. On Trial 1, the change in associative strength of A is the K value (0.2) multiplied by the λ value (100, because the effect was present) minus the strength of all the stimuli that were present on this trial. The only stimulus that was present was A, and its associative strength was 0. Hence $\Delta V_A = 20$. Trials 2 and 3 follow that same logic—the only thing that changes is the associative strength of the stimuli present.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

At this point, we want to point out something really cool about this model, which hopefully reflects an intuition you have about the world: As you are exposed to the same association more and more, you learn from it less and less each time. On Trial 1, the change in the associative strength of A was 20, but on Trial 2, it was 16. Even if you pair a stimulus with an effect thousands of times individually, you can't ever achieve an associative strength beyond the λ value. This was an important point of the model; it was focused on the idea that the slope of a learning curve changed as the animal was given more trials. An analogy in human learning is automaticity: Some things are learned so well that further exposure does little to strengthen the relation. In the human case, that can make these associations hard to unlearn (e.g., Shiffrin & Schneider, 1977).

On Trial 4, you should notice a pretty big change to the calculations. On this trial, the λ value is 0. That is because the effect *did not occur* on this trial. In this case, B does not accrue any associative strength, because there is nothing to associate with its presence. This feature of the model illustrates another important part of the learning process: Associations can also be unlearned, a process known as *extinction*. But the more strength an association accrues, the harder it is to unlearn.

So at the end of this process, the associative strength between object A and the machine's activation is 48.8 and the associative strength between object B and the machine's activation is 36. Children are then asked, "Is this (A/B) a blicket?" Presumably, to answer that question, the associative strength has to exceed a certain value. Because the numbers are arbitrary, we can choose that value, as long as it stays the same throughout the demonstration. So let's pick a value of 30. In this scenario, then, A and B are both blickets because their associative strength exceeds that threshold. This matches nicely with the empirical data from Gopnik et al. (2001).

But what about the one-cause trial? Again, let's run the numbers:

Initially: $V_A = V_B = 0$
Trial 1 (A+): $\Delta V_A = .2(100 - 0) = 20$; $V_A = 20$; $V_B = 0$
Trial 2 (B–): $\Delta V_B = .2(0 - 0) = 0$; $V_A = 20$; $V_B = 0$
Trial 3 (AB+), $\Delta V_A = \Delta V_B = .2(100 - 20) = 16$; $V_A = 36$; $V_B = 16$
Trial 4 (AB+), $\Delta V_A = \Delta V_B = .2(100 - 52) = 9.6$; $V_A = 45.6$; $V_B = 25.6$

Notice a few things here: First, Trials 1 and 2 result in the same effects as the first times A and B were put on the machine in the two-cause case. On

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

Trial 3, the change in associative strength for A and B is the same, but it is lessened by the fact that A already has some associative strength. It's lessened even more on Trial 4 because A and B are both present, and both have some associative strength. If we use the same arbitrary threshold for "blicketness" as we used above (an object is a blicket if its V is greater than 30), then A is a blicket in this example, and B is not. Again, this result matches the data from Gopnik et al. (2001). This conclusion lends credence to the idea that children's performance on that task (and on similar tasks) could be the result of this kind of associative learning, hence might not be due to children's abilities to reason according to the tenets of the causal graphical model framework or to use the Markov assumption.

An important objection to this example is the arbitrariness of the values that we assigned, particularly our threshold for deciding that something is a blicket. This is true; these numbers are arbitrary. But instead of thinking of it as a bug, think of it as a feature. For one thing, these numbers are constants, so the results will replicate regardless of their exact value. But more importantly, we can choose these numbers to be as generous to this associative modeling framework as possible, so that it has the best chance of explaining children's performance (although we are going to show in the next section that it does not work all the time). The point right now is that, based on the modeling, the associative strength of object B in the two-cause condition exceeds that of the strength of object B in the one-cause condition. As long as that is the case, there is a concern that simple statistical learning mechanisms—and not the more sophisticated graphical model approach that we favor—can underlie children's causal inferences.

Bayesian Inference

As discussed above, one of the main motivations behind adopting the causal graphical model framework as a description of children's learning is that it brings precision to the otherwise vague description of children's learning provided by theory theory. These models allow us to specify how children represent causal structures and the connections among events that they observe in the world. But our review of the Rescorla-Wagner model in the last section suggests that these processes might not require such a sophisticated computational framework. It might be that young children are just

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

engaging in a form of *blocking*, where they are discounting one potential cause in favor of another event they already know is a cause.

Although it is well-established that this kind of discounting or blocking can be explained by associative models, other results are more challenging for these models. One example is "backward blocking," which just reverses the order of the blocking paradigm. In studies on backward blocking, similar to the blicket detector studies described above, adult participants observed two potential causes (A and B) produce an outcome. Then, one of those causes alone (A) produced the same outcome. Participants were less likely to judge that B was a cause than when they only observed A and B together (Shanks, 1985; Shanks & Dickinson, 1988). The Rescorla-Wagner model has difficulty explaining these data, because the associative strength of B is the same in both conditions.

To examine this developmentally, we (Sobel, Tenenbaum & Gopnik, 2004) implemented a backward blocking procedure with preschoolers. We first showed 3- and 4-year-olds two objects (A and B), which activated the machine together. Then we placed object A on the machine alone. In some conditions, A did not activate the machine by itself. Here, children should make a similar inference as they did in the Gopnik et al. (2001) study described above: object A was just "along for the ride" in this case, and only object B was efficacious. Unsurprisingly, given the earlier results, this was exactly the inference that they made. (It is worth noting that if you just show children that objects A and B together make the machine go, and then ask them about object A, children usually say that it is a blicket.)

The more interesting case is when A did activate the machine by itself; this is a version of the backward blocking paradigm. So A and B activate the machine together, and then A activates the machine by itself. In this case, what is the efficacy of object B? The correct answer is, "I have no idea." This is because object B's causal efficacy is ambiguous. But given that children now have an explanation for why the machine activated when A and B were on it together (i.e., object A unambiguously makes the machine go), they might explain away object B as a potential cause. And this is basically how children responded. Four-year-olds, for example, stated that object B had efficacy only 13% of the time. Other researchers (e.g., McCormack et al., 2009) have generated similar findings on slightly older children, further demonstrating that children are retrospectively reevaluating the probability that objects have causal efficacy.[14]

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

Although these data are a challenge for the Rescorla-Wagner model, many other theories of causal inference, particularly those from the literature on adult cognition, can account for them (e.g., Krushke & Blair, 2000; Van Hamme & Wasserman, 1994; Wasserman & Berglan, 1998). Similarly, other investigations have proposed that causal learning relies on the estimation of causal parameters, again based on the frequency with which events co-occur (Allan, 1980; Cheng, 1997; Jenkins & Ward, 1965; Shanks, 1995). What these models all have in common is that they use statistical learning principles to make a calculation about the causal strength between stimuli. What differs among these accounts is the math, not necessarily the way in which causal inferences are made. Further, most of these models could be categorized as making a lower-level kind of inference, as opposed to using the kinds of graphical representations we are advocating for (although see Glymour & Cheng, 1998).

Even if backward blocking can be accommodated on these models, we believe that there are other patterns in children's reasoning that cannot straightforwardly be explained within an associative learning framework. The example we work through here involves children's use of *base rates*—the frequency with which an event tends to occur in the environment. Some work with the blicket detector shows that even young children can track the frequency with which objects activate the machine and use that information when making inferences about ambiguous cases. That is, instead of just explaining away ambiguous data, children assume that they should default to the base rate of blickets when they do not know whether something is a blicket.

In one study on children's use of base rate information, we (Sobel, Tenenbaum & Gopnik, 2004, Experiment 3) showed 3- and 4-year-olds a set of identical objects and put 12 of those objects on the machine, one at a time. Either 2 of the objects or 10 of the objects activated the machine; there were different base rates of efficacious objects across the two conditions. Then, we showed children that the 13th and 14th objects (we'll call these objects A and B) together made the machine go, and then that object A made the machine go by itself. When the base rate of blickets was low (2 out of 12 objects had activated the machine in the initial demonstration), 4-year-olds judged object B to be a blicket approximately 16% of the time. When the base rate of blickets was high (10 out of 12 objects had activated the machine in the initial demonstration), 4-year-olds judged object B to be a blicket approximately 83% of the time. This was almost identical to the base rates.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

These results cannot straightforwardly be explained via association because the associative relation between object B and the machine is the same in both conditions, yet children treat them quite differently. These results are more parsimoniously explained by algorithms that involve the causal graphical model framework. Specifically, the modeling approach that Sobel et al. (2004) suggested was based on Bayesian inference (following Griffiths & Tenenbaum, 2009; Tenenbaum & Griffiths, 2001), which takes us back to theory theory. According to theory theory, children have some knowledge about a situation (i.e., their theories). In modeling terms, this knowledge allows them to construct a set of hypotheses about the world, in which each hypothesis has a probability value attached to it (*priors*). Then, they observe new data, and they use those data to update their priors so that each hypothesis now has a new probability value (*posteriors*). This allows children to construct new representations of the world—that is, to learn. Bayesian inference provides a formal approach for describing exactly how this learning happens, specifically, how individuals update their priors based on data. It is expressed using this formula:

$$P(H \mid D) = (P(D \mid H) * P(H)) \,/\, P(D)$$

Here, H stands for "hypothesis" and D for "data." The initial term, P(H | D), expresses the probability that the hypothesis H is true given the current set of data (D). Bayes' theorem allows us to calculate this probability as a function of three other terms: the probability that we would observe the data that we have observed if the hypothesis were true, P(D|H), the probability of observing the data, P(D), and the probability that the hypothesis is true, P(H). Bayes' theorem provides a formal description of how different hypotheses are weighed against each other given the data from the world and given how existing knowledge can be rationally integrated with observed data to make novel inferences. Soon after this idea of Bayesian inference was incorporated into theory theory, the theory was rebranded as a new form of constructivism we will refer to as *rational constructivism*. The goal of rational constructivism is to focus on the ways that children use their existing knowledge to make new inferences about the world based on the data they observe in this rational way, which allows for a formal definition of how theories can change (Gopnik & Wellman, 2012; Xu & Kushnir, 2012; Xu & Tenenbaum, 2007[15]).

Algorithms that use Bayesian inference can provide a good description of how children reason. Specifically, they can explain how young children's

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

existing knowledge (their priors) can constrain future inferences. Children's use of base rates in their reasoning provides a particularly powerful example of this, because this procedure makes it clear how different sets of prior information lead children to make different inferences at the end of the procedure.

Interestingly, although the 4-year-olds tested in the base rates study described above were able to do this, the 3-year-olds were not. They just said that B was a blicket in both conditions. Why were only the older children in this procedure able to attend to the base rate with which objects have causal efficacy? One possibility is that the 3-year-olds did not have the same priors (i.e., knowledge) as the older children.[16]

To succeed at the base rates task (and any causal reasoning task involving the blicket detector), children must register that there are at least two types of entities in the environment: objects and detectors. Blickets are objects that have the ability to activate a detector. Children are told all of this information in the beginning of the experiment. What they have to infer is an unobserved attribute of the objects—that some of the objects in front of them are blickets. To do so from the evidence, they must possess three pieces of prior knowledge. The first two are *temporal priority* and *spatial independence*. Temporal priority allows children to understand that certain objects being placed on the detector are responsible for the detector's activation, as opposed to the idea that the detector's activation causes the experimenter to place an object on it. Spatial independence allows children to understand that the identity of an object is independent of its spatial location and the spatial location of all other objects. That is, placing one object on the detector does not cause another object to become a blicket or cause another object to be moved in space. These are fairly basic assumptions about how causality works in general; as discussed earlier, there is some evidence that young children and even infants make such assumptions (e.g., Bullock et al., 1982; Leslie & Keeble, 1987; Oakes & Cohen, 1990; Sophian & Huber, 1984).

The third assumption is what Sobel et al. (2004, following Tenenbaum & Griffiths, 2001) called the *activation law*: the machine activates if and only if at least one blicket is placed on top of it, and only blickets make the machine activate. Children had to recognize that there was something about the object that connected its being placed on the machine to the machine's activation. A candidate cause for that "something" was a nonobvious property that the object possessed.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

To test whether children made this assumption, we (Sobel, Yoachim, Gopnik, Meltzoff & Blumenthal, 2007) ran a new blicket detector study where we presented 3- and 4-year-olds with three objects. Two were identical; one was unique. One of the two identical objects and the unique object activated the machine. The other object did not. Children were then shown that the member of the identical pair that had activated the machine also had an internal property (e.g., a piece of hard plastic in its center). They were asked which among the two other objects had the same inside. Four-year-olds in this study responded based on the causal efficacy and chose the unique object. However, the 3-year-olds responded based on perceptual similarity. These data suggest that, between the ages of 3 and 4, the children are developing the understanding that there is something inherent about objects that make the machine go. This in turn can explain the difference in their performance in Sobel et al. (2004): Only 4-year-olds could attend to the base rates because only they could apply their understanding of the hidden properties of blickets to the objects.

To make this case more strongly, one of our later experiments (Sobel & Munro, 2009) connected these two ideas. In this experiment, we replicated the Sobel et al. (2004) procedure in which causes were rare and the Sobel, Yoachim et al. (2007) "insides" procedure and found strong correlations, controlling for age. It did not matter whether children were 3 or 4; what mattered was their understanding of the relation between causes and insides. If children understood that an object's causal efficacy related to its internal properties, then they were likely to use the base rate information and say that object B was not efficacious in the final trial. This experiment also manipulated the way the machine's activation was described to children. Sometimes, it was presented as a machine; other times it was presented as an agent ("Mr. Blicket") who liked things. Desire is a mental state that 3-year-olds understand well (e.g., Repacholi & Gopnik, 1997; Wellman & Liu, 2004). In this condition, 3-year-olds were almost always able to correctly reason about the relation between the activation and the internal properties of the objects, as well as use base rates in their reasoning.

This study provides a clear illustration of how content knowledge constrains causal reasoning, as discussed in chapter 1. It also continues to show how the causal graphical model framework and the Bayesian inference procedure for updating beliefs on the basis of new evidence provide a better

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

match to children's behavior than do models of associative learning. It also illustrates another feature of this framework: the idea that children represent a *hypothesis space* of potential causal models. Under the causal graphical model framework, one looks at the data that one has available and constructs a set of possible explanations for how those data could have come about. Bayesian inference can then be used to select which explanation (that is, which hypothesis in the space of possibilities) is the best fit to the data. This process can be used to describe children's reasoning behavior in tasks like the ones described above. The exact details of Bayesian inference are not important for the argument we want to make here (see Gopnik & Bonawitz, 2015, for a good introduction to this topic). What is important is the question of how a hypothesis space is formed in the first place. We propose that this mechanism relies on our imagination: We must extrapolate away from the observed data in order to think about all of the *possible* mechanisms that could have caused it. This proposal illustrates the tight links between imagination and causal reasoning, which we explore further in chapter 10.[17]

In general, though, these data all suggest that children's causal reasoning is well-described by Bayesian inference, which allows children to use their existing knowledge (e.g., about base rates) to construct hypotheses about what they will observe in a new situation and to update those hypotheses on the basis of new information. In addition, numerous researchers have applied these ideas to other domains (e.g., Schulz et al., 2007) and to more domain-general learning problems (e.g., Goodman et al., 2011). More recently, Bayesian approaches have offered interesting and important contrasts (as well as supplements) to approaches in machine learning and artificial intelligence (e.g., Lake et al., 2017). They thus constitute a promising way for describing how children learn in general.

A Concern about Mechanisms

There is one thing that we have left out of our discussion of causal graphical models and Bayesian inference so far (actually, there are quite a number of things that we have left out, but we are trying to restrict our discussion to what will be useful for our later arguments about scientific thinking). This is the issue of *mechanisms*, which are important for using causal graphical

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

models as representation of human causal knowledge (see Goodman et al., 2011). A mechanism is a description of how the nodes in a graph are related to one another. Knowing something about mechanisms helps you relate the data that you observe to the model that you represent.

More specifically, in a graphical model, a vertex (the arrow that goes from one node to another) represents a dependency: a probabilistic relation between one object, event or property and another object, event or property. Although this seems simple, it raises a problem: Even the simplest causal graph ($X \rightarrow Y$) can be consistent with an infinite set of interpretations because we do not necessarily know what the probability of that relation is. The dependency between Y and X might be near deterministic (e.g., the probability that Y occurs given that X occurs might be .99), or it could be .63, or .17, or any real number between 0 and 1. But in all of these cases, the graph itself is the same. If graphical models are meant to be representations of our causal knowledge, how do we know how to interpret them, given that any particular representation is compatible with so many different possible structures?

It gets worse. Think about a graph in which two events each cause a third (see figure 2.7). Here, events X and Y are independent, but each cause Z. But what is the nature of the relations among X, Y, and Z? A natural way to think about this graph is that Z occurs either if X occurs or if Y occurs, with some probability. That is, this graph could represent a set of disjunctive relations, called a "noisy-or" parameterization when the relations among the variables are not deterministic. Such relations are fairly common in everyday causal reasoning, which might be why it is so easy to think this way. But this is only one possible interpretation of this graph. It could also be the case that X and Y both have to occur to cause Z (i.e., a conjunctive parameterization). Or it might be the case that only either X or Y is necessary for Z to occur, but the other is probabilistically related to Z.

Yet another issue with understanding the mechanisms represented by a graphical model is the question of abstraction. To illustrate this issue, consider two possible causal models. In one, you posit that diseases cause symptoms. In another, you posit that having a cold causes you to have a runny nose. When you observe that individuals who have colds also have runny noses, which model do you learn? This distinction between learning "specific" theories and "framework" theories (Wellman & Gelman, 1992) is one of the most difficult to describe in causal learning.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

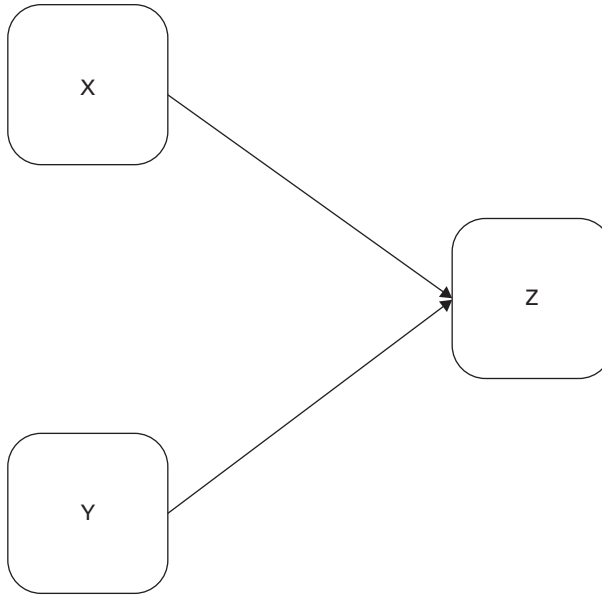


Figure 2.7
A common effect model. Even simple models like this one can represent a wide array of possible causal structures.

The answer might be that you learn both of these models; as you observe data in the world, you can formulate a representation not only of the specific causal structure (between runny noses and colds), but also an abstraction of broader frameworks or types of causal relations (between diseases and symptoms). Goodman et al. (2011) refer to this as the "blessing of abstraction."[18] They suggest that hierarchical modeling can allow learners to do both at once: to reconstruct both the specific causal model specified by the data and also the general principles (which they call "theories") that govern how such representations of causal structure should be constructed. Once those general principles are in place, they can also constrain subsequent causal learning given new data. Although human reasoners can do this, at least to a certain extent, it remains a challenge to understand which level(s) of abstraction people are thinking about in any given circumstance.

These problems occur largely because causal graphical models were not originally designed to represent human cognition; they are just formal tools.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

But to make them do useful work for psychology, we need to sort out exactly what kinds of relations people are representing. The good news is that, at least with respect to the issue of different types of parameterizations as illustrated in figure 2.7, it turns out that young children can recognize different forms that these relations can take. For example, when shown evidence for disjunctive or conjunctive causal relations, 4- and 5-year-olds have no difficulty telling the difference (Lucas et al., 2014). They also do not have much difficulty discerning the difference between deterministic and probabilistic environments (Griffiths et al. 2011).

Because of this, we think that even young children posit that there are some kinds of mechanisms that relate different events to each other, and they do so with reasonable sophistication even from a very young age (similar to how the infants in the Madole and Cohen (1995) experiment we described above recognized which regularities to attend to). Although this is not a formal part of the causal graphical model framework (at least not in the computations), it provides a good starting point for thinking about the relation between a representation of causal knowledge and the world. But this is merely a starting point, and it does not fully resolve all the complexities of using causal graphical models to describe human learning that we outlined above. We return to these issues in chapter 3, where we talk about some potential challenges to the causal graphical modeling framework as a description of human causal reasoning.

More generally, what this discussion of mechanism tells us is that children are faced with a challenge, which is how they instantiate the models that they represent. In blicket detector tasks, this is basically done for the children. They can see that there is a machine that activates and that there are objects that are potential causes of that activation. Moreover, they are often told by a (presumably) knowledgeable experimenter that the machine is a blicket machine and that objects that activate the machine are blickets. The pedagogical situation between the experimenter and the child establishes that there are blickets in the world and that this machine will help figure out whether objects get that label. In situations that involve richer real-world contexts, however, there might be an additional problem for children: They not only have to learn how the variables relate to one another, they also have to grapple with the information presented by the context. We return to this discussion when we explore the role of contextualization in children's scientific thinking in chapter 6.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

Levels of Explanation

The causal graphical model framework outlined above provides a precise way to discuss children's (and adults') causal reasoning abilities and how causal relations between objects and events can be represented and updated (i.e., learned). The blicket detector was used to test whether children's causal inferences indeed match with this formalism, and results from studies using the detector indicate that they generally do. In later chapters, we examine how more complex causal reasoning tasks can be presented using blicket detectors, in order to test other aspects of children's causal reasoning. For now, the important point is that this reasoning can be well represented with causal graphical models, providing a more precise language for describing how children learn about the causal structure of the world.

We (along with many other researchers) are in favor of the causal graphical model description of causal reasoning because these models—in combination with algorithms like Bayesian inference that describe how one might build up these models from observed data and prior beliefs—allow us to accurately and formally represent children's reasoning processes. But it is important to recognize what these models are supposed to be doing in our theory of cognitive development.

In his classic book on vision, Marr (1982) noted that it was possible to speak about a symbolic system at different levels. One could focus on exactly how that system is instantiated in the substrate that makes it up (e.g., neurons for humans; silicon chips for computers); this is the implementation level. Or one could focus on the steps that the system takes to perform its computations; this is the algorithmic level. Or one could focus on the general processes that the system performs and what it does; this is the computational level. To take a concrete example, both a smartphone and an abacus can be used to do addition, but they have different algorithms for doing so, which are implemented in different ways. But at the computational level, they are similar. They take representations of two numbers and produce a sum.

We take the causal graphical model framework to be describing children's reasoning processes at the computational level. These models offer a description of how children learn causal knowledge and make causal inferences, but they are neutral about exactly what steps children go through in order to implement this process and also about exactly how those steps are represented and carried out in the brain.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

We discuss these distinctions among levels of explanation in order to note that the causal graphical model framework is not meant to describe exactly how children go through their reasoning processes (i.e., at the algorithmic level). The framework is meant to provide observations about children's reasoning processes, and it is meant to be used to make predictions about how children will behave in certain circumstances, given certain kinds of data. But even though it describes those processes precisely and formally, it is not meant to say anything further about exactly *how* those processes are carried out in human minds. That is, it can be a successful representation of children's reasoning processes (at the computational level) without being a fully detailed description of all the processes children undergo when reasoning (at the algorithmic level).

This discussion makes it clear that we do not yet have an algorithmic-level or an implementation-level description of children's reasoning. That is, we do not yet know exactly how the causal graphical models are instantiated in children's neural structures (or even if they are—it is probable that the instantiation is quite different and only simulates such a modeling framework). This leaves open the question of what is happening at the algorithmic level. If our brains are not genuinely representing causal graphs, then what underlies our abilities to use these computational tools?

Some researchers are currently investigating this issue (e.g., Bonawitz et al., 2014). One likely possibility is that the brain is sampling information in a way that mimics the representational structures and algorithms described by causal graphical models (e.g., Sanborn & Chater, 2016). Although this view is satisfying to some, others are not convinced. The dissatisfaction has mostly centered on the Bayesian inference part of the framework (Jones & Love, 2011; Marcus & Davis, 2013). Without reviewing all of these arguments, our goal for this section is simply to say that not all researchers believe that the causal graphical model framework is the be-all and end-all of human causal reasoning. However, this framework does present an abstract, coherent way to represent causal knowledge and describe causal inference, one that can be useful in constructing explanations and making predictions. For developmental psychologists interested in looking at the relation between children's causal reasoning capacities and other facets of cognitive development, this suffices.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

As an example, one of our studies (Yang, Bushnell, Buchanan & Sobel, 2013) presented 15-month-olds with a novel object (X) that had a manipulandum, such as a bright red button (call this A). In this study, we showed the infants that activating the manipulandum (pushing the button) caused a nonobvious effect from the toy (e.g., it would make a mooing sound). Infants were allowed to imitate this novel effect on the object, and most did so. We then showed the infants an identical object (X) with a different manipulandum, such as a blue lever (call this B), which did nothing when manipulated. Again, infants imitated the action on this toy. The critical part of the study was when we showed infants a third identical object (X) that had both manipulanda on it (A and B; see figure 2.8). Which one would infants go for? Could they generalize what they had observed in imitation to produce a causal action on a novel object?

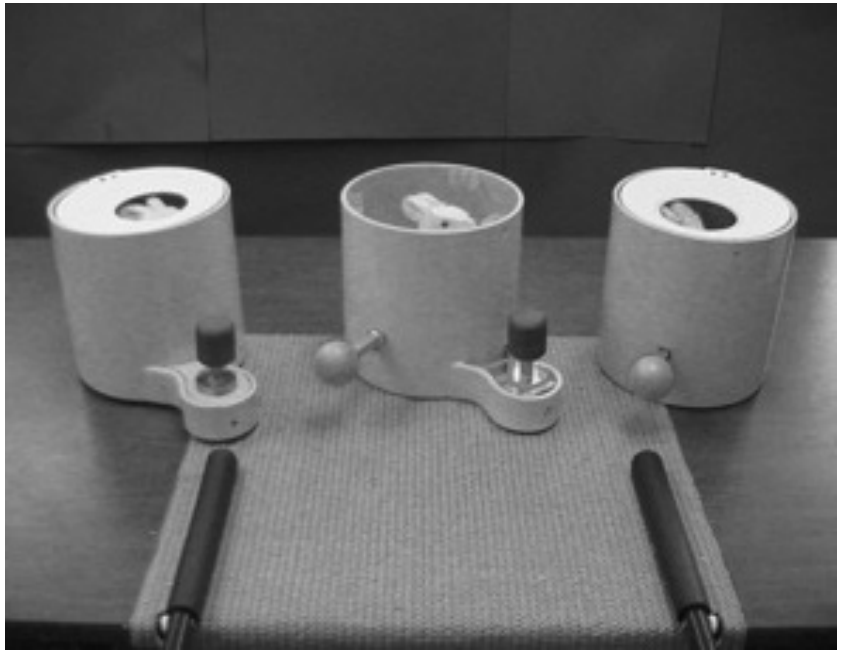


Figure 2.8
Stimuli from Yang et al. (2013). The objects used in the demonstration phase are on the left and right, and the object used in the test phase is in the center.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

The answer is yes—infants acted on whichever manipulandum they had previously seen make the toy do something interesting (in this example, the button). But they only did this when the third toy—the one with both manipulanda—was identical to the first two toys. In a second experiment, we showed infants the same demonstration, but now the test object (Z) looked completely different, although it had the same two manipulanda (A and B). In this case, infants did not generalize—they went for either the button or the lever at random. These data suggest that infants' generalization capacities are relatively weak, and they were not able to identify that their actions on the manipulanda were the cause of the effect. When they saw the novel stimulus, they resorted to an irrational basis for deciding what to do.

But perhaps infants were behaving rationally in both cases. In both experiments, infants had observed a presumably knowledgeable adult model the efficacy of two objects that looked quite similar but behaved differently. Infants also imitated both of those actions, indicating that they could perform both actions and that they had some motivation to interact with both objects. When the test object with both manipulanda was the same, they generalized because they recognized that the action related to that object. But when the test object was different, there was no reason to generalize. Because the base object was different, this could plausibly indicate that it operated according to a completely different set of causal relations. It is reasonable to assume that pushing a button might work on toys of type X, but not all toys in general. There is potentially also a Bayesian analysis of this explanation: The demonstration phase of the experiment does not provide enough information to support the posterior hypothesis that one should generalize when the test object changes shape.

This line of reasoning suggested a third experiment. Instead of presenting two similar objects with different manipulanda during the demonstration and then a novel object with both manipulanda at test, this new study presented three different objects: First, children saw object X with manipulandum A, which caused an effect, and object Y with manipulandum B, which did not do anything. Then, children were shown the test object, Z, with both A and B on it. Fifteen-month-olds generalized in this case; they chose to intervene first on the manipulandum that was previously shown to be efficacious.

We also presented a formal computational model of these data, based on Bayesian inference over causal graphical models. The model wound up

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

explaining the data nicely, including one completely unanticipated aspect of the results. In the first experiment, in which all three manipulanda had the same base shape, the model predicts that the likelihood of the previously seen efficacious manipulandum being efficacious on the test object is about 80%. But in the third experiment, in which all three objects have different bases, the model predicts that this likelihood is only 65%. In both cases, given that they have two choices at test, children should still attempt to act on the manipulandum that was previously shown to be efficacious; in both cases, it is more than 50% likely to work.

But in both of these studies, the test object—the one with two manipulanda—was inert, so that even when the infant manipulated the previously seen efficacious manipulandum, the object did not do anything. The question is then what children should do when they observe that this manipulandum is not efficacious on the test object. The model suggests that they should be quicker to switch to acting on the other manipulandum in the third experiment (where the likelihood of the first manipulandum working is 65%), as opposed to in the first experiment (where the likelihood of the first manipulandum working is 80%). This is exactly what the infants did. They persisted more with the first manipulandum when all of the objects were the same than when all of the objects were different.

We do not know if this is a good computational model or a bad one, or a good set of experiments or a bad set. That's not the point. Rather, the point is that the model helped explain some data we had collected and made novel predictions about new data. It even explained something completely unexpected in the data—something that we did not even know we were looking for.

Computational modeling is now a large part of cognitive science. Explaining behavior via such models is useful because it provides a formal account of what kinds of inferences children are making—not necessarily how they are making them, but what they are and are not capable of. We have presented a particular framework, but that does not mean that other computational frameworks are not also good descriptions. Regardless of whether causal graphical models or Bayesian inference mechanisms are instantiated as such in our brains, or whether other computational models provide a better description of children's reasoning, our goal is to use these frameworks productively in developmental science to better explain children's behavior and learning.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/11939.001.0001

One of the things that we are going to do throughout the rest of this book is use this framework to describe aspects of the development of children's causal reasoning, particularly as it relates to scientific thinking. This does not mean that one has to know about these computational frameworks or even pay attention to them in order to understand our arguments. Rather, we articulate this framework here because we find it helpful in guiding the theoretical arguments we are trying to construct.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

This is a section of doi:10.7551/mitpress/11939.001.0001

Constructing Science

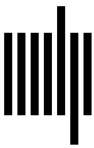
Connecting Causal Reasoning to Scientific Thinking in Young Children

By: Deena Skolnick Weisberg, David M. Sobel

Citation:

Constructing Science: Connecting Causal Reasoning to Scientific Thinking in Young Children
By: Deena Skolnick Weisberg, David M. Sobel
DOI: 10.7551/mitpress/11939.001.0001
ISBN (electronic): 9780262370615
Publisher: The MIT Press
Published: 2022

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open



The MIT Press

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024

MIT Press Direct

© 2022 Massachusetts Institute of Technology

This work is subject to a Creative Commons CC-BY-ND license. Subject to such license, all rights are reserved.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Weisberg, Deena Skolnick, author. | Sobel, David M., author.
Title: Constructing science : connecting causal reasoning to scientific thinking in young children / Deena Skolnick Weisberg and David M. Sobel.
Description: Cambridge, Massachusetts : The MIT Press, [2022] | Includes bibliographical references and index.
Identifiers: LCCN 2021045987 | ISBN 9780262044684 (paperback)
Subjects: LCSH: Science—Methodology. | Reasoning in children. | Scientific ability. | Science—Study and teaching—Psychological aspects. | Constructivism (Education)
Classification: LCC Q175.32.R45 W45 2022 | DDC 501—dc23/eng/20211214
LC record available at https://lccn.loc.gov/2021045987

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253982/c001100_9780262370615.pdf by guest on 09 August 2024