Considering alternatives facilitates anomaly detection in preschoolers

Jae Engle (jengle@ucsd.edu)
Department of Psychology, 9500 Gilman Drive, UC San Diego
La Jolla, CA 92093 USA;
Department of Neuroscience, 10550 N. Torrey Pines Road, The Scripps Research Institute
La Jolla, CA 92037

Caren M. Walker (carenwalker@ucsd.edu)
Department of Psychology, 9500 Gilman Drive, UC San Diego
La Jolla, CA 92093 USA

Abstract

Here we explore whether drawing upon preschooler's intuitive causal reasoning abilities may bolster their attention to the presence of conflicting data. Specifically, we examine whether prompting children to think counterfactually about alternative outcomes facilitates their anomaly detection in a causal reasoning task. The current task assesses whether children in two conditions successfully differentiate between potential causes: one that accounts for 100% of the data (no anomalies), and one that accounts for 75% of the data (anomalies observed). Results indicate that counterfactual prompts lead 5-year-olds to privilege the hypothesis that accounts for more of their observations, and also support transfer of this hypothesis to inform their inferences about novel cases. Findings suggest that counterfactual scaffolds may be beneficial in promoting causal reasoning in children.

Keywords: cognitive development; causal learning; counterfactuals; scientific reasoning; anomaly detection

Detecting Anomalous Data

As learners, we frequently encounter evidence that is incompatible with our existing theories (Carey, 2009; Chinn & Brewer, 1998; Zimmerman, 2007). Although the presence of anomalous data plays a key role in overturning incorrect hypotheses, errors at the stages of observation, interpretation, generalization, or retention could prevent belief revision (Chinn & Brewer, 1998). In particular, a learner might ignore, reject, or exclude anomalous data to maintain an existing theory. They could also choose to keep the anomalous data in abeyance, or even reinterpret the conflicting data to force alignment.

Despite these challenges, decades of empirical work have demonstrated that even preschool-aged children intuitively draw causal inferences from observed patterns of co-variation (e.g., Gopnik, Sobel, & Schulz, 2001), selectively explore unexpected evidence (Legare, Gelman, & Wellman, 2010), and readily update prior hypotheses in light of new data (e.g., Gopnik & Wellman, 2012; Schulz, Bonawitz, & Griffiths, 2007). Given this early competence, why do much older children often struggle to revise existing hypotheses in light of anomalous evidence (e.g., Penner & Klahr, 1996)?

One well-established factor that makes it difficult to process anomalous data is the effect of biases that result from prior knowledge: both children and adults have difficulty setting aside their prior knowledge in light of contradictory evidence (Brewer and Chinn, 1991; Kuhn et al., 1989; Kunda, 1990). However, difficulties can also arise in cases of rapidly formed beliefs, even from relatively sparse data. For instance, Schulz et al., (2008) found that preschool-aged children infer abstract physical causal laws from minimal evidence, and maintain these newly inferred laws when confronted with anomalous observations. Schulz and colleagues note that this ability to learn robust abstract principles from sparse data is part of what makes learning so powerful and efficient, even in early childhood. However, this response to new data may also make detecting and responding to informative anomalies more difficult, leading to inflexibility when confronted with counterevidence (Koslowski, 1996; Kuhn, 1989; Kuhn, Amsel, & O'Laughlin, 1988; Schauble, 1990). That is, after a learner's initial exposure to even a small amount of data, all subsequent casual judgments reflect these data as prior knowledge, influencing their interpretation of additional observations. The salience of these newly formed prior beliefs can thus slow the process of hypothesis revision, leading initial beliefs to become rapidly entrenched.

Another factor that has been proposed to impede anomaly detection, which can operate simultaneously or independently of prior knowledge, is cognitive load. For example, Koerber, Osterhaus and Sodian (2017) suggest that children may fail to engage in belief revision because a task is too complex or insufficiently salient. To illustrate this point, they demonstrated that when data was presented in the context of bar graphs, second-graders were more likely to revise their initial beliefs in light of available counterevidence. They proposed, therefore, that providing external representations of the data might serve to decrease cognitive demands, making information more accessible.

Other types of scaffolding have instead focused on manipulating the learner's internal representations of the data (Walker, Lombrozo, Williams, Rafferty, & Gopnik, 2016; Williams & Lombrozo, 2010). For example, Walker and colleagues (2016) found that prompts to explain led preschoolers to process anomalous data more effectively, and successfully revise their beliefs in light of new evidence. In this study, children observed patterns of data suggesting two different causal properties: one that accounted for 100% of their observations and one that accounted for only 75%. Results demonstrated that the majority of 5-year-olds in the control condition ignored the presence of anomalous data associated with the 75% cause, thereby failing to differentiate between the two candidate causes when asked to generalize to a novel set. When prompted to explain however, children were significantly more likely to generalize according to the 100% cause. The authors argue that the process of generating an explanation leads even very young learners to privilege those hypotheses with greater "scope" (i.e., those that account for the greatest proportion of the data). Similar results have also been found in adult learners (e.g., Williams & Lombrozo, 2010).

Counterfactuals as scaffolds for anomaly detection

We hypothesize that because children's early learning mechanisms resemble the basic inductive process underlying scientific theory change, it may be possible to harness children's intuitive causal reasoning to bolster anomaly detection, facilitating belief revision in light of counter-evidence. In particular, we explore the use of explicit counterfactual prompts in highlighting the presence of anomalies in a novel data set.

The central distinguishing feature of causal knowledge is the fact that causal relations have the additional requirement of counterfactual dependence (i.e., the statement *X causes Y* implies the counterfactual that *a change to X would lead to a change in Y*) (Pearl, 2000; Woodward, 2003). Counterfactuals therefore act as input to causal judgments (Lewis, 1986; Mackie, 1974), and previous work suggests that encouraging children to think counterfactually leads them to engage in more sophisticated forms of causal inference (e.g., McCormack, Simms, McGourty, & Beckers, 2013).

In addition to prompting reflection about the potential outcomes of specific interventions, counterfactual questions may also enable consideration of multiple, alternative hypotheses in order to select the one that is most consistent with the data observed. While we do not usually engage in conscious tracking of these implied counterfactuals (Sloman, 2005), the process of reasoning about anomalies should theoretically invoke counterfactual reasoning. That is, if a learner holding theory *A* observes incompatible evidence *X*, they should infer that *if theory A were true, evidence X would not have occurred,* thereby promoting belief revision. Despite this link, previous research indicates that children often fail to spontaneously consider this pattern of causal contingency (Chinn & Malhotra, 2002; Chinn & Brewer, 1998).

One possible explanation for this failure may be that children have difficulty interpreting the relevance of the anomalous data they observe if they fail to generate high-quality alternative hypotheses (Chinn & Brewer, 1993). By explicitly prompting a learner to consider a counterfactual possibility, it may not only serve to highlight the availability of plausible alternatives, but also lead them to simulate whether the existing evidence would hold if some alternative hypothesis were true (Hirt & Markman, 1995).

While there has been prior empirical work examining the development of counterfactual thinking (Buchsbaum, Bridgers, Weisberg, & Gopnik, 2012; Harris, 1996; Riggs & Peterson, 2000), and the development of anomaly detection (Chinn & Malhotra, 2002; Chinn & Brewer, 1998; Zimmerman, 2007), as well as prior theoretical work drawing similarities between the two (Bauchsbaum et al., 2012; Gopnik &Walker, 2013; Walker & Gopnik, 2013a; Wenzlheumer, 2009), there are no empirical studies to date that directly examine the link between these abilities in children. In fact, to our knowledge, there is only one study that has examined the relationship between scientific reasoning and counterfactuals (Galinsky & Moskowitz, 2000). Results of this study indicate that adult participants who were primed to think about counterfactuals were more likely to consider alternative hypotheses, prompting them to seek hypothesis-disconfirming evidence.

The current study uses a version of the causal learning paradigm described above (Walker et al., 2016) to examine the role of counterfactuals (if any) in differentiating between two candidate causes: one that accounts for all of the observed data (i.e., no anomalies), and one that accounts for most, but not all, of the observed data (i.e., anomalies observed). Using this paradigm, we will assess whether strategically placed counterfactual prompts foster anomaly detection and hypothesis revision in 5-year-olds. To avoid some of the difficulties associated with processing anomalies described above (i.e., the effects of prior knowledge and cognitive load), children were introduced to a novel causal system that does not rely upon prior knowledge, and were not required to maintain the information that they observed in working memory.

We predict that counterfactual prompts will facilitate children's ability to detect anomalies, leading them to (1) privilege the hypothesis that accounts for a greater proportion of the data, and (2) generalize this newly learned rule to novel cases. The ability to generalize would indicate that counterfactuals not only help children to notice anomalies and revise their hypotheses in the specific setting in which they were observed, but also to apply newly learned causal relations to novel situations, which is a hallmark of scientific reasoning.

Methods

Participants

Forty-eight 5-year-olds (*M*=64.6 months, *SD*=3.67 months, range=58.8-71.4; 19 females) were included in the study. Children were randomly assigned to *control* (*n*=24) or *counterfactual* (*n*=24) conditions, with no significant difference in age between conditions, *p*=.955. An additional 6 children were tested, but excluded due to inattention (4) or experimenter error (2). Children were recruited from local preschools, museums, and a university subject pool.

Materials

Machine. The machine used in the training phase was a "blicket detector" (Gopnik & Sobel, 2000), which consisted of a black (5 x 5 x 4 ¾) wooden box containing a light that was surreptitiously controlled by a remote. Certain blocks were said to *cause* the toy to light up when placed on top.

Training and observation blocks. A total of 8 1.5-inch painted wooden cube blocks were used (4 causal and 4 inert) in the *training phase* (2 blocks) and *observation phase* (6 blocks) (see Figure 1). Only the causal blocks "activated" the toy, causing it to light up. Each block had two painted sides of different colors (red, white, blue, yellow). One of the two colors represented the 100% hypothesis, and the other color represented the 75% hypothesis. During the 8 observations, one color would always cause the machine to light up, and the other color would only cause the machine to light up 75% of the time.

For example, consistent with the 100% hypothesis, all four causal blocks might have a red top, and all four inert blocks might have a blue top. Consistent with the 75% hypothesis, three out of the four causal blocks would have a white front, and one would have a yellow front, and three out of the four inert blocks would have a yellow front, and one would have a white front. The placement of the colors and the color representing the 100% cause were counterbalanced.

Two memory cards were used to help children remember which blocks were causal and which were inert, minimizing cognitive load. The causal blocks were placed next to a card with a picture of the toy with the light on. The inert blocks were placed next to a second card with a picture of the toy with the light off.

Test blocks. Two novel blocks were used to test for generalization. One block included the 100% causal color (e.g., red) plus a novel color (purple), and the other block included the 75% causal color (e.g., white) plus another novel color (green).

Procedure

Training Phase. The training phase served to familiarize children to the experiment by introducing the materials and the novel causal system. The experimenter brought out the machine and said, "This is my toy. When I put some things on top of my toy, my toy will light up. When I put other things on top of my toy, it will not light up. Let's try to find out what things will make my toy light up."

The experimenter then brought out the first block and said "Let's try this one," and placed it on the toy. The first block was always causal (e.g., red top, white front), and the toy lit up. After the child observed what happened, the experimenter asked, "Did this one make my toy light up or not light up?" After providing a verbal response, children were asked to sort the block in front of the appropriate memory card (causal or inert). This process was then repeated for an inert block (e.g., blue top, yellow front). Here, the second block was consistent with the rule that blocks with a blue top and yellow front are inert.

Observation Phase. During the observation phase, participants were randomly assigned to *control* or *counterfactual* conditions. Children observed six trials in which the experimenter placed the remaining blocks on top of the toy. After each demonstration, the child was asked two questions. The first question was the same, regardless of condition: "Let's try this one!" Then, after the demonstration they were asked, "Did this red/white one make my toy light up or not light up?" This was to ensure children were paying attention and answering correctly. If they did not answer correctly, the experimenter repeated the demonstration.

The second question differed by condition. For instance, in the *control* condition, if the child observed a red/white block causing the toy to light up, the child was asked, "Now I want you to remind me what happened. What happened when I put this *red/white* one on top of my toy? Did my toy light up or not light up?" In the *counterfactual* condition, the experimenter asked, "Now I want you to imagine something different. What if this block had been *blue/yellow*? What would have happened to my toy? Would my toy have lit up or not lit up?" The *counterfactual* prompt served to explicitly call attention to an alternative scenario (i.e., the maximally opposite color combination).

The presentation of the six blocks in the *observation phase* was pseudorandom (see Figure 1): Children first observed two blocks that followed the same pattern as the two blocks in the training phase (e.g., one block with a red top and white front activating the toy [causal] and another block with a blue top and yellow front failing to activate the toy [inert]). Next, this rule was challenged by the presentation of anomalous data. Children observed two anomalies: e.g., 1 block with a red top and yellow front activating the toy (causal), and 1 block with a blue top and white front not activating the toy (inert). Critically, this evidence violated the hypothesis that blocks with a white front are causal (75% hypothesis), but not the hypothesis that blocks with a red top are causal (100% hypothesis). Children would therefore have to recognize this violation in order to correctly infer that red (but not white) is the causal property that accounts for all of the data. Finally, children observed two blocks that were again consistent with the original pattern. All blocks remained sorted and in full view throughout the remainder of the experiment.

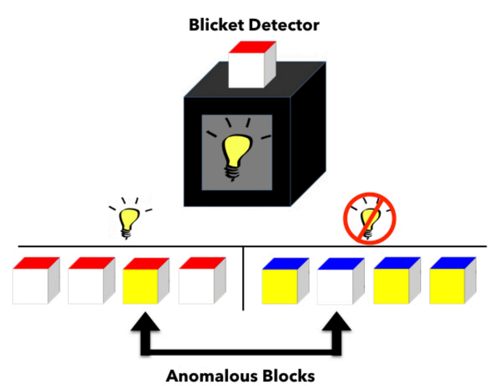

Figure 1: Schematic of the experimental paradigm

Generalization Phase. During the generalization phase, the experimenter said, "Now that you've seen how my toy works, I need your help finding more things that will make my toy light up. I have some more blocks inside of this bag. I'm going to tell you about some of these blocks and ask you some questions about them."

The experimenter then looked inside an opaque bag (the contents of which were not visible to the participant) and asked two *no-conflict* questions (*100% no-conflict* and *75% no-conflict*), followed by two *conflict* questions (*verbal conflict* and *transfer conflict*).

In all *no-conflict* questions, the causal feature corresponding to each hypothesis (*100%* or *75%*) was pitted against the inert feature corresponding to that same hypothesis. For example, in the *100% no-conflict* question, red was pit against blue, because red always activated the toy (i.e., 100% of the time), and blue always failed to activate the toy (i.e., 100% of the time). The experimenter said, "When I look inside this bag, I see one with a red part and I see one with the blue part. Which one will make my toy light up?" The order of presentation was counterbalanced. Participants were never shown the contents of the bag, and were asked to make an inference on the basis of this verbal description alone. These *no-conflict* questions were included to ensure that children in both conditions inferred the co-variation pattern between block color and activation of the machine.

For all *conflict* questions, the 100% causal color (e.g., red) was pit against the 75% causal color (e.g., white) to assess which (if any) hypothesis would be favored. The procedure for the *verbal conflict* question was identical to the *no-conflict* questions: The experimenter looked inside the bag and described features of the two blocks. However, in this case, the features were *both* associated with the effect, with varying probability (e.g., 100% red and 75% white).

Finally, in the *transfer conflict* question, the experimenter placed two novel blocks on the table in front of the child. The first block was composed of the 100% causal color (e.g., red), paired with a novel color (e.g., purple). The second block was composed of the 75% causal color (e.g., white), also paired with a different novel color (e.g., green). The experimenter said, "I have two new blocks. Which one will make my toy light up? The one with the red/purple part or the one with the green/white part?" This question similarly served to pit the two hypotheses against one another, but also required that children generalize the privileged hypothesis to a novel set of blocks. In order to answer both types of *conflict* questions correctly, it was necessary for children to notice the anomaly to the 75% causal color, and to use this information to inform their inference.

Results

Analysis considered three questions: (1) Did all children learn the 75% and 100% causal hypotheses? (2) Do counterfactual prompts facilitate early anomaly detection and hypothesis revision? and (3) Do counterfactual prompts help children to generalize a newly learned causal rule to a novel case?

Average proportions of responses for all question types in each condition appear in Figure 2. As predicted, children in both the *control* and *counterfactual* conditions performed above chance on the 100% *no-conflict* questions (*control*: M=.96, SD=.20; *counterfactual:* M=.96, SD=.20), p<.001 (exact binomial test), and the 75% *no-conflict* questions (*control*: M=.79, SD=.41; *counterfactual:* M=.79, SD=.41), p<.01 (exact binomial test). These results indicate that, as predicted, children in both conditions successfully learned the novel causal structure, with no difference between conditions, p=1.

We next analyzed performance on both the *verbal* and *transfer conflict* questions. In contrast to our prediction, children in both the *control* (M=.79, SD=.41) and *counterfactual* (M=.75, SD=.44) conditions responded in line with the 100% hypothesis (p<.01 and p<.05, respectively) on the *verbal conflict* question with no significant difference between conditions, $X^2(1)$=.73, p=.50 (one-tailed), φ=.05 (Fishers exact).[1]

However, in response to the *transfer conflict* question, children in the *counterfactual* condition (M=.83, SD=.38) privileged the 100% hypothesis, p<.01 (exact binomial test), while those in the *control* condition (M=.54, SD=.51) selected between the two options at chance, p=.31. There was also a significant difference between conditions, with children in the *counterfactual* condition more likely to privilege the 100% color at transfer, $X^2(1)$=4.75, p=.02 (one-tailed), φ=.31.

Discussion

The current study investigated the effects of strategically placed counterfactual prompts in facilitating anomaly detection in children as young as 5 years of age. Results indicate that counterfactual prompts helped young learners to privilege the hypothesis that accounted for a greater proportion of their observations when predicting a novel causal outcome. Most importantly, children who were prompted to think counterfactually were more likely to transfer this newly learned causal rule to inform their judgments about a set of novel objects. These findings indicate that counterfactual prompts offer an effective means of constraining some of the key processes underlying causal reasoning, including anomaly detection, belief revision, and the generalization of causal knowledge.

Surprisingly, however, when the 100% causal color was *verbally* contrasted with the 75% causal color, children in the *control* condition performed equally well, even in the absence of additional scaffolding. This result contrasts with children's baseline performance in previous findings, in which they selected between the two hypotheses at

[1] One-tailed tests reflect the directional nature of our hypothesis (i.e., that counterfactual prompts would improve performance).

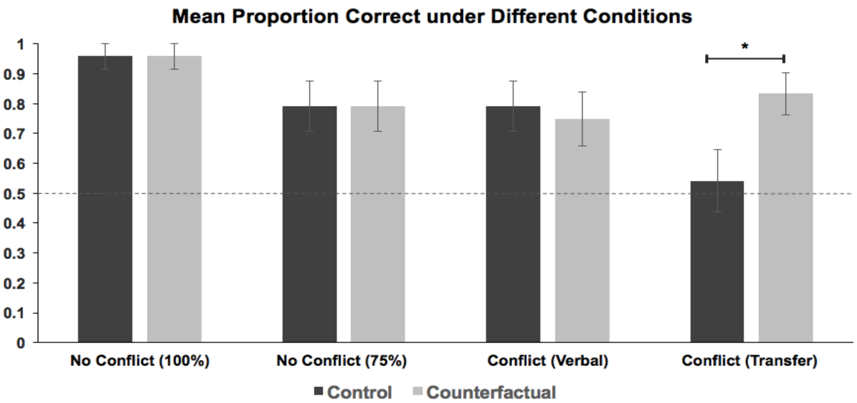

Figure 2: Mean proportion of correct responses for *No-Conflict* and *Conflict Items* in *control* and *counterfactual* conditions.

chance (Walker et al., 2016). It is possible that a feature of the current control condition may have inadvertently boosted performance, leading to this inconsistent result. In particular, unlike in the previous research (Walker et al., 2016), the experimenter in the current study both verbalized and highlighted the specific features of the blocks in the scaffolding phase (e.g., "Did this *red/white* one make my toy light up or not light up?), and repeated this prompt twice for each observation. We believe that this may have reinforced a verbal representation of the specific causal rules.

Of course, this pattern of data is also compatible with the possibility that the control prompt served to hinder children's performance on the *transfer conflict* question. Although the current data cannot definitively rule out this alternative, this interpretation is unlikely when considered in conjunction with Walker and colleagues prior work. That said, ongoing research is aimed at reassessing performance across conditions using picture cards to bypass verbal repetition of the block features.

Future work will also consider cases in which anomalous observations challenge more entrenched prior knowledge, and will examine the specific effects of different types of scaffolds (e.g., explanation vs. counterfactuals) on the efficiency of belief revision. In this case, counterfactual prompts produced similar effects on children's inferences as prompts to explain (Walker et al., 2016), although we expect that the underlying mechanisms are likely quite different. Finally, additional work is required to investigate the robustness of these findings across contexts, and to consider the nature of the particular counterfactual question that is posed.

In sum, there are clear benefits to successfully identifying simple prompts that engage particular cognitive processes since they are relatively easy to incorporate into any learning environment, with few resources and little teacher training (Williams et al., 2016). The current findings provide evidence supporting the use of *counterfactual* scaffolds in facilitating anomaly detection and fostering belief revision. These scaffolds not only facilitate the detection of anomalous data, but also help children to generalize and transfer newly learned causal rules to novel contexts, which is an essential component of scientific reasoning.

Acknowledgments

We thank Leslie Carver, Patricia Ganea, and Angela Nyhout, for their early contributions to this work; Hazel Baker-Harvey, Kathryn Knowlton, Andrea Nishimi, and Nicky Sullivan for assistance with data collection; members of the Early Learning and Cognition Laboratory at UCSD, the Ruben H. Fleet Science Center, and participating parents and children.

References

Brewer, W. F., & Chinn, C. A. (1991). Entrenched beliefs, inconsistent information, and knowledge change. In *Proceedings of the 1991 International Conference on the Learning Sciences* (pp. 67-73). Association for the Advancement of Computing in Education, VA.

Buchsbaum, D., Bridgers, S., Weisberg, D. S., & Gopnik, A. (2012). The power of possibility: Causal learning, counterfactual reasoning, and pretend play. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2202-2212.

Carey, S. (2009). *The Origin of Concepts*. New York: Oxford University Press.

Chinn, C. A. & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of educational research, 64 (1), 1-49.*

Chinn, C. A., & Brewer, W. F. (1998). An empirical test of a taxonomy of responses to anomalous data in science. *Journal of research in science teaching*, *35*(6), 623-654.

Chinn, C. A., & Malhotra, B. A. (2002). Children's responses to anomalous scientific data: How is conceptual change impeded?. *Journal of Educational Psychology*, *94*(2), 327.

Galinsky, A. D., & Moskowitz, G. B. (2000). Counterfactuals as behavioral primes: Priming the simulation heuristic and consideration of alternatives.

Journal of Experimental Social Psychology, *36*(4), 384-409.
Gopnik, A. (2012). Scientific thinking in young children: Theoretical advances, empirical research, and policy implications. *Science*, *337*(6102), 1623-1627.
Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, *37*(5), 620.
Gopnik, A., & Walker, C.M. (2013). Considering counterfactuals: The relationship between causal learning and pretend play. *American Journal of Play, 6*(1), 15-28.
Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological bulletin*, *138*(6), 1085.
Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, *61*(3), 233-259.
Hirt, E. R., & Markman, K. D. (1995). Multiple explanation: A consider-an-alternative strategy for debiasing judgments. *Journal of Personality and Social Psychology*, *69*(6), 1069.
Koerber, S., Osterhaus, C., & Sodian, B. (2017). Diagrams support revision of prior belief in primary-school children. *Frontline Learning Research, 5(1), 76-84*
Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. Mit Press.
Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin*, *108*(3), 480.
Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96, 674-689
Kuhn, D., Amsel, E., O'Loughlin, M., Schauble, L., Leadbeater, B., & Yotive, W. (1988). *The development of scientific thinking skills*. Academic Press.
Lewis, D. (1986). Counterfactuals. Cambridge, MA: Harvard University Press.
Legare, C. H., Gelman, S. A., & Wellman, H. M. (2010). Inconsistency with prior knowledge triggers children's causal explanatory reasoning. *Child development*, *81*(3), 929-944.
Mackie, J. L. (1974). Truth, probability, and paradox: A reply to James E. Tomberlin's review. *Philosophy and Phenomenological Research, 34*(4), 593-594.
McCormack, T., Simms, V., McGourty, J., & Beckers, T. (2013). Encouraging children to think counterfactually enhances blocking in a causal learning task. *Quarterly Journal of Experimental Psychology*, *66*(10), 1910-1926.
Pearl, J. (2000). Causal inference without counterfactuals: Comment. *Journal of the American Statistical Association*, *95*(450), 428-431. doi: 10.2307/2669380
Penner, D. E., & Klahr, D. (1996). The Interaction of Domain- Specific Knowledge and Domain- General Discovery Strategies: A Study with Sinking Objects. *Child Development*, *67*(6), 2709- 2727.
Riggs, K.J., & Peterson, D.M. (2000). Counterfactual and mental state inferences in pre-school children. IN P. Mitchell & K.J. Riggs (Eds). *Children's Reasoning and the Mind*. Psychology Press.
Schauble, L. (1990). Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of experimental child psychology*, *49*(1), 31-57.
Schulz, L. E., Bonawitz, E. B., & Griffiths, T. L. (2007). Can being scared cause tummy aches? Naive theories, ambiguous evidence, and preschoolers' causal inferences. *Developmental psychology*, *43*(5), 1124.
Schulz, L.E., Goodman, N.D., Tenenbaum, J.B., & Jenkins, A.C. (2008). Going beyond the evidence: Abstract laws and preschoolers' reponses to anomalous data. *Cognition,* 109(2), 211-223
Sloman, S. (2005) *Causal models: How people think about the world and its alternative*s. Oxford University Press.
Walker, C. M., & Gopnik, A. (2013). Causality and imagination. *The Oxford handbook of the development of the imagination*, 342-358.
Walker, C.M., Lombrozo, T., Williams, J. J., Rafferty, A.N., & Gopnik, A. (2016). Explaining constrains causal learning in childhood. *Child Development*.
Wenzlhuemer, R. (2009). Counterfactual thinking as a scientific method. *Historical Social Research/ Historische Sozialforschung*, 27-54.
Williams, J. J., & Lombrozo, T. (2010). The role of explanation in discovery and generalization: Evidence from category learning. *Cognitive Science*, *34*(5), 776-806.
Williams, J. J., Lombrozo, T., Hsu, A., Huber, B., & Kim, J. (2016). Revising Learner Misconceptions Without Feedback: Prompting for Reflection on Anomalous Facts. *Proceedings of CHI (2016), 34th Annual ACM Conference on Human Factors in Computing Systems*
Woodward, J. (2003). *Making things happen: A theory of causal explanation.* Oxford University Press: Oxford.
Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review, 27*(2), 172-223.