CHAPTER SEVEN

Edge Replacement and Minimality as Models of Causal Inference in Children

David W. Buchanan*,1, David M. Sobel†

*IBM T.J. Watson Research Center, Westchester, New York, USA
†Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, Rhode Island, USA

1Corresponding author: e-mail address: dbuchan@us.ibm.com

Contents

1. Introduction 184
2. Causal Graphical Models, Minimality, and Edge Replacement 185
   2.1 Stream Location 191
3. Evidence for Determinism 193
   3.1 Hidden Inhibitors and Time 194
   3.2 Variability and Complexity 196
4. Evidence for Branching: Stream Location in Preschoolers 198
   4.1 Experiment 1 199
   4.2 Experiment 2 203
   4.3 Experiment 3 206
5. General Discussion 209
Acknowledgments 212
References 212

Abstract

Recently, much research has focused on causal graphical models (CGMs) as a computational-level description of how children represent cause and effect. While this research program has shown promise, there are aspects of causal reasoning that CGMs have difficulty accommodating. We propose a new formalism that amends CGMs. This edge replacement grammar formalizes one existing and one novel theoretical commitment. The existing idea is that children are determinists, in the sense that they believe that apparent randomness comes from hidden complexity, rather than inherent non-determinism in the world. The new idea is that children think of causation as a branching process: causal relations grow not directly from the cause, but from existing relations between the cause and other effects. We have shown elsewhere that these two commitments together, when formalized, can explain and quantitatively fit the
otherwise puzzling effect of nonindependence observed in the adult causal reasoning literature. We then test the qualitative predictions of this new formalism on children in a series of three experiments.

1. INTRODUCTION

Before children enter elementary school, they come to possess a rich array of causal knowledge across a variety of domains (e.g., Gelman & Gottfried, 1996; Gopnik & Meltzoff, 1997; Inagaki & Hatano, 2004; Kalish, 1996; Shultz, 1982). Cognitive scientists have considered causal graphical models (CGMs) as a computational-level (Marr, 1982) description of how children (and adults) represent this causal knowledge (e.g., Gopnik et al., 2004; see also Griffiths & Tenenbaum, 2005; Rehder & Burnett, 2005; Sloman, 2005). In the CGM framework, events, properties, or objects are represented as nodes and causal relations are represented as directed edges linking these nodes. A graph’s structure thus specifies dependence and conditional independence information relating events.

Our goal is to compare two ways of deploying the CGM framework as a psychological theory about human causal representations. The first theory is a relatively direct adaptation of CGMs as they are used in the statistics and artificial intelligence community. It is also the version of CGMs that has been most widely applied within cognitive psychology and cognitive development (e.g., Gopnik et al., 2004). This theory uses the principle of minimality: when constructing a representation of causal structure, use as few edges as possible. The second theory uses a novel formalism called edge replacement (Buchanan, Tenenbaum, & Sobel, 2010): representations of causal knowledge are built from a particular set of rules governing how causal graphs can be constructed, which we will specify below. The theories differ most acutely in their mechanisms for constructing initial representations of novel causal systems. For this reason, we test their differing experimental predictions on children, who have less experience with causal systems than adults do.

We begin by giving an overview of CGMs, with a focus on the differing implications of minimality and edge replacement. We then briefly review a portion of the adult literature that shows better quantitative fits to edge replacement. We then turn to testing qualitative predictions on children. We first review two sets of experiments, published elsewhere, that support the idea that the core predictions of edge replacement are correct. The main
body of the paper presents in detail three new experiments that test for a distinction between edge replacement and minimality, which we call *stream location*. We conclude that edge replacement better explains various aspects of children’s causal reasoning and discuss the implications of this conclusion.

2. CAUSAL GRAPHICAL MODELS, MINIMALITY, AND EDGE REPLACEMENT

A CGM consists of nodes that represent events, properties, or features, and edges, which represent causal relations. Edges are always directed in a CGM and point from *head* to *tail*. There are multiple ways to parameterize a CGM (i.e., define the rules by which nodes interact via their edges), but we will focus on the canonical *noisy-logical* formulation (Cheng, 1997; Yuille & Lu, 2007). Under this formulation, a node is on if it is caused, but not prevented, and off otherwise.

There are two kinds of edges: generative and preventative. Both generative and preventative edges have a *power*. In the generative case, this is the probability that the head, if on, causes the tail to be on (if the tail is not prevented). In the preventative case, this is the probability that the head, if on, causes the tail to be off. Nodes can also have a background rate—the probability that they are caused by something not otherwise represented in the graph.

The semantics of a CGM are such that all dependencies between nodes are assumed to be represented in the graph. Through the rules of the CGM formalism, we can infer the patterns of dependency between any pair of nodes given any state of knowledge. Importantly, if we cannot find any such dependency in the graph, the graph specifies the relation to be one of statistical independence.

To clarify the meaning of CGMs, consider an example: Alice uses her cellular phone to call her friend Cindy’s phone. When Alice does this, Cindy’s phone rings 80% of the time. Figure 7.1a shows a simple CGM that captures this relation: a single directed edge connects these two events, with the edge having a power of 0.8. We can also create graphs that relate multiple variables. For instance, imagine that Alice and Bob can both call Cindy, each with 80% efficacy. Figure 7.2a shows this set of causal relations: two edges connected to the outcome, each with an independent causal power of 0.8. If both Alice and Bob call Cindy, we calculate the probability that Cindy’s phone rings with a simple rule of probability (based on a
noisy-OR logic): \( p(\text{Alice’s call got through}) + p(\text{Bob’s call got through}) - p(\text{both}) = 0.8 + 0.8 - (0.8)^2 = 0.96. \)

CGMs are highly expressive. Yuille and Lu (2007) have shown that the above formulation can capture any causal–functional relation (i.e., any particular pattern of conditional dependencies between events) that we might care to represent. This expressivity, however, creates a problem: many possible graphs can capture the same set of causal relations. For instance, Fig. 7.1a–c shows several graphs that could capture the cell phone example. Instead of a single probabilistic relation of power 0.8, there could be a set of

![Figure 7.1](image1)

**Figure 7.1** Three CGMs that all represent the cellphone example. The node marked A corresponds to the event where Alice calls Cindy. The node marked “C” corresponds to the event where Cindy’s phone rings. Edges marked with a number have the causal power indicated by the number. Unmarked edges have power 1.0. Nodes marked with a number have that indicated background rate. Unmarked nodes have a background rate of zero. Dashed edges indicate an inhibitory relation. Note that all three graphs capture the same causal–functional relation between A and C: when A is active, C is active 80% of the time, and inactive otherwise.

![Figure 7.2](image2)

**Figure 7.2** A minimal (a) and nonminimal (b) representation of the cell phone example involving three phones.
deterministic relations, mediated by a hidden inhibitor that fires 20% of the time (Fig. 7.1b). There could also be a series of edges connecting Alice’s call to Cindy’s phone (Fig. 7.1c). There is actually an infinite set of CGMs that can capture this relation, most of which are quite complex. This is because for any CGM that expresses a given relation, we can always construct a more complex CGM that also expresses the same relation.

To make use of CGMs, we need to determine which representations to use. We will call this the complexity problem. A natural solution is to use a version of Occam’s Razor: prefer the simplest representation that fits the data. Applying this principle, however, requires a definition of simplicity. One natural choice is to count the number of edges in the representation. This is known as minimality: a graph is minimal if no other graph that fits the data has fewer edges. Minimality allows us to choose Fig. 7.1a to represent the cell phone example in a principled way—because it has fewer edges than Fig. 7.1b or c.

Minimality is one solution to the complexity problem. If we are going to treat CGMs as a good description of the way children (and adults) represent causal knowledge, either we must specify that they obey minimality when they construct graphs to represent causal structure, or we must describe an alternative principle by which they do so. The overarching goal of this chapter is to suggest that minimality makes a set of counterintuitive and incorrect commitments about how children represent causal knowledge. We outline two of these commitments below and then suggest an alternative principle that we believe is more structured, coherent, and consistent with the way in which children make causal inferences.

The first commitment concerns causal systems where the relations sometimes fail. In these situations, minimality prescribes that human beings represent the causal relation as inherently probabilistic. This is because representations that involve hidden inhibitors have more edges than representations that have no inhibitors. In our example, minimality strictly prescribes that we prefer Fig. 7.1a over Fig. 7.1b, unless we have independent evidence for the existence of the node marked “?.”

A second prediction concerns conditional independence. Consider the common effect representation in Fig. 7.2a: if either Alice or Bob call Cindy, Cindy’s phone will ring with a probability of 0.8. Consider a case where Alice tries to call Cindy and fails. Does this datum make it more likely that Bob’s call will fail as well? Although we may find it intuitive to say “yes,” minimality says “no.” In the minimal representation of Fig. 7.2a, the first failure is a random occurrence that occurs independently of any relation
between Bob’s phone and Cindy’s phone. Under the statistical rules of CGMs, causes produce their effects independently of each other. This means that the efficacy of Alice’s attempt should give us no information about whether Bob’s attempted call will succeed. In contrast, if we think that Alice’s failure makes it more likely that Bob’s call will fail as well, then we must be using a more complex graph, such as the one shown in Fig. 7.2b (i.e., one contrary to minimality). The belief that Alice’s failure gives information about Bob’s potential failure, or more generally, the idea that the efficacy of one relation in a common cause gives information about another relation (or that one common effect gives information about another, given the cause) is known as \textit{nonindependence}. In contexts where minimality is considered normative, nonindependence is seen as a fallacy of reasoning. Data from multiple experiments (e.g., Rehder & Burnett, 2005; Walsh & Sloman, 2005) indicate that adults show a robust non-independence effect in causal reasoning.

These two counterintuitive aspects of minimality are potentially concerning, but in the absence of alternatives, minimality is widely used. Elsewhere (Buchanan et al., 2010), we have proposed an alternative to minimality, called \textit{edge replacement}. We advocate for it here as a psychological theory of children’s causal representations. The core commitment of edge replacement is that the human mind solves the complexity problem not with minimality, but by using a particular process to generate causal representations (see Fig. 7.3 for an example):

1. Begin with a single, deterministic edge (i.e., power = 1) connecting a single cause and a single effect.
2. Replace the edge with an edge–node–edge combination (also deterministic) that incorporates a relation with a new node. The replacements are either inward or outward, and either generative or inhibitory. For instance, an inward inhibitory replacement creates an event that will prevent causal power from flowing through the relation that it replaces. An outward generative replacement creates a new side effect. An inward generative replacement creates a new alternative cause. (Outward inhibitory replacements are allowed, but are only needed for formal completeness.)
3. Repeat this process on the newly created edges until the representation is consistent with the observed data.

There are further formal details in the full treatment of edge replacement (Buchanan et al., 2010), but the above specification will be sufficient for our purposes in this chapter.
To elucidate the process in more detail, we can apply edge replacement to the cell phone examples. In the simple example where Alice (A) calls Cindy (C), we begin with a single edge with power 1 connecting A to C. Because the relation sometimes fails to hold, the data are inconsistent with this single deterministic edge. Unlike minimality, edge replacement does not allow probabilistic edges. Instead, apparently probabilistic relations are represented as involving hidden inhibitors that are active some of the time. We apply the edge replacement process to add one step of complexity to the graph. We replace the single edge with a path incorporating the influence of a hidden inhibitor, as in Fig. 7.1b. Because Cindy’s phone rings 80% of the time, we posit that the hidden inhibitor is active 20% of the time. Because this representation is consistent with the data, we stop.

A further commitment of edge replacement is that the hidden inhibitors we posit exist not just functionally but physically. That is, they have extent in space and time and respect spatial and temporal relations. This means that the node marked “?” in Fig. 7.1b really represents something in the world that we expect to be able to discover. If it prevents the relation between Alice and Cindy’s phone, then it must exist in some location in space between Alice’s phone and Cindy’s phone, and must be active at the same the time the

Figure 7.3 A series of edge replacements that construct a graph consistent with the cell phone example involving three phones. Each of (c) and (d) represents a different replacement that could have been made from (b). The node marked “M” is a bridge node, whose purpose is to incorporate the influence of B under edge replacement.

To elucidate the process in more detail, we can apply edge replacement to the cell phone examples. In the simple example where Alice (A) calls Cindy (C), we begin with a single edge with power 1 connecting A to C. Because the relation sometimes fails to hold, the data are inconsistent with this single deterministic edge. Unlike minimality, edge replacement does not allow probabilistic edges. Instead, apparently probabilistic relations are represented as involving hidden inhibitors that are active some of the time. We apply the edge replacement process to add one step of complexity to the graph. We replace the single edge with a path incorporating the influence of a hidden inhibitor, as in Fig. 7.1b. Because Cindy’s phone rings 80% of the time, we posit that the hidden inhibitor is active 20% of the time. Because this representation is consistent with the data, we stop.

A further commitment of edge replacement is that the hidden inhibitors we posit exist not just functionally but physically. That is, they have extent in space and time and respect spatial and temporal relations. This means that the node marked “?” in Fig. 7.1b really represents something in the world that we expect to be able to discover. If it prevents the relation between Alice and Cindy’s phone, then it must exist in some location in space between Alice’s phone and Cindy’s phone, and must be active at the same the time the
relation is prevented. There might be a single, specific event, such as that Cindy’s battery is low or she is in an area with poor reception. Alternatively, we might have a vague placeholder for some mechanism that we do not understand (and that itself might be a more complicated graph). Edge replacement is committed to the idea that people expect these hidden causes to exist in the world, in a physical and temporal arrangement that is consistent with their functional properties in the causal representation.

To illustrate the process of edge replacement further, consider the three-phone network among Alice, Bob, and Cindy. Fig. 7.3 shows a step-by-step visualization of how edge replacement constructs these graphs. We begin with a single edge connecting Alice’s call (A) to Cindy’s phone ringing (C).1 The first thing to do is to incorporate Alice’s action into the graph. We do this by replacing the A → C edge with a path that incorporates Bob (B), as in Fig. 7.3b. Under this representation, Bob and Alice can both cause Cindy’s phone to ring, but so far the relations never fail. We can accommodate the presence of failures by introducing one or more hidden inhibitors, as we did in the first example. One way of doing this is shown in Fig. 7.3c. This representation is consistent with the data, so we stop.

Importantly, there are multiple representations consistent with the data that edge replacement could generate. For instance, in Fig. 7.3, we could have added the inhibitor to a different part of the graph, such as on the edge from A to M, instead of on the edge from M to C. The results of such a replacement are shown in Fig. 7.3d. When using the edge replacement process to examine how to represent a set of causal relations, the process generates a distribution over graphs: each individual graph is assigned a particular probability based on the frequency with which it is generated by the rules of edge replacement.2 Edge replacement predicts that participants will be most likely to choose the graph that best fits the data that they observe.

Consider now the difference between Fig. 7.2a (a representation consistent with minimality) and Fig. 7.2b (a representation more consistent with edge replacement). The conditional dependence relations in these graphs are subtly different. In Fig. 7.2b, if Alice fails to reach Cindy, we can rationally be less confident that Bob will be able to reach Cindy. This is because Alice’s

---

1 We could have started with a relation from Bob’s call as well—we would end up with the same representation.

2 We would refer readers interested in the appropriate level of mathematical detail to Buchanan et al. (2010). We hope this level of detail is unnecessary to appreciate the argument put forward in this chapter.
failure may have occurred because the node H is active. If H is active, it will also prevent Bob from reaching Cindy. In Fig. 7.2a, we cannot make this inference, and thus we should not expect to see nonindependence. In contrast, under edge replacement, the nonindependence effects from the adult literature are not only allowed, but prescribed as rational. Edge replacement often creates representations where causal relations are expected to share inhibitors, and these shared inhibitors create rational nonindependence. Indeed, edge replacement provides good quantitative fits to a set of published nonindependence effects in adults (Buchanan et al., 2010).

2.1. Stream Location

Edge replacement also makes an important qualitative prediction, which we will test empirically. Consider the cell phone example again: imagine that we have a piece of information that suggests why the call failed. For instance, after Alice and Bob try to call Cindy and fail, Alice changes the battery in her phone and is successful. If Bob tries again, will he be able to reach Cindy?

Contrast your prediction with a second case. Again, Alice and Bob both fail to reach Cindy. This time, Cindy changes the battery in her phone. Alice tries again and is now successful. If Bob tries again, will he be able to reach Cindy?

Edge replacement predicts that judgments of Bob’s efficacy should be higher than in the second case than the first. Why? The generative rules of edge replacement implicitly tend to construct graphs that have a branching character: causal relations in a common effect structure share some but not all of their causal structure with sibling relations, and the common part of the structure is more likely to physically close to the common effect. The intervention that changed the efficacy early in the causal stream (changing Alice’s battery) is more likely to have affected an inhibitor that only affects the Alice–Cindy relation (as in Fig. 7.3d). The intervention on a node late in the causal stream (changing Cindy’s battery) is more likely to have affected an inhibitor that is shared by both relations (as in Fig. 7.3c). We will call edge replacement’s prediction here the stream location effect.

What does minimality predict in the stream location example? Figure 7.4 shows the minimal graph for these causal relations. We created this figure by starting with a minimal common effect structure as in Fig. 7.2a and then adding a single edge to represent the effects of the intervention. Note that there is no difference in the structure of the graph between the early and late
interventions under minimality. These differences require intermediate structure, which minimality minimizes away.

There are some data from experiments on nonindependence that suggest that adults reason about stream location in a manner more consistent with edge replacement than minimality. Mayrhofer, Hagmayer, and Waldmann (2010) told participants a cover story involving four mind-reading aliens. The first alien (the cause alien) could send his thoughts to the three other effect aliens, causing them to think of food. Mayrhofer et al. asked participants to judge the probability that the first effect alien was thinking of food, given that the cause alien was thinking of food, and zero, one, or two other effect aliens were thinking of food. According to minimality, judgments should be independent of the number of other effect aliens that were thinking of food. Given that the cause is present, all direct effects in a CGM are independent. Participants, however, violated this prescription, generating a nonindependence effect: they were systematically more likely to predict that the first alien was thinking of food, as the experimenters increased the number of other effect aliens thinking of food.

Crucially, Mayrhofer et al. added a manipulation: they used two different cover stories in different conditions of the experiment. In the sending condition, they described the cause alien as sending his thoughts and suggested that he sometimes had trouble concentrating. In the reading condition, they described the effect aliens as reading the thoughts of the cause alien and suggested that the effect aliens sometimes had trouble concentrating. They found a significantly stronger nonindependence effect in the sending than in the reading condition.

Our interpretation of this difference is that the cover story Mayrhofer et al. used allowed participants to reason as if inhibitors were located at one of two different locations in their representations of the causal system.
In the sending condition, a single inhibitor was placed close to the common cause and was thus shared among the causal relations according to edge replacement. In the reading condition, participants represented individual inhibitors that influenced each effect node separately. While this experiment is an indirect and post hoc test of edge replacement, this finding does suggest further investigation. Below we describe an attempt to test edge replacement more directly by looking for a stream location effect in children.

Before we do, we wish to highlight two key differences between the prescriptions of minimality and edge replacement. The first is determinism: in the case of a novel causal relation that sometimes fails, minimality prescribes that children should expect an intrinsically stochastic relation, whereas edge replacement prescribes that children should posit the existence of a hidden inhibitor that has existence in space and time.

The second is branching: minimality prescribes that children represent novel relations as independent by default. In contrast, edge replacement prescribes that in an initial representation of multiple causal relations, the relations share some of the path from cause to effect. With determinism, branching gives rise to the stream location effects just described.

### 3. EVIDENCE FOR DETERMINISM

The most straightforward difference between edge replacement and minimality concerns what we have termed determinism: in novel relations that sometimes fail, minimality prescribes that initial representations posit a probabilistic relation. By contrast, edge replacement prescribes that in these cases, initial representations be of a hidden inhibitor that exists in space and time. Note that this theoretical stance of determinism is not new. Multiple researchers (e.g., Goldvarg & Johnson-Laird, 2001; Schulz & Sommerville, 2006) have advocated for similar versions of psychological determinism. Edge replacement formally instantiates determinism in a way that allows us to directly contrast it with minimality.

Note also that we are only endorsing one meaning of the word “determinism.” “Determinism” is sometimes used to refer to the principle of sufficient reason, where every event must have a cause (e.g., Bullock, Gelman, & Baillargeon, 1982). We are agnostic on this issue. Another common usage is that a cause must always produce its effects. Edge replacement makes a similar, but slightly more precise claim, which is that for all direct causal relations, the cause always produces the effect. Treating edge replacement as a psychological model means that when children represent causal
relations that sometimes fail, they posit at least one intermediate step between cause and outcome.

This claim is grounded in work by Schulz and Sommerville (2006). They introduced 4- and 5-year-olds to a machine that lights up and plays music when a switch is pressed. The machine also had a ring on the top surface. Children saw three trials in which pressing the switch caused the machine to activate. They then saw three trails in which the ring was removed; in these trials, the machine did not activate when the switch was pressed. This gave strong evidence that removing the ring was an inhibitor. At this point, the experimenter and confederate changed places, and the confederate pressed the switch eight times. This resulted in a sporadic pattern: the machine only activated in two nonconsecutive trials. The experimenter then revealed that she had been hiding a flashlight in her hand the whole time. She placed the flashlight next to the machine (the ring remained on the machine) and announced that she was about to press the switch. Children were asked to inhibit the activation of the machine. (“Can you make it so that the switch won’t work and the toy won’t turn on?”)

Note that at this stage children have two choices: they can choose to remove the ring, which they have previously seen to inhibit the causal relation. This is the reliable, conservative action. They can also choose to intervene on the flashlight. Intervening on the flashlight indicates that children inferred that the flashlight was inhibiting the machine during the sporadic period; the ring was visible and present on top of the machine.

Results indicate that most children chose the flashlight. This was in contrast to a condition in which the machine activated eight of eight times for the confederate; children in this condition tended to choose the ring. They also chose the ring when the sporadic activation was explained away by showing that the confederate was not pressing down hard enough on the switch. These data are consistent with edge replacement’s predictions about hidden inhibitors: during the sporadic period, children should represent a hidden inhibitor, which they later learn is the flashlight. Minimality’s predictions are less clear because it is not clear whether the flashlight should be counted as an observed variable.

3.1. Hidden Inhibitors and Time

Edge replacement also makes predictions about the temporal aspect of a hidden inhibitor, which are inconsistent with minimality. If a causal relation is probabilistic, we should rationally expect successes and failures of causal
efficacy to be independent in time, similar to coin tosses. If heads comes up on a fair coin the 10th time it is flipped, it does not mean heads is more or less likely on the 11th flip. Instead, the best predictor of the probability of heads is the proportion of heads that have come up before.

In the case of a real, instantiated inhibitor, we might expect that the status of the inhibitor depends on its previous states. In this way, successes and failures might not be more dependent in time, like the weather. If it rained yesterday, it is more likely to rain today than if it was sunny yesterday. The previous day’s weather is a better predictor of today’s weather than the proportion of days in which it has rained in the past. As adults, we know that the coin is random but the weather is due to a (set of) hidden inhibitors. The question at hand is which is a better description of the default form of basic causal representations, as can be measured through children’s judgments on a novel relation.

Buchanan and Sobel (2011b) showed children a scenario designed to put minimality and edge replacement at odds with each other. We introduced 3- and 4-year-olds to a set of lights that activated when the experimenter pressed their top. Two cause lights (A and B) were connected to an effect light (C) via wires visible to the child. In the solid condition, we pressed the A light, which activated the C light three times, then showed that B activated C three times, and then that A activated C again three times. The experimenter then lifted up the C light, told the child that he was “fiddling” with it, and then placed it back on in the system. The experimenter then pressed the A light once, which failed to activate the C light. The question of interest was whether children would think B activated C?

Contrast this case to one in which during the first part of the procedure, A and B did not always activate C, but did so two of three times (the sporadic condition). In this condition, the failures occurred at places in the sequence that was designed to appear random, but we ensured that the last trial was a failure, as in the solid condition. Thus, in the solid condition, the system succeeded 90% of the time, but failed on the most recent trial. In the sporadic condition, the system succeeded 60% of the time and had also just failed.

Edge replacement makes a strong, falsifiable prediction here: children should be more likely to predict success in the sporadic than in the solid condition. This is because in the solid condition, they can infer from the pattern of success and failure that there is a hidden inhibitor (evidenced by the one failure) that rarely changes state (evidenced by the string of successes). In the sporadic condition, the hidden inhibitor often changes state, evidenced by the changing pattern of success. In both conditions, children can infer that
the inhibitor was active on the previous trial, but in the solid condition, its active state is more likely to persist. Thus, they should be more likely to predict failure in the solid condition.

Minimality makes different predictions. We can deploy minimality in one of two ways in the scenario. The first is to predict that the probability of success on the next trial is proportional to some function of the previous pattern of successes and failures. Under this proportional approach, minimality must predict that children will be more likely to predict success in the solid than in the sporadic condition because there were more successes in the solid condition. Even if we weigh trials based on recency, minimality must make the same prediction because every failure in the solid condition is also a failure in the sporadic condition. A second way of deploying minimality is to take an agnostic approach: say that we know nothing about the power of the relation and predict random guessing. In either case, minimality makes a different prediction than edge replacement: children should be at least as likely to predict success in the solid as in the sporadic condition.

The data showed that preschool-aged children were significantly more likely to predict success in the sporadic than in the solid condition. That said, only 4-year-olds showed a pattern of performance that was different from chance; 3-year olds’ performance was indistinguishable from chance. We will return to this age-related difference in our discussion at the end of this chapter, but in general, this pattern of performance is more consistent with edge replacement than minimality.

3.2. Variability and Complexity

In the case of a causal relation that sometimes fails, edge replacement predicts that people will represent a hidden inhibitor that persists in time. More generally, edge replacement predicts that a more variable causal relation implies the existence of a causal mechanism that is more physically complex. In many ways, asking children to make judgments about success and failure on a particular trial (as we examined earlier) is an indirect test of this commitment. We can directly test edge replacement by asking children about possible mechanisms that underlie causal systems.

Erb, Buchanan, and Sobel (2013) showed 3- and 4-year-olds two novel toys. Both toys illuminated when pressed. One toy, the variable toy, illuminated in a series of flashing colors that changed over time. The other toy, the solid toy, illuminated in only one color. Children were then shown two
pictures of the insides of the toys (see Fig. 7.5). One picture showed a simple inside containing a battery and a light, whereas the other picture showed the battery and the light and additionally three novel objects. Children were asked to map each inside to a light.

Edge replacement predicts that children should map the more complex inside to the more variable causal relation and the simple inside to the less variable causal relation. This is because more hidden causes are needed to explain the variable effect, and these hidden causes must exist in space and time. Minimality instead requires the simplest possible relation: a single edge in both cases that does not distinguish between the two insides. We will treat minimality as predicting chance performance.

Four-year-olds’ performance was significantly more likely to map the complex inside to the variable pattern of activation. Three-year-olds’ performance was indistinguishable from chance. A series of controls and tests were conducted to support the result that the inference was indeed causal and not just a perceptual mapping. For instance, in one follow-up experiment, children never observed the activation of the light; instead, the experimenter described the activation pattern verbally. The same results were observed: 4-year-olds were significantly more likely to make the predicted mapping than chance would predict, while 3-year-olds’ ability to make the predicted mapping was not significantly different from chance.

In these experiments, young children showed that they can infer a more complex inside from a more variable causal relation. This phenomenon is consistent with edge replacement, in that children seem to believe that a more variable effect requires a more complicated internal structure—that is, more edges and actual, physically instantiated nodes between the observed cause and effect(s).
4. EVIDENCE FOR BRANCHING: STREAM LOCATION IN PRESCHOOLERS

We have so far examined data about the most basic contrasting prediction of minimality and edge replacement, that of determinism. We have seen that in multiple experiments, children appear to posit the existence of hidden inhibitors to explain relations that do not always hold, and they posit the existence of hidden causes to explain variable relations. We have yet to test a more subtle aspect of edge replacement: its organization of causal models into branching streams. This section will describe a series of experiments designed to test branching. We will do this through the stream location prediction described earlier. Stream location relies on determinism, but adds the implications of branching, bringing together multiple aspects of the theory.

To illustrate the ground for our experiments, consider again the cellphone example from the previous section. Figure 7.3 depicts two distinct cases. In both, Alice and Bob can call Cindy’s phone and cause it to ring, but each relation sometimes fails. In the early scenario, we imagined that Alice and Bob both called Cindy and failed. Then Alice changed the battery in her phone. Then Alice tried to call Cindy and succeeded. In the late scenario, we similarly imagined that Alice and Bob both called Cindy and failed. But in this scenario, it was Cindy who then changed the battery on her phone. After this change, Alice tried to call Cindy and succeeded. In each scenario, at issue is the probability that Bob will succeed in reaching Cindy on his next attempt at calling her. Edge replacement predicts that Bob will have a higher probability of success in the late than in the early scenario.

Edge replacement’s two key commitments work together to make this prediction: determinism requires that there be a physical reality to the inhibitor that caused the relation to fail. Branching makes it likely that the late inhibitor (Fig. 7.3c) will affect both relations, and likely that the early inhibitor (Fig. 7.3d) will affect just one relation. Recall that minimality makes no such prediction because we have no data about the effect of the intervention on the second independent relation with Alice’s phone. The minimal graph is shown in Fig. 7.4.

We conducted three experiments designed to test for a stream location effect in children. We created a scenario analogous to the previous example using toys designed to be novel to the children. Over the course of three experiments, we increased the novelty of the casual system.
4.1. Experiment 1

After familiarization with a novel causal system, we showed children a causal structure in which two causes were connected to a single effect. Critically, both causal relations initially failed: neither cause produced the effect. We then made a change (specifically, in this experiment, adding batteries) either close to one of the causes (the early condition) or close to one of the effects (the late condition). The change enabled one of the causal relations—children saw that one of the causes was now effective in producing the effect. We then asked children to predict the efficacy of the other relation. Our hypothesis was that experimentally manipulating the location of the intervention would affect responses—children who observed a change close to the effect would be more likely to generalize than children who observed a change close to one of the individual causes.

4.1.1 Methods

4.1.1.1 Participants and Design
Thirty-two preschoolers (22 boys, 10 girls, $M = 46.32$ months, range = 36–59 months) were recruited from birth records, a local preschool, and a children’s museum. Three additional children were tested, but were excluded due to experimenter error or equipment failure. Children were randomly assigned to either the early ($n = 16$) or late ($n = 16$) condition (between-subject). There were an equal number of 3- and 4-year-olds in each condition. The racial and ethnic breakdown of the sample was as follows: 29 children were Caucasian, 2 were African-American, and 1 child was Hispanic/Latino. No information about socioeconomic status was collected, but most children were from middle- to upper-middle class backgrounds.

4.1.1.2 Materials
We used two sets of commercially available closet lights, modified for the experiment. One set (the cause lights) consisted of eight lights, each 10 cm in diameter, with a button that could illuminate when pressed (see Fig. 7.6). The underside of each light consisted of a battery compartment enclosed by a removable cover that could be left on or off (leaving the batteries visible). The casing of each light was painted a different color.

Four lights made up the other set (the effect lights). These were 14 cm in diameter. One was modified so that it was activated only via a remote control. The experimenter (a trained magician) practiced an effect in which the light appeared to activate when depressed, even though the remote was the
actual cause of the activation. In pilot testing, this effect was convincing to both children and adults who often reported surprise when the remote was revealed. The remote allowed the experimenter to control which actions, if any, would appear to cause the light to activate. Each large light had a pipe cleaner around its casing so that children could tell them apart, but the apparent “identity” of the light could be switched surreptitiously. All four lights were shown together at the outset of the experiment, but the one light that was activated with the remote was the only one actually used in the test phase. Because of the deception, children believed that they were being shown the four lights in succession.

4.1.1.3 Procedure
The experimenter showed the lights to the children, arranged as in Fig. 7.6. There were two small cause lights, apparently connected to one large effect light. The experimenter familiarized the children with the causal system, by saying: “I have some of these lights. When you push on them, they light up. See [pushes on effect light, and it illuminates]. Here, you try.” All children pushed the effect light, which illuminated when depressed. “Sometimes, when I push on the little lights, they make the big light go. Watch.” He then pushed each cause light and each appeared to cause the large effect light to illuminate. There was no delay between the activation of cause and effect lights. The experimenter then pointed to each of the cause lights and asked,
“Does this one make the big light go?” Most children (26 of 32) correctly answered “yes” to these questions. The remaining children responded correctly after one instance of corrective feedback. Excluding children who required feedback on this or any other training question did not change the statistical significance of the results we report.

The experimenter then removed the three initial lights and arranged three different lights in the same configuration. This began the first of three test trials. In the late condition, the effect light in each test trial was missing a battery. In the early condition, one of the cause lights in each test trial was missing a battery. Lights without batteries did not illuminate when pressed. We left the covers off so that children could see the absence of batteries when the lights were flipped over.

For each test trial, the experimenter began by pushing the cause lights, demonstrating that they failed to activate the effect light. The experimenter asked the child to verify whether each cause light made the effect light activate. Most children (26 of 32) correctly answered “no” to these questions on all three trials. Five then responded correctly after one round of feedback, and one child required two rounds. Excluding these children does not change the significance of reported results. Note that at this point in the procedure, children had correctly answered “no” to two questions with feedback, and “yes” to two questions with feedback. Thus, children were not coached on a strategy that would allow them to answer the test questions correctly.

The experimenter then modified the causal system, depending on the condition. In the early condition, he flipped over one of the cause lights and said, “Look, this light has room for a battery, but there’s no battery there. Let’s put a battery in.” He inserted a battery and flipped the cause light back over. The battery was actually inserted backward so that the light would not begin illuminating when pressed; this would have allowed children to deduce that the other cause light, which did not illuminate when pressed, would be ineffective. In the late condition, the experimenter inserted a battery into the effect light (as opposed to one of the cause lights). No child spontaneously indicated that they had noticed that the battery had been inserted backward in either condition.

The experimenter then said, “Now let’s see what they do.” He pressed one of the cause lights (in the early condition, always the one he had changed) and it caused the effect light to activate. He then asked, pointing to the light he had just pressed: “Does this one make the big light go now?” All children correctly answered “yes” to this control question. They were then
asked the test question about the other light: “What about this one? Will this one make the big light go now?”

Responses to this test question were recorded and analyzed by an undergraduate assistant blind to the experimental hypotheses. A second undergraduate assistant coded a random sample of 25% of the data; interrater agreement was 100%. After the test question, the experimenter moved to the next test phase without giving feedback or showing the actual efficacy of the lights: he removed all three lights, brought out three different lights, and repeated the procedure.

4.1.2 Results

Responses to the test questions are shown in Table 7.1. For preliminary and chance analyses, we considered the proportion of children that gave correct answers to all three test questions. This was defined as three “yes” answers in the late condition or three “no” answers in the early condition. Twenty-nine of the 32 children made this kind of response. There was no significant difference in this distribution between genders, Fisher’s exact test, \( p = 1.00 \); between age groups, Fisher’s Exact test, \( p = 1.00 \); or between conditions, Fisher’s exact test, \( p = 0.23 \). We compared this pattern of responding to the proportion of all correct answers that would be expected if children guessed on each trial (0.125), if they randomly chose an initial response and then perseverated (0.50), or if they simply said “yes” to all questions (0.50)—the data were significantly different in each case, Binomial tests, all \( p \)-values <0.01.

We then compared the number of “yes” responses between the two conditions. The overall distribution of “yes” responses differed between the

<table>
<thead>
<tr>
<th>Number of “Yes” Responses on Test Questions</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experiment 1 (familiar mechanism)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early ( n = 16 )</td>
<td>13</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Late ( n = 16 )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td><strong>Experiment 2 (unfamiliar mechanism)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early ( n = 16 )</td>
<td>11</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Late ( n = 16 )</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>14</td>
</tr>
</tbody>
</table>
conditions, \( \chi^2(3, N=32) = 32.00, \quad p < 0.01, \quad \phi = 0.71 \). Because of low expected counts, we supplemented this analysis by collapsing children into two categories: those who answered “yes” to every test question and those who showed any other pattern of responding. Children in the late condition generated this pattern of response 100% of the time, while children in the early condition generated this pattern of response 0% of the time, Fisher’s exact test, \( p < 0.01 \). A similar result was found if we considered the proportion of children who said “no” to every test question as opposed to those who showed any other pattern of responding. Children in the early condition generated this pattern of response 81% of the time, while children in the late condition generated this response 0% of the time, Fisher’s exact test, \( p < 0.01 \).

4.1.3 Discussion

We essentially observed floor and ceiling results: the majority of children in the early condition said “no” to all test questions; the majority of children in the late condition said “yes” to all the test questions. Thus, Experiment 1 confirmed that preschoolers were not simply reasoning according to a minimal model when asked to reason about a novel causal system with a familiar mechanism (i.e., batteries).

Of course, at this age, children are quite familiar with batteries (e.g., Buchanan & Sobel, 2011a; Gottfried & Gelman, 2005; Sobel, 2009), which might have motivated their responses. Children might simply recognize that putting batteries into ineffective toys makes them activate. To address this concern, in Experiment 2, we examined whether children would make similar inferences given the same scenario, but with an action on an unfamiliar mechanism. Instead of adding batteries, all the lights had their full complement of batteries, but some did not have a battery cover. Adding a cover to the lights apparently enabled their causal efficacy. At issue is whether we replicate the findings from Experiment 1 (as edge replacement predicts) or more consistent with minimality framework, the pattern of performance becomes less akin to floor and ceiling levels.

4.2. Experiment 2

4.2.1 Method

4.2.1.1 Participants and Design

We recruited 32 preschoolers (8 girls, 24 boys, \( M = 46.44 \) months, range = 36–57 months) from birth records, a local preschool, and a local children’s museum. One additional child was tested, but was excluded.
due to experimenter error. Children were randomly assigned to either the early ($n=16$) or late ($n=16$) condition, with an equal number of 3- and 4-year-olds in each condition. The racial and ethnic breakdown of the sample was as follows: 28 children were Caucasian, 1 was African-American, 1 was Asian, 2 were Hispanic/Latino, and 2 were of mixed race. No information about socioeconomic status was collected.

### 4.2.1.2 Materials and Procedure

The materials and procedure were the same as in Experiment 1 except for two changes. The first was that all of the lights had each battery slot filled (the battery covers were still initially left off). Second, the act that apparently changed the efficacy of the relation was not adding batteries, but adding a cover to the opening that housed the batteries. The experimenter said, “Look, this one is missing a cover. Let’s put a cover on here.” Otherwise, the procedure was the same as in Experiment 1: children were initially shown a set in which two cause lights were effective, and were asked to verify the efficacy of the lights. Most children (30 of 32) answered correctly; the remaining two children answered correctly after one round of feedback. They were then shown three test trials. In each test trial, both cause lights were ineffective, and children were asked to verify this. Most children (22 of 32) correctly answered “no” to both of these questions on all three trials. Six children required one round of feedback, two children required two rounds of feedback, and two children required three rounds or more. Excluding all children who required any feedback does not change the statistical significance of the results reported below.

The experimenter then showed that adding a cover (either to the cause or effect light, depending on the condition) appeared to enable one relation, and children were asked to verify that the light was now effective (no children failed to answer this question correctly). Children were then asked about the efficacy of the other relation. Responses to this test question were recorded and analyzed by an undergraduate research assistant who was blind to the hypothesis. A second rater coded a random sample of 25% of the data; interrater agreement was 96%, with disagreements resolved by the first author.

Finally, to ensure that we did not negatively affect children’s causal knowledge, all children were briefed on the deception involved in the experiment: they were shown the remote and allowed to play with it at the end of the procedure.
4.2.2 Results
Table 7.1 displays children’s responses to the test questions, which were analyzed in the same manner as in Experiment 1. Twenty-five of 32 children answered as predicted to all three test questions. There was no significant difference in this distribution between genders, Fisher’s exact test, \( p = 0.63 \); between age groups, Fisher’s exact test, \( p = 1.00 \); or between conditions, Fisher’s exact test, \( p = 0.39 \). We compared this pattern of responding to the proportion of all predicted answers that would be expected if children guessed on each trial (0.125) or if they randomly chose an initial response and then perseverated (0.50), or if they showed a yes bias (0.50)—the data were significantly different in each case, Binomial tests, all \( p \)-values <0.01.

We then compared the number of “yes” responses between the two conditions. The distribution of the number of “yes” responses differed between the early and late conditions, \( \chi^2(3, \ N=32) = 25.67, \ p < 0.01, \ \phi = 0.51 \). Because of low expected counts, we supplemented this analysis by collapsing the data as in Experiment 1—comparing the number of children in each condition who said “yes” on all three test questions. In the early condition, children generated this pattern of results 0% of the time; in the late condition, children generated this pattern 88% of the time, a significant difference, Fisher’s exact test, \( p < 0.01 \). A similar finding was obtained when we analyzed whether children said “no” to every test question (69% in the early condition vs. 6% in the late condition), Fisher’s exact test, \( p = 0.01 \).

4.2.3 Discussion
As in Experiment 1, children were more likely to generalize efficacy from one light to the other in the late condition than in the change early condition, as predicted by edge replacement. These data are potentially problematic for minimality because it is not clear how children would have learned to expect these different actions to have these particular different effects.

But, it is possible that children are still responding based on their knowledge of batteries (or electrical systems more generally). Adding a cover might not be a completely novel action for young children; adults often remove the cover from a toy, exchange the batteries for fresh ones, and then replace the cover, making these two actions highly correlated and potentially (to the child) diagnostic of one another. Thus, it is difficult to rule out the possibility that children in Experiment 2 might have developed a mistaken association between changing the cover and changing efficacy, making these data consistent with minimality. To address this
issue, Experiment 3 was designed to ask preschoolers to reason about an action that is in conflict with this mechanism knowledge: removing batteries enabled the causal relation.

4.3. Experiment 3
4.3.1 Methods
4.3.1.1 Participants
We recruited 32 preschoolers (17 girls, 15 boys, $M = 48.00$ months, range = 36–60 months) from birth records and a local children’s museum. Three additional children were tested, but were excluded because they failed the control trials. The racial and ethnic breakdown was as follows: 25 children were white, 2 were African-American, 2 were native Hawaiian or other Pacific Islander, and 2 were of mixed or other races. No information about socioeconomic status was collected.

4.3.1.2 Materials
We modified the lights for this experiment, such that they each had a false bottom. The visible compartment on each light could contain batteries, but their presence was unrelated to the activation of the light itself, which was powered by batteries deeper inside (unbeknownst to the child). This made the lights about 2 cm taller, but they were otherwise similar to those in Experiments 1 and 2. We used two effect lights and four cause lights. The effect lights were individuated by a pipe cleaner placed around their body. Each cause light was painted a different color.

The false bottoms allowed us to address a potential concern with Experiments 1 and 2. In these experiments, cause lights illuminated in the late but not in the early condition. Although we know of no reason why this would bias our results, we welcomed the opportunity to make the two conditions as similar as possible. In both conditions of Experiment 3, each light always lit up when pressed, regardless of whether it had batteries in the visible compartment.

4.3.1.3 Procedure
Because of the robust effects in previous experiments, children were given a single test trial. During the familiarization phase, children were shown two cause lights that activated an effect light, and they were asked training questions (i.e., whether each cause light would activate the effect light) and
provided with feedback. Most of the children (28 of 32) required no feedback here; the remaining four children required one round of feedback before they answered correctly. All reported results remain statistically significant when these four children are excluded.

For the test trial, the experimenter showed children two new cause lights that failed to activate a new effect light and asked children to verify their efficacy. Most children (27 of 32) correctly answered “no”; four children required one round of feedback before answering “no,” while one child required two rounds. Two children (both 3-year-olds) were excluded due to a failure to generate the correct response within four trials of feedback.

Then the experimenter said, “Let’s look underneath this light. Look, this light has batteries. Let’s take the batteries out.” For the half of the children randomly assigned to the early condition, he removed the batteries of one of the cause lights, showed the empty compartment to the child, and replaced the light to its original location. For the remaining children in the late condition, he performed the same action to the effect light. He then demonstrated that this enabled one of the relations—that specific cause light (in the early condition) or one of the cause lights (in the late condition) now made the effect light activate. Children were asked to verify this change in efficacy. One child’s data were replaced because she failed to answer this correctly (no feedback was provided).

Children were asked about the efficacy of the other causal relation, and their answer was recorded and coded by an undergraduate assistant who was blind to the hypothesis. A second rater coded a random sample of 25% of the data; inter-rater reliability was 100%. At the end of the procedure, all children were briefed on the deception: they were shown that the lights had more batteries deeper inside and were allowed to play with the remote.

4.3.2 Results

Responses to the question about whether the other light would be effective are shown in Table 7.2. We found no effects of gender or age group on the proportion of children responding “yes” to this question, Fisher’s exact tests, all $p$-values $>0.32$. One of 16 children (6%) said this light would be effective in the early condition, compared with 12 of 16 children (75%) who made this response in the late condition, Fisher’s exact test, $p < 0.01$. Responses in both conditions differed from chance levels (50%), Binomial tests, both $p$-values $<0.01$. 
We repeated these analyses on the first trials of Experiments 1 and 2. Results from the first trial of Experiments 1 and 2 are also shown in Table 7.2. In all three experiments, significantly more children answered “no” in the early condition and “yes” in the late condition on the first trial than would be predicted by chance, binomial tests, all \( p \)-values < 0.01. In both Experiments 1 and 2, children were significantly more likely to answer “no” on the first trial in the early than in the late condition, Fisher’s exact tests, all \( p \)-values < 0.01.

4.3.3 Discussion

Children were significantly more likely to predict that a change to the effect light would enable both relations while a change to the cause light would only enable that relation. Critically, they generated this response even when the efficacy of this change conflicted with their mechanism knowledge. There was no difference between performance here and in comparable conditions in Experiments 1 and 2.

These results make the minimal account of the data particularly difficult to sustain, as the minimal account relies on children using their experience to construct a model that allows for stream location effects. However, children could not have had experience with this type of causal system because it operates in a way that directly contradicts any experience they might have previously had.

Across three experiments, we found a robust stream location effect in 3- and 4-year-olds’ causal inferences. The effect persisted across increasingly extreme attempts to make the causal system unfamiliar to the children: from toys that were enabled by adding a battery, to toys that were enabled by adding a battery cover, to toys that were enabled by removing a battery. Edge replacement not only accounts for but also predicts these effects, whereas minimality cannot provide an account of these data.
5. GENERAL DISCUSSION

An advantage of computational modeling in cognitive science is that it allows theories to be specified in an operationalized way. This advantage compels theories to make specific, concrete predictions that can also be tested with behavioral data. In rare cases where two competing theories can be operationalized within the same framework, this allows for instructive comparisons. Edge replacement and minimality constitute such a rare case: with respect to children’s most basic representations of novel causal systems, each theory is committed to opposing, specific predictions for a set of phenomena. We have presented an overview of these phenomena, along with experimental results designed to test these predictions. All of the results lend more support for edge replacement than minimality as describing the basic underlying principle around which children build their representations of causal structure.

Before we consider the relevance of these findings for minimality or edge replacement as the principle by which children construct CGM representations of their causal knowledge, we want to address a particular aspect of the data we have reviewed so far: some experiments that we reviewed have found differences between 3- and 4-year-olds and some have not. We found that 4-year-olds showed a significant difference between the solid and sporadic conditions of the hidden inhibitor experiments, but 3-year-olds did not. Erb et al. (2013) also found that 4-year-olds could map a variable casual relation to a more complex inside, but 3-year-olds’ performance was indistinguishable from chance. We did find stream location effects for both 3- and 4-year-olds (i.e., no development).

A tempting conclusion to make from this pattern of data is that 3-year-olds have not yet developed the capacity to use edge replacement. It is not clear that this conclusion is warranted. One reason is that 3-year-olds do show strong responses in stream location, and stream location incorporates all the important elements of edge replacement. That is, determinism is required to infer that a hidden inhibitor caused the change, and branching is required to infer and use the implications of the change’s location. Furthermore, 3-year-olds might have responded at chance levels due to poorer memory capacities or an inability to appreciate other kinds of substantive causal knowledge required to make inferences in the variability/complexity and solid/sporadic procedures.

Our general conclusion is that edge replacement is a more coherent principle than minimality as the way in which children (and adults) construct
representations of causal knowledge. Edge replacement does seem more consistent in both qualitative and quantitative ways with the data we have presented (see also Buchanan et al., 2010).

That said, one could argue that minimality is a much simpler account of how causal representations are constructed and should therefore be preferred on just this basis. In response, we question whether minimality is actually simpler. This may seem counterintuitive given that minimality is in essence a preference for simplicity, but there are some reasons to think that it leads in fact to unnecessary complexity. For instance, minimality states a preference for simplicity, but does not specify exactly how the mind might construct the minimal representation. Just because the end result is simple, does not mean that the generative process is simple. For instance, writing a 500-word summary of a line of research can sometimes take more time and effort than a 5000-word summary. In an analogous way, given a large set of experiences and descriptions of a causal system, there is no guarantee that it will be simpler to construct a minimal representation than a nonminimal one. On the other hand, we have described an algorithm that instantiates the edge replacement process for constructing representations (Buchanan et al., 2010).

The second reason to think that minimality may not be simpler is that it hides its complexity in edges. Quick visual inspection of graphs, such as those shown in Fig. 7.2a and b might lead a reader to think that edge replacement is more complex. But one could argue that Fig. 7.2b is actually simpler because all of the edges are deterministic. Edge replacement instantiates the same causal-functional relations, but forces all complexity into the structure of the graph, where it can easily be seen. Minimality instead offloads most of the complexity into the functional forms of edges, which is not apparent to visual inspection.

A final consideration with respect to complexity: Most scientists are justifiably suspicious of complex theories because they often have more free parameters that allow it to fit a wider range of data sets. Edge replacement does not have parameters that give it this freedom (again, for full computational details, see Buchanan et al., 2010). Rather, edge replacement makes a set of highly specific and falsifiable predictions. The edge replacement algorithm might be more difficult to instantiate than minimality, but it is not inherently more complex.

Finally, there is a further question of whether these data need to be fit using a model at all. For each experiment that we present here, one could construct a qualitative theory that explains the data. We do not object to
such exercises, nor do we think that edge replacement is the only theory that could explain these data. However, edge replacement deserves unique credit for two reasons: (a) it is the model from which the predictions were generated; (b) it also gives us an operationalized, coherent account of these multiple experiments. Even more important, edge replacement makes further predictions that have yet to be tested. We outline some of these predictions.

The first experiments to consider are the most direct extensions of the experiments we present here. For instance, stream location has been shown only in one domain in children and only using a common effect structure. A natural extension is to attempt to replicate this effect in other causal systems, or using a common cause structure. Edge replacement predicts that stream location effects should be found in these cases as well. Similarly, the data from Buchanan and Sobel (2011b) on hidden inhibitors, and the data from Erb et al. (2013), should be replicated in other domains. Experiments with adults could be valuable as well because they can more readily provide finer-grained quantitative judgments. Another place for application of edge replacement is in the investigation of functional forms. Most causal relations follow an OR functional form: each cause is individually sufficient. But it is conceivable for instance that causation can follow an AND functional form: multiple variables must each be in a particular state in order to cause another event to occur. For instance, Lucas, Gopnik, and Griffiths (2010) show that children may be better at adapting to new functional forms than adults (see also Lucas & Griffiths, 2010). Edge replacement can generate any functional form, but some are particularly awkward to construct; it assigns these a low probability. Thus, it makes predictions about which functional forms people should choose to apply first and may be useful to researchers in this area.

Less direct extensions of edge replacement are also necessarily more speculative, but are worthy of discussion. Edge replacement makes predictions about learning not just in the context of one experiment, but over longer periods of time. It suggests a specific progression to the way that children learn about a given causal system: begin with a simple deterministic representation and then amend the representation by adding structure that represents hidden causes. Once these hidden causes are uncovered, elaborate the structure and seek out more hidden causes. Edge replacement makes specific, operationalized predictions about how this process should unfold. Note that this is a theory of learning, not a theory of development: we do not propose that the core ability to use edge replacement changes as
children grow older. Instead, we predict that the application of edge replacement will lead to representations of particular causal systems that change in particular ways as children gain more experience with those systems.

The data and models here are part of a larger, and older, debate and research program about human causal inference. They are especially relevant to the relatively unresolved question of causal mechanisms: to what extent do people necessarily represent events that occur between cause and effect? There exists a continuum of positions with respect to causal mechanisms: at one extreme, it might be that people always represent causal mechanisms at a level of detail that can be grounded in intuitions about force and mechanical causality (see for instance Shultz, 1982; Wolff, 2007). For instance in representing how cell phones work, we might always need to imagine something like bits or energy moving through wires or space (even if the specifics are incorrect). At the other extreme, it might be that we never represent intermediate events unless they are necessitated by the statistical properties of the relation we observe. Recall that this latter stance is carried, intentionally or not, by minimality. Edge replacement allows us to navigate a middle way, without being vague. It holds that human representations are recursive: for every relation, there is a mechanism, but the mechanism need only be filled in (generating new relations) as that level of detail is needed.

ACKNOWLEDGMENTS

This work was supported by NSF (0518161 and 1223777 to D. M. S.). We would like to thank all of the parents and children who participated in this research and the Providence Children’s Museum for allowing us to work with children on-site. We would also like to thank Karis Casagrande, Brianna Doherty, Katie Green, Caroline Kleeman, Karina Ikezoe, Deanna Simeone, Rachel Shelley-Abrahamson, and Kristen Swan for help with participant recruitment and data analysis.

REFERENCES


