

Causal stream location effects in preschoolers

David W. Buchanan (david_buchanan@brown.edu) and David M. Sobel (dave_sobel@brown.edu)

Department of Cognitive and Linguistic Sciences
Box 1978, Brown University, Providence, RI, 02912

Abstract

Based on the predictions of a computational model, we test preschoolers' ability to reason about *stream location effects*: reasoning that interventions that occur on a common part of a causal process should be more likely to affect multiple relations, than interventions which occur on independent parts of a causal process. In two experiments, we show that 3- and 4-year-olds both show a stream location effect. Children show this effect for both familiar and unfamiliar interventions.

Keywords: causal reasoning, cognitive development, models of causal reasoning.

Introduction

Even when we do not explicitly understand the details of how a causal system works, we often have strong intuitions about the effects of different interventions on that system. Consider a phenomenon that is a mystery to most adults – the way that a remote turns on a television. Even though most of us could not verbally describe the mechanism, we know that removing the batteries from the remote would make it fail. We also know that the television must be plugged in. We even know about correlations between relations: Say you have two remotes that both turn on your television – for instance, the one that came with your television, and a universal remote you bought to control all your devices. One day, both fail to turn on the television. You replace the batteries in remote A, and it now succeeds in turning on the television. You would not expect this intervention to change the efficacy of remote B – it would be odd if replacing the batteries in one remote made both effective. On the other hand, if you had noticed that the television was unplugged, and plugging it in had restored the efficacy of remote A, you would not be surprised if remote B started working as well. These inferences seem obvious, even to a person who knows nothing about how these devices operate. Interventions in one place in a causal system are expected to have wide-ranging effects, while interventions in other locations are not. Why is this?

To explain this intuition, we will use the metaphor of a *causal stream*. For instance, imagine we introduce pollution into a river that has several branches. The further upstream the pollution occurs, the more branches of the stream will be polluted. When we think of causation as flowing down a branching path, we can start to formalize these intuitions. Elsewhere (Buchanan, Tenenbaum, & Sobel, 2010) we have proposed a computational model that generates causal structures that have a branching, stream-like character. We call it the *causal edge replacement process*, or CERP. While the details are beyond the scope of this paper, we will outline its main implications.

CERP makes use of causal graphical models, a way of representing causal relations using graphs (Gopnik et al., 2004;

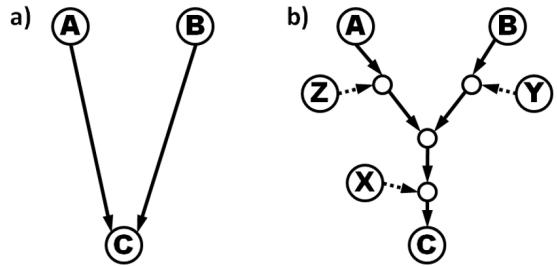


Figure 1: Examples of causal graphical models. Nodes represent events, and edges represent causal relations. Dashed edges indicate inhibitory relations. On the left, the simplest graph that captures a common effect relation: A and B both cause C. On the right, a graph generated by CERP, which allows us to make predictions about stream location effects. Intervening on X disables both relations, but intervening on Y or Z disables only one.

Pearl, 2000; Spirtes, Glymour, & Scheines, 2001). Nodes represent events, and directed edges represent causal relations. Figure 1a shows an example, the simplest way of representing the common effect relation of the television (C) and the two remotes, A and B. There are other ways of representing this relation; CERP tends to generate graphs that are more complex, like Figure 1b. In this graph, we have enough detail to represent interventions on the mechanism that relates cause and effect. For instance, X is such an intervention, which disables both relations, preventing causation from flowing down the edge on which it falls. (The dashed edge indicates an inhibitory relation.) CERP implies that when causal relations share a node, they always share part of the path from cause to effect. In a common cause relation such as the television example, interventions (like X) that occur late in the causal stream (close to the common effect) are more likely to fall on this shared path, changing both relations. Interventions (like Y and Z) that occur early in the causal stream, are more likely to fall on the independent path, changing only one relation. We call this difference a *stream location effect*. Note that these implications are general and structural, and do not depend on the specific causal system involved.

The stream location hypothesis is that human beings should expect stream location effects even about systems for which we have little or no knowledge of the causal mechanism involved (like the television remote, for most of us). While this hypothesis was inspired by CERP, it is not the only model which is consistent with these predictions. For instance, evidence that supports the stream location hypothesis

is not inconsistent with a general approach to causal graphical models. But only models (like CERP) that have a branching character directly and specifically predict stream location effects.

There is already some empirical evidence that suggests that stream location effects may exist in adults. Mayrhofer, Hagmayer, and Waldmann (2008) told participants a cover story involving mind-reading aliens: When the “cause” alien thought of food, he often caused the three “effect” aliens to think of food as well. The experimenters manipulated the number of other aliens that thought of food, and asked subjects to judge the probability that a given alien would also think of food, given that the cause alien was thinking of food. For instance, when the cause alien and two other effect aliens were thinking of food, participants judged it highly likely that the third alien was thinking of food. When the cause alien was thinking of food, but the two other effect aliens were not, subjects judged it less likely that the third effect alien was thinking of food.

This difference is known as a *nonindependence* effect, which CERP fits well in general¹: Adults predict that collateral effects of a common cause should be correlated, even given their common cause. Crucially for the stream location hypothesis, the strength of nonindependence could be manipulated by changing the cover story. In the “sending” condition, participants were told that the cause alien sometimes had trouble concentrating; there was a strong nonindependence effect in this condition. In the “receive” condition, the effect aliens sometimes had trouble concentrating; there was a significantly weaker nonindependence effect in this condition. Mayrhofer et al. succeeded in showing that by changing the description of the mechanism, they could change the degree of nonindependence observed. We hypothesize that a stream location effect was responsible for this difference: the location of the described inhibitor (trouble concentrating) in the causal stream was different between conditions. Of course, we are only explaining their data in hindsight. The experiments in this paper present a more direct predictive test of the stream location hypothesis.

Because CERP makes such strong predictions about situations in which we have little or no knowledge, the best tests of the stream location hypothesis will be in children’s causal reasoning. This is because children often have little specific causal knowledge about individual causal systems; we can see their reasoning as revealing the expected form of causation more than the expected content of causation. For instance, infants seem to initially expect that novel abstract objects need to make physical contact in order to interact causally (Leslie & Keeble, 1987). Among preschoolers, Bullock, Gelman, and Baillargeon (1982) found that even 3-year-olds expect that causes must precede their effects. They also found that 3-year-olds could reason appropriately about interventions on causal systems: They could recognize that some interventions would change a relation, whereas some

would not. Buchanan and Sobel (submitted) showed that this ability depended on the specific causal system involved. For instance, 3-year-olds could not reason correctly about interventions on electrical connection, but they could reason correctly about interventions on batteries. On the other hand, 4-year-olds could reason appropriately about both connection and batteries. Because of these developmental differences, and numerous other studies on preschoolers’ causal reasoning, we chose to test 3- and 4-year-olds in these experiments.

Our overarching hypothesis, which we test in two experiments, is that preschool-aged children will show stream location effects, expecting different changes to arise from interventions at different locations in the causal stream. Further, we predict that these differences will continue to hold regardless of the familiarity of the intervention involved, as long as that unfamiliar intervention appears to change the causal relation in the same way.

Experiment 1

In this experiment, we tested the hypothesis that children would reason differently about interventions to a causal system, depending on the location of the intervention in a causal stream. We presented children with a novel common effect relation in which both relations failed. Then we made a change, either early or late in the causal stream, which apparently enabled one of the relations. We asked children whether this change would enable the other relation as well. Our hypothesis was that in accordance with CERP, children would judge the late intervention as more likely to affect both relations than the early intervention.

Methods

Participants We tested 16 three-year-olds, (8 girls, mean age = 40.3 months, range = 36-45 months) and 16 four-year-olds (2 girls, mean age = 52.25 months, range = 48-59 months). Three additional children were tested, but were excluded due to experimenter error or equipment failure. About half the children were recruited from birth records, and the other half were recruited and tested either at a children’s museum or at a local preschool. Children were randomly assigned to either the “early inhibitor” ($n = 16$) or “late inhibitor” ($n = 16$) condition. There were an equal number of 3- and 4-year-olds in each condition.

Materials Materials consisted of two sets of commercially available closet lights, modified for the experiment. In one set (the “cause lights”) there were 8 lights, each 10 cm in diameter, with a large white button that illuminated only when actively depressed. These lights had a battery compartment on the underside that could hold two batteries; they required the presence of both batteries, inserted properly, in order to illuminate when pressed. It was possible to insert one battery backwards, in order to be able to show the presence of two batteries, without the light illuminating when pressed. The compartment also had a cover, which could be left on or off. The casing of each light was painted a different color, so that

¹For details of this fit, see Buchanan et al. (2010)

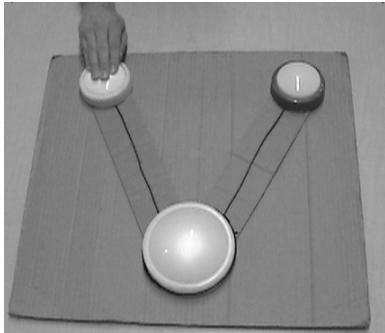


Figure 2: The push lights environment used in the present experiments, shown from the point of view of the child.

children could easily differentiate the lights.

Another set (the “effect” lights) consisted of four similar but larger lights, each about 14 cm in diameter. The effect lights were rendered distinguishable from one another by placing pipe cleaners of different colors around their casing. These lights each held four batteries. One of these lights was modified using radio-controlled car components, such that it illuminated only when a hidden remote was activated – depressing the light was not the actual cause of its illumination. It was possible to give adults and children the impression that they were causing the light to activate by depressing it, by activating the remote only when they pressed the light. The experimenter (a trained magician) practiced this effect until a convincing causal impression was achieved. The remote allowed the experimenter freedom to control which actions, if any, appeared to cause the effect light to illuminate. In post-tests, the experimenter was often able to use the remote to convince children that pressing their nose activated the effect light. All such children were subsequently debriefed, and allowed to play with the remote.

The lights were mounted on a piece of cardboard, together with wires that appeared to connect the lights. The cardboard was used so that the experimenter could easily retract the whole setup, controlling when and if children could intervene on the lights. The setup is shown in Figure 2. We refer to this setup as the “push lights environment.”

Procedure The experimenter began by showing children all the lights to be used in the procedure, in order to establish that there were a large number of them. Then he showed them the push lights environment, arranged as shown in Figure 2: There were two small “cause” lights, connected to one large “effect” light using the wires. This began the training phase. The experimenter said: “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.” Children pushed the effect light, which illuminated. “Sometimes, when I push on these little lights, they’ll make the big light go. Watch.” He then pushed each cause light, both of which appeared to cause the large light to illuminate simultaneously. He then pointed to each of the cause lights and asked: “Does this one

make the big light go?” Most children (26 out of 32) correctly answered “yes” to this question. The remaining children answered “yes” 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 visibly different lights in the same configuration. This began the first of three test phases. In the “late inhibitor” condition, the effect light in each test phase light was missing one battery. In the “early inhibitor” condition, the cause lights in each test phase were each missing one battery, and thus did not illuminate when pressed. The battery covers were left off so that the absence of batteries was visible, but only when the lights were flipped over. These new lights failed to activate the effect light. Children were asked about the efficacy of the relations. Most children (26 out of 32) correctly answered “no” to these questions on all three trials. Five responded correctly after one round of feedback, and one child required two rounds. Excluding these children did 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.

At this point in the test phase, the experimenter made a modification to the causal system, which depended on the condition. In the “late inhibitor” condition, the experimenter flipped over the large light, exposing the fact that there was a battery missing. He said: “Look, this light has room for a battery, but there’s no battery in there. Let’s put a battery in.” He then inserted a battery into the space, and replaced the light in its original position. In the “early inhibitor” condition, he instead flipped over and added a battery to one of the cause lights. Then he said “Let’s try this light now.” He pressed one cause light (side counterbalanced, and in the early inhibitor condition, always the effect light that had been intervened on), which made the effect light illuminate.² The experimenter then asked, pointing to this light: “Does this one make the big light go now?” All children answered “yes.” He then pointed to the other light: “What about this one? Will this one make the big light go now?” Children’s responses to this test question were recorded and analyzed. The experimenter repeated the test phase three times with three visibly different sets of lights, for a total of three answers from each participant. This meant that we collected three yes/no answers from each child, making 24 for each age group/ condition combination.

²In the “early inhibitor” condition, the cause light did not illuminate even when it was effective. Otherwise the illumination of the cause light would be diagnostic of its efficacy in causing the effect light to illuminate. That is, in the “early inhibitor” condition, when the experimenter pressed on an effective cause light, only the effect light illuminated.

Results

No effects were found for the age or gender of the children, or whether the question was initially asked about a light that was on the left or on the right. Results are shown in Table 1. In the “early inhibitor” condition, only 2 out of 24 responses from 3-year-olds, and 1 out of 24 responses from 4-year-olds were “yes.” Both of these patterns were significantly below the proportion of “yes” responses predicted by chance, Binomial test, $p < 0.01$ in both cases. In the “late inhibitor” condition, all 3- and 4-year olds answered “yes” to every question, meaning that both age groups answered “yes” to 24 out of 24 questions. This was significantly above chance, Binomial test, $p < 0.01$ for both conditions.

Table 1: Mean number of “yes” answers in Experiment 1.

Age Group	Condition	“yes”/trials	Mean	SD
3-year-olds	Early ($n = 8$)	2/24	0.25	0.46
	Late ($n = 8$)	24/24	3.00	0.00
4-year-olds	Early ($n = 8$)	1/24	0.12	0.35
	Late ($n = 8$)	24/24	3.00	0.00

Children of both ages were significantly more likely to answer “yes” in the late inhibitor than in the early inhibitor condition. For 3-year-olds, the average number of “yes” responses out of three was 0.25 in the early inhibitor condition and 3.00 in the late inhibitor condition. Among 4-year-olds, the means were 0.12 and 3.00, respectively. We ran a 2(age group) x 2(condition) ANOVA, which revealed a main effect of condition, $F = 746.05$, $p < 0.01$, partial $\eta^2 = 0.96$, but no main effect of age, $F = 0.37$, $p = 0.55$, partial $\eta^2 = 0.01$ or interaction, $F = 0.37$, $p = 0.55$, partial $\eta^2 = 0.01$. Because of the apparent difference in the variances, we supplemented this analysis using a Mann-Whitney U test: We found a significant difference between conditions, $U = 0.00$, $Z = 5.37$, $p < 0.01$, but not in the number of correct answers (“yes” in the late condition, and “no” in the early condition) between age groups, $U = 20.00$, $Z = 0.60$, $p = 0.78$.

Discussion

Children were significantly more likely to predict a change in both relations in the “early inhibitor” than in the “late inhibitor” condition. These results indicate that both 3- and 4-year-olds are sensitive to the location of an intervention in a causal stream. An open question is whether this is due primarily to structural inferences that apply to causal streams in general, or acquired knowledge about this specific intervention. Previous research (i.e. Buchanan & Sobel, submitted; Gottfried & Gelman, 2005) indicates that even 3-year-olds understand enough about batteries to make appropriate inferences about relevant and irrelevant modifications to a causal system when batteries are involved. To support stream location as a general structural principle, we needed to show a stream location effect for an unfamiliar intervention.

Experiment 2

The goal of this experiment was to show a stream location effect in a similar environment, but using an intervention that was not usually associated with a change in causal relations. In this experiment, instead of adding batteries, we added a battery cover. Since the presence of battery covers is not actually causally related to the efficacy of toys, children could not base their inferences on previous causal learning. We hypothesized that we would find the same effect in this experiment as in Experiment 1. We were agnostic as to whether children would be more uncertain in this experiment, generating a significantly more variable pattern of responses.

Participants As in Experiment 1, we tested 16 three-year-olds, (3 girls, mean age = 40.67 months, range= 36-46 months) and 16 four-year-olds (5 girls, mean age = 52.18 months, range = 48-57 months). One additional child was tested, but was excluded due to experimenter error. About half the children were recruited from birth records, and the other half were recruited at a children’s museum or local preschool. Again, children were randomly assigned to either the “early inhibitor” ($n = 16$) or “late inhibitor” ($n = 16$) condition, with an equal number of 3- and 4-year-olds in each condition.

Methods Experiment 2 was identical to Experiment 1, using the same materials and procedure, except for two changes: First, all battery slots were filled, but as in Experiment 1 the battery covers were left off initially. Second, during the procedure, the experimenter did not add batteries; instead, he pointed out that each light was missing a cover, and added one. Thus, he said: “Look, this one does not have a cover. Let’s put a cover on there.” Just as in Experiment 1, in the early inhibitor condition, the cause lights did not illuminate. This was done in order to maintain similarity between experiments. Also as in Experiment 1, intervening on the light (the cause light in the late inhibitor condition, and the effect light in the early inhibitor condition) apparently changed the efficacy of one of the cause lights. Children were asked to verify this. Most children (22 out of 32) required no corrective feedback during this procedure. Six children required one round of feedback, two children required two rounds of feedback, one child required three rounds, and another four. Excluding all children who required any feedback does not change the statistical significance of the results reported below. In the test question (for which no feedback was provided), children were asked to predict the efficacy of the other cause light. To avoid negative effects on children’s causal learning, all children were debriefed on the deception at the end of the procedure: They were allowed to play with the remote, and observed that replacing the covers did not in reality make the lights effective.

Results

Results are shown in Table 2. Again, no effects were found for gender, or the location of the intervened-on light. As in

Experiment 1, chance analyses showed that in all four condition/age group combinations, the proportion of “yes” responses was significantly different from what would be expected by chance, Binomial test, $p < 0.01$ in each case. In the “early inhibitor” condition, children were below chance, and in the “late inhibitor” condition, they were above chance.

Table 2: Number of “yes” answers in Experiment 2.

Age Group	Condition	“yes”/trials	Mean	SD
3-year-olds	Early ($n = 8$)	4/24	0.5/3	0.27
	Late ($n = 8$)	23/24	2.87/3	0.12
4-year-olds	Early ($n = 8$)	3/24	0.37/3	0.74
	Late ($n = 8$)	21/24	2.62/3	1.06

For 3-year-olds, the average number of “yes” responses out of three was 0.25 in the early inhibitor condition and 3.00 in the late inhibitor condition. Among 4-year-olds, the means were 0.12 and 3.00, respectively. As in Experiment 1, we ran a 2(age group) x 2(condition) ANOVA, which revealed a main effect of condition, $F = 72.05$, $p < 0.01$, partial $\eta^2 = 0.72$, but no main effect of age, $F = 0.47$, $p = 0.497$, partial $\eta^2 = 0.017$ or interaction, $F = 0.05$, $p = 0.82$, partial $\eta^2 = 0.002$. Because of the apparent difference in the variances, we supplemented this analysis using a Mann-Whitney U test: We found a significant difference between conditions, $U = 11.50$, $Z = 4.27$, $p < 0.01$, but not in the number of correct answers (“yes” in the late condition, and “no” in the early condition) between age groups, $U = 123.00$, $Z = 0.26$, $p = 0.87$.

Anecdotally, several 4-year-olds seemed surprised that merely changing the cover had changed the efficacy of the relation. Some initially responded “maybe” to the test question – they were asked to choose either a “yes” or “no” response. All children eventually responded appropriately to the test question.

Because of this phenomenon, we also tested for differences between the experiments: We gave each child a score based on the number of correct (“yes” in late inhibitor, and “no” in early inhibitor) responses they made. We then performed a t-test on the difference between scores in the two experiments. In Experiment 1, the mean score was 2.90, ($SD = 0.29$), and in Experiment 2, the mean score was 2.65, ($SD = 0.74$). This difference was only marginally statistically significant, $t = 1.76$, $df = 62$, $p = 0.08$. Because of the difference in the variances, we supplemented this analysis using a Mann-Whitney U test, which also failed to show a significant difference, $U = 443.50$, $Z = 1.48$, $p = 0.14$.

Discussion

Even in the case of an intervention that is not normally causally related to efficacy, 3- and 4-year-olds were able to reason appropriately about stream location. That is, when the unfamiliar intervention that resulted in a change in efficacy was early in the causal stream, children predicted that

the other causal relation would be unaffected, but when the unfamiliar intervention was late in the causal stream, they predicted that both relations would be affected. The data are inconclusive about whether there is an effect of familiarity, possibly making children’s responses more variable. Even if this effect exists, it is probably small, and manifestly not large enough to eliminate the stream location effect we observed.

General Discussion

Both experiments supported the stream location hypothesis: Children were significantly more likely to predict a change in both relations in the late intervention than in the early intervention condition. The results in both age groups indicate that children have a strong understanding of stream location, even as early as three years old. Furthermore, Experiment 2 showed that children would make these inferences even for an unfamiliar intervention. This suggests that stream location may reflect knowledge of the structure of causation in general, rather than just experience with a specific causal system. Further work is necessary to provide more support for this possibility. For instance, we may be able to find stream location effects when the intervention is not just unfamiliar but opposite to past associations – if batteries disable rather than enable the relation, for example.

CERP predicts and supports these findings. The model prescribes that early interventions on a common effect structure are likely to fall on the independent portion of the path from cause to effect, changing only one relation, whereas late interventions on a common effect structure are likely to fall on the shared portion of the path from both causes to the effect, changing both relations. While the data we present are consistent with the general causal graphical model framework – we have shown that children prefer Figure 1b over Figure 1a – only CERP explains *how* this preference is generated.

In Experiment 2, adding a cover appeared to change the efficacy of a the causal relation, a situation that would be counter to children’s experiences (if any) with such causal systems in the real world. Why, then, did they not show a significantly different pattern of responding in Experiment 2? For instance, we might expect children to guess. The answer comes from noticing that the intervention was perfectly correlated with a change in efficacy: the light never worked until we added a cover. It seems that children required an explanation for this change, and the addition of the cover was the only explanation available. This is in line with previous research (i.e. Schulz & Sommerville, 2006) that shows that children are determinists: they attribute such changes in efficacy to human interventions, rather than attributing them to randomness. In work currently underway, we are exploring the interaction between this type of determinism, and inferences about hidden interventions on a causal stream. CERP makes clear predictions here: for instance, if failures sometimes occur without an intervention, a given failure is less indicative of a changed causal relation. Thus, more variable relations should show weaker stream location effects.

The existence of stream location effects in preschoolers provides support for CERP as a model of causal reasoning. Although CERP arose from attempts to make quantitative fits to data on a different phenomenon with adults (namely, the nonindependence phenomenon mentioned above), it nonetheless predicted a novel, qualitative effect that could be detected in children. We see this as one of many examples (i.e. Sobel, Tenenbaum, & Gopnik, 2004; Thelen, Schoner, Scheier, & Smith, 2001) of a productive dialog between experiments and models in cognitive development and cognitive science.

Acknowledgments

This work was supported by NSF (DLS-0518161 to DMS). We would like to thank all of the parents and children who participated, and the Providence Children's Museum and Brown/Fox Point preschool for allowing us to recruit and test on-site. We also thank Karis Casagrande, Caroline Kleeman, Brianna Doherty, Katie Green, Rachel Shelley-Abrahamson, Kristen Swan, for help with data collection and analysis.

References

Buchanan, D., & Sobel, D. (submitted). Mechanism-based causal reasoning in young children.

Buchanan, D., Tenenbaum, J., & Sobel, D. (2010). Edge replacement and nonindependence in causation. In *Proceedings of the 32nd annual conference of the cognitive science society*.

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 209–254). Academic Press.

Gopnik, A., Glymour, C., Sobel, D., Schulz, L., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, 111, 3–32.

Gottfried, G., & Gelman, S. (2005). Developing domain-specific causal-explanatory frameworks: The role of insides and immanence. *Cognitive Development*, 20(1), 137–158.

Leslie, A., & Keeble, S. (1987). Do six-month-old infants perceive causality. *Cognition*, 25(3), 265–288.

Mayrhofer, R., Hagmayer, Y., & Waldmann, M. (2008). Violations of screening off: A Bayesian error attribution model of causal reasoning. *Unpublished presentation at Mathematical Psychology 2008*.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press.

Schulz, L., & Sommerville, J. (2006). God does not play dice: Causal determinism and preschoolers' causal inferences. *Child development*, 77(2), 427–442.

Sobel, D., Tenenbaum, J., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science*, 28(3), 303–333.

Spirites, P., Glymour, C., & Scheines, R. (2001). *Causation, prediction, and search*. The MIT Press.

Thelen, E., Schoner, G., Scheier, C., & Smith, L. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and brain sciences*, 24(01), 1–34.