CHAPTER SEVEN

Scientific reasoning and counterfactual reasoning in development

Angela Nyhout[a,*] and Patricia A. Ganea[b]
[a]School of Psychology, University of Kent, Canterbury, United Kingdom
[b]Department of Applied Psychology and Human Development, University of Toronto, Toronto, ON, Canada
*Corresponding author: e-mail address: a.nyhout@kent.ac.uk

Contents

1. Introduction and overview 224
2. Scientific thinking and reasoning 224
3. Counterfactual reasoning 228
4. Connecting the dots: Counterfactual, causal, and scientific reasoning 234
   4.1 Developmental evidence 236
5. Counterfactual thought experiments 240
   5.1 Counterfactuals and the inquiry process 241
   5.2 Counterfactuals and scientific knowledge 245
6. Future directions for counterfactual thought experiments in research and education 247
7. Conclusions 249
References 249

Abstract

In this chapter, we bridge research on scientific and counterfactual reasoning. We review findings that children struggle with many aspects of scientific experimentation in the absence of formal instruction, but show sophistication in the ability to reason about counterfactual possibilities. We connect these two sets of findings by reviewing relevant theories on the relation between causal, scientific, and counterfactual reasoning before describing a growing body of work that indicates that prompting children to consider counterfactual alternatives can scaffold both the scientific inquiry process (hypothesis-testing and evidence evaluation) and science concept learning. This work suggests that counterfactual thought experiments are a promising pedagogical tool. We end by discussing several open questions for future research.

Advances in Child Development and Behavior, Volume 61
ISSN 0065-2407
https://doi.org/10.1016/bs.acdb.2021.04.005

1. Introduction and overview

Why do spiders make webs? How can we effectively combat this virus? Is there life on other planets? Scientific questions pervade our everyday lives, whether we are young children in classrooms, laypeople consuming scientific media, or scientists in laboratories. Through scientific inquiry, we come closer to answering such questions and gaining a richer and more accurate understanding of the world. However, extensive past work has documented limitations in scientific inquiry in development, including in hypothesis formation, hypothesis-testing, and evidence evaluation. Given the pervasiveness and importance of scientific inquiry and understanding, identifying how best to foster these abilities is a core objective for developmental psychologists and educators alike (Zimmerman, 2007).

Although the goal of scientific inquiry is to understand the world, we are not confined to reasoning about the world *as it is*. From early in development, children are capable of reasoning about the world as it *could be*. This ability to imagine and reason about the world as it could be is termed counterfactual reasoning. But why spend time imagining things that *could be* when we want to understand what *is*? In this chapter, we build on existing theories and evidence to explore the claim that counterfactual reasoning functions as an imagined experiment, allowing the learner to mentally manipulate their representations of the world to glean new knowledge and understanding.

We begin by reviewing research on the development of scientific and counterfactual reasoning in development, taking into account key debates about the nature and development of these abilities. We then consider the relation between counterfactual, causal, and scientific reasoning and discuss several different theoretical proposals on their connection. In the final sections, we turn to findings on counterfactual thought experiments which indicate that engaging individuals—from young children to adults—in counterfactual reasoning can influence both the scientific inquiry process and the scientific inferences they make. We conclude with considerations for future research in this emerging area.

2. Scientific thinking and reasoning

The skills required to reason scientifically develop over many years through both formal and informal educational experiences. Scientific thinking has been characterized and studied in terms of two main features: (1) the knowledge base or theoretical content that one is relying on to make sense of the world around, and (2) the processes or strategies used to bring about this knowledge, such as hypotheses generation, experimentation, and evidence evaluation. Decades of research on the development of scientific thinking have demonstrated that scientific reasoning is challenging. At all ages, rather than generating evidence to test existing theories, individuals generate evidence to merely confirm them. Even when observing disconfirming evidence, individuals either ignore or attempt to fit it within their current theories, as opposed to adjusting their theories (Kuhn, 1989). Misconceptions about science are also prevalent across ages and co-exist alongside valid scientific theories of a domain, even after formal training (Coley & Tanner, 2012; Potvin & Cyr, 2017; Shtulman & Valcarcel, 2012). Because misconceptions influence how we view the world and the acquisition of new knowledge, the question of how we address them when teaching children domain-specific knowledge is essential.

Children start formal school with a good base of structured, abstract knowledge in various domains, including biology (Gelman & Wellman, 1991), physics (Bullock, Gelman, & Baillargeon, 1982; Spelke, Breinlinger, Macomber, & Jacobson, 1992) and psychology (Perner, 1991). Their naive, intuitive theories contain misconceptions that are resistant to revision and thus can make science teaching challenging. Examples of such misconceptions in the domain of physics include the idea that heavy objects sink and light ones float (Penner & Klahr, 1996; Potvin & Cyr, 2017), that the weight of an object determines the speed with which it falls (Hast & Howe, 2012; Kavanagh & Sneider, 2007; Pine, Messer, & St. John, 2001), that the shape of earth is flat (Vosniadou & Brewer, 1992), or that objects in motion have impetus or "internal force" and as a result they continue to move perpendicular to the ground when in free fall (McCloskey, 1983). Overcoming scientific misconceptions has proven difficult and can sometimes require revision of the naïve theory or mental model that the individual holds about a domain (Hardy, Jonen, Möller, & Stern, 2006). The predominant view in developmental psychology has been that young children engage in exploratory testing behaviors to examine the fit between their naive theories of a domain and new evidence (Gopnik & Meltzoff, 1998; Gopnik & Wellman, 2012).

Just as formal scientists do, children actively and spontaneously formulate, test, and revise hypotheses in light of new evidence. This view of the child as scientist is consistent with the influential classical model of conceptual change proposed by Posner, Strike, Hewson, and Gertzog (1982), which argues that presenting learners with contradictory or anomalous evidence is an essential instructional goal. The anomalous evidence has the potential to highlight an explanatory gap in one's theory and lead the learner to hypothesize about alternative explanations.

In line with these influential theories of knowledge acquisition and conceptual change, it is often assumed that children learn best from hands-on concrete experiences and many educational boards recommend an inquiry-approach to teaching of scientific concepts (Hardy et al., 2006; National Research Council, 2000). This approach assumes that children learn best through their active exploration of the world by engaging in hypothesis-testing behaviors when presented with new evidence. However, a large body of developmental and educational research has shown that this hands-on approach to science learning is ineffective because of limitations in children's ability to revise beliefs based on evidence alone. The failure to learn on the basis of counterevidence is due to two main factors. First, naïve beliefs influence the type of evidence that both children and adults attend to, and in the face of contradictory evidence they tend to either disregard or reinterpret this evidence to fit their naïve beliefs (Chinn & Brewer, 1993; Koslowski, 1996; Penner & Klahr, 1996). Second, both children and adults have poor experimentation skills, in that they have difficulty coordinating theory and evidence (Kuhn, 1989), design confounded experiments (Klahr & Nigam, 2004) and draw invalid conclusions even when presented with valid experiments (Renken & Nunez, 2010). Children need formal training on how to engage in hypothesis-testing behaviors, especially when robust misconceptions are involved (Klahr & Nigam, 2004; Koslowski, 1996; Kuhn et al., 1995; Zimmerman, 2007).

In a recent study from our lab (Ganea, Larsen, & Venkadasalam, 2021) we used a tightly-controlled design that contrasted anomalies-based instruction with explanation-based instruction when children held the misconception that weight, rather than density determines whether an object sinks or floats. We aimed to identify the individual and additive effects of these approaches on scientific belief revision, and the effect of their sequence in the teaching process. The results showed that 5-year-old children did not revise their misconceptions about sinking and floating when observing evidence that contradicted their naive beliefs—they continued to rely on their naïve belief about weight as the determining factor in causing an object to sink or float to make new predictions and justifications. In contrast, when children were exposed to rich conceptual explanations about what makes objects sink and float, they were more likely to set aside their naïve beliefs and rely on this information to make new predictions and justifications. Our study adds to the body of literature showing that simply presenting conflicting evidence does not necessarily lead children to revise their naïve misconceptions about science concepts. In the absence of an alternative theory, children revert to their naïve beliefs when asked to explain the anomalies. We also found that the timing of when an alternative viable explanation is received is important. Having access to an alternative theory as the conflicting evidence is observed leads to more belief revision than if the alternative theory is received after the evidence is observed (Ganea et al., 2021). These results speak to the limitations of anomalies-based instruction and the importance of rich conceptual explanations and their timely delivery for science education.

Given the emphasis of science curricula on learning through direct, hands-on experience with evidence and concrete materials, and the lack of evidence that such learning is effective, it is worth considering other ways in which children can acquire science knowledge. In this chapter we argue that imagination, in the form of counterfactual prompts, can engage children's ability to consider alternative possibilities that will lead them to overcome current misconceptions. According to Harris (2021), children's imagination—early pretend play, future thinking, thinking about what is possible and not, and figurative drawing—is reality-based and it may serve the purpose of helping children navigate reality. The process of mentally considering alternative realities to better understand the structure of the real world is evident in early pretend play in the second year of life. Children's pretend play tends to follow familiar scripts, reflects day-to-day routines, and is influenced by the same causal constraints that influence reality itself (Harris, 2021). Its reality-based structure and focus suggest that children might be able to learn about the world without engaging in direct experimentation with concrete materials, and instead use their imagination to make new discoveries. In the domain of science learning this idea has been fairly unexplored.

There is some evidence that when children are prompted to use their imagination, they are better able to overcome biases that are rooted in existing misconceptions. For example, when asked to make predictions about the outcome of dropping a ball through crisscrossed tubes, 2- and 3-year-old children commit the gravity error, predicting that the ball will end up directly in the cup positioned under the location where was inserted—rather than following the path of the tube to a different cup

(Hood, 1995). This error is robust and persists across trials even after feedback of the ball's final location. Nevertheless, if children are instead prompted to first imagine the ball's path, they are more likely to make correct predictions about its final location (Joh, Jaswal, & Keen, 2011; Palmquist, Keen, & Jaswal, 2018). Work by Bascandziev and Harris (2010) has shown that children's improvement in performance on the tube task when prompted to use their imagination lasts even on subsequent trials when the prompts are withdrawn. These findings are promising because they suggest that children's imagination can be used to get them to override their naïve beliefs by considering alternatives. Children's imagination "… can serve as mental laboratory, or Twin Earth, where they can make discoveries and undermine misconceptions about the real world." (Harris, 2021, pp. 20).

In some of the work we describe here we have begun to explore the pedagogical benefits of prompting children to engage in thought experiments and whether the outcome of such thought experiments promotes their understanding of reality. We have focused on a particular type of imaginative prompt that recruits counterfactual reasoning. Counterfactual reasoning involves reasoning from premises that are counter to what one knows about reality at a certain point in time. Children as young as 2 years old can reason from false premises. For example, when asked to entertain the premise that fish live in trees on a syllogistic reasoning task, children make the correct deductive inference that a fish, Tot, lives in a tree (Dias & Harris, 1988). Other work shows that 3-year-olds can use their imagination to manipulate representations of past and current reality when prompted to consider counterfactual alternatives (e.g., Harris, German, & Mills, 1996). Below we outline the development of children's counterfactual reasoning and describe studies that explore the benefit of exploiting the imagination in the service of science learning.

3. Counterfactual reasoning

Before considering how and whether counterfactual reasoning contributes to science learning, it is important to first consider its development. There is substantial debate among developmental psychologists about the breadth and nature of counterfactual reasoning (for a debate, see Beck, 2016; Weisberg & Gopnik, 2013, 2016). This debate focuses on whether different abilities to think about alternatives to reality are qualitatively different from one another, and the related issue of when in development children can be credited with the ability to reason counterfactually. In this section, we review the ongoing theoretical debate and relevant developmental evidence, given that the types of representations children can manipulate counterfactually and the types of counterfactual inferences they draw at different stages in development are important considerations for understanding how and whether counterfactuals can be leveraged for science learning.

Some researchers conceptualize counterfactual thinking broadly and argue that it encompasses a diverse set of abilities to reason hypothetically about past, present, future, and fictional possibilities (Buchsbaum, Bridgers, Skolnick Weisberg, & Gopnik, 2012; Weisberg & Gopnik, 2013). On this view, counterfactual thinking is available to children early in development, with the first evidence usually appearing in pretend play. According to Weisberg and Gopnik (2013), counterfactual reasoning involves three key components: disengaging from the current reality, introducing and making inferences about an alternate reality, and keeping representations of reality and the alternative separate from one another. In support of this view, Buchsbaum and colleagues found a strong correlation between children's reasoning about novel causal relations during pretend play and their ability to answer explicit hypothetical questions (Buchsbaum et al., 2012). Three- and four-year-olds were introduced to a "birthday machine" that played the song "Happy Birthday" when certain types of objects, called zandos, were placed on it, but did not play a tune when another type of block was placed on it. During the test phases, children were asked counterfactual questions (e.g., "What if this zando were not a zando; what would happen if we put it on the machine?"). In another test phase, children were told the real birthday machine had to be taken away, and were asked to pretend with a stand-in box and blocks that did not play music. Children who answered counterfactual questions correctly were those who also made correct inferences during the pretend play phase.

More generally, but also in support of the claim that counterfactual reasoning develops early, a substantial body of research finds that children begin passing counterfactual reasoning tasks in the preschool years. The sophistication of this ability and its underlying mechanisms are still the subjects of substantial debate. In a seminal study, Harris et al. (1996) presented children with a story about a character who left her muddy shoes on and made the kitchen floor all dirty. Children as young as 3½ years old were able to reason that if she had removed her shoes, the floor would have been clean. Several other studies have found that children begin passing counterfactual reasoning tasks around their fourth birthday (e.g., Beck, Robinson, Carroll, & Apperly, 2006, standard counterfactuals; Guajardo & Turley-Ames, 2004; Perner, Sprung, & Steinkogler, 2004; Riggs, Peterson, Robinson, & Mitchell, 1998).

Other researchers subscribe to a narrower view of counterfactual thinking and contrast real world counterfactuals and general counterfactuals (Beck, 2016; Beck & Riggs, 2014). Real world counterfactuals are those about past episodes, and involve holding in mind (and in some cases comparing between) representations both of reality and a counterfactual alternative. In contrast, general counterfactuals include the broader category of abilities described by Weisberg and Gopnik, including pretend play and future hypotheticals. Beck and colleagues argue that the ability to reason about real world counterfactuals develops around age 6. For instance, Beck et al. (2006) showed 3–6-year-old children an apparatus that included three slides a mouse could go down, and the child was tasked with placing pillows at the bottom to cushion the mouse's landing. After the mouse went down a slide, children were asked counterfactual or future hypothetical questions. Preschoolers were able to reason about close-ended possibilities (e.g., "What if he had gone the other way?"), but had considerably more difficulty reasoning about open-ended ones (e.g., "Where else could he have gone?"). Other work indicates that children find it easier to reason about future hypothetical compared to past counterfactual alternatives (Robinson & Beck, 2000). These findings suggest not only that counterfactual and future hypothetical thinking may be separable constructs, but also that counterfactual thinking itself may not be a unitary construct. When asked to reason about multiple possibilities (compared to just one alternative possibility), children had more difficulty.

An even stricter definition of counterfactual thinking comes from Rafetseder and colleagues, who argue that children do not develop counterfactual thinking until middle childhood to adolescence when they respect the *nearest possible world constraint* (Rafetseder, Cristi-Vargas, & Perner, 2010; Rafetseder & Perner, 2014; Rafetseder, Schwitalla, & Perner, 2013). On this view, the reasoner must change only one aspect of a past event and hold all other features constant (Edgington, 2011). When the reasoner stipulates a counterfactual antecedent, they should change only those features that are causally dependent on the new antecedent, holding all else constant. To test this claim, Rafetseder et al. (2013) presented children with causally overdetermined scenarios in which two antecedent events lead to the same outcome. Children were then asked how the outcome would have been different if one of the antecedent events had not happened.

For example, a kitchen floor got muddy by two separate characters walking on it with muddy boots. It was not until the age of 13 that children correctly answered that if one of the characters had removed their boots, the floor would still have been muddy due to the other character's boots. Rafetseder et al. (2010, 2013) and Rafetseder and Perner (2014) have argued that young children who correctly answer counterfactual questions in single-cause scenarios (e.g., only one character with muddy shoes) are relying on *basic conditional reasoning*. In this simpler form of reasoning, children use general causal knowledge to arrive at a correct answer (e.g., clean feet are associated with clean floors) and do not construct a genuine counterfactual representation. However, there are other possible interpretations of these findings, which we discuss below.

In the following sections, we take the view we developed in Nyhout and Ganea (2019a) that counterfactual reasoning is a cognitive process involving introducing a change to a *mental representation*—this mental representation may be a past event from episodic memory or a model of a causal system (e.g., an ecosystem) from semantic memory. Thus, we characterize counterfactual thinking according to the *process*, rather than the type (e.g., episodic or semantic) and complexity of the input. We argue that counterfactual reasoning involves (1) retrieving a representation of a system of two or more causally related variables, (2) manipulating the representation by positing a false premise, and (3) inferring the causal implications of the false premise. Children's ability to reason counterfactually in a given situation will depend on the complexity of the event or causal model they are reasoning about.

In Nyhout and Ganea (2019a), children's ability to reason counterfactually is argued to rely on their ability to encode, retrieve, and represent events and causal systems. In particular, we argue that a child's model of reality will partly determine the counterfactual inferences the child makes. How the child interprets an event or system determines the parameters of his or her model in terms of the type, number, and nature of causal relations between events or entities. For instance, in the scenarios used by Rafetseder et al. (2013) in which two individuals walked on a floor with muddy shoes, children may represent the events in different ways; one child may represent the two antecedent events as independent (e.g., the two went into the kitchen separately) and another may represent them as connected (e.g., one character followed the other). These differences in their *model of reality* will influence the counterfactual inferences the two children make.

To test this prediction, in Nyhout, Henke, and Ganea (2019) we gave 6- to 8-year-old children events with more versus less clearly specified causal structures and asked them counterfactual questions. Eight-year-olds made correct counterfactual inferences both when presented with narratives in which antecedent events were disconnected as in a common effect model (e.g., the two characters go into the kitchen separately) and connected as in a causal chain model (e.g., one character follows the other). Six-year-olds showed success on scenarios in which antecedent events were causally connected to one another, and therefore the causal models were more clearly specified, but not when they were disconnected. Performance was chance-like for children of all ages when the causal relations between events were ambiguous, as in Rafetseder et al.'s (2013) study.

A comparison across studies also supports the claim that the complexity of a causal model influences children's ability to reason counterfactually. Children appear to reason earlier about singly- versus over-determined causal models. As described above, whereas 4-year-olds in Harris et al.'s (1996) could reason counterfactually about the muddy floor scenario when there was a single character causing the mess (Harris et al., 1996; Rafetseder et al., 2013), children do not appear to reason about overdetermined versions of such scenarios until age 6 at the youngest (Nyhout, Henke, & Ganea, 2019; Nyhout, Iannuzziello, Walker, & Ganea, 2019).

Studies looking at children's reasoning about direct versus indirect causes provide further evidence that the complexity of the causal model predicts children's counterfactual reasoning. For instance, German and Nichols (2003) presented 3- to 4-year-old children with stories involving a causal chain of events. In one story, Mrs. Rosy has just planted a flower in her garden and calls her husband to come see it. Her husband opens the door to come outside, letting their dog out, who then tramples on the flower, making Mrs. Rosy sad. Children were asked how Mrs. Rosy would feel after a counterfactual alternative that was direct (e.g., if the dog had not trampled on the floor) or indirect (e.g., if her husband had not opened the door). Four-year-olds were better able to answer direct than indirect questions about counterfactual changes (German & Nichols, 2003). Beck, Riggs, and Gorniak (2010) attempted to replicate this finding, but found that 4-year-olds answered both direct and indirect counterfactual questions incorrectly. The authors suggested this may have been due to the language in which children were studied (Greek in German & Nichols, English in Beck et al.). In a recent study, we found a similar pattern to German and Nichols (2003) using a novel ecosystem task (Nyhout, Sweatman, & Ganea, 2021). This study is described in more detail in Section 5. Together, these results show that reasoning about indirect effects of counterfactual alternatives is harder for children than reasoning about direct effects.

The causal domain under consideration also plays a role in children's ability to reason counterfactually, and children's counterfactual inferences appear to track with their domain-specific causal reasoning (Sobel, 2011). In most studies, researchers have presented children with short narratives mostly pertaining to psychological or physical causal relations. However, there may be something particularly difficult about reasoning about fictional events involving third parties. Children can reason counterfactually about overdetermined *physical* events that they have directly witnessed at a younger age compared to fictional narratives. In Nyhout and Ganea's (2019b) study, 3-, 4- and 5-year-olds were familiarized with a "blicket detector" toy that lit up when certain types of blocks were placed on it and did not light up when other types of blocks were placed on it. On critical *overdetermined* trials, two causal blocks were placed on the box, the box lit up, and children were asked if the lights would still be on if one of the two blocks had not been placed on the box. Four-year-olds answered these questions at a level significantly above chance, and 5-year-olds' performance was at ceiling. Another study using a more complex apparatus found that children could reason counterfactually about overdetermined events by 6 years of age (McCormack, Ho, Gribben, O'Connor, & Hoerl, 2018).

We hope it is clear at this point that counterfactual reasoning is not simply something that children can or cannot engage in. A 4-year-old may be able to reason about simple causal models, whereas an 8-year-old may struggle with more complex ones. Their ability to do so depends substantially on the nature and complexity of the subject matter they are reasoning about (Nyhout, Henke, & Ganea, 2019; Nyhout, Iannuzziello, et al., 2019). This will impact what educators can ask them to think about in educational contexts. Many adults would be hard-pressed to reason counterfactually about a complex historical event or scientific model, and indeed many such examples do not have a single "correct" answer. Questions like "What would have happened if JFK had not been assassinated?" are frequent subjects of counterfactual inquiry, but one could generate a multitude of different answers to this question.

Having reviewed key developmental findings in the areas of scientific inquiry and counterfactual reasoning, in the following section we lay out the theoretical foundations connecting counterfactual, causal, and scientific reasoning.

4. Connecting the dots: Counterfactual, causal, and scientific reasoning

Although our focus in this chapter is on counterfactual and scientific reasoning, many relevant findings on their relation come from studies of causal reasoning. Counterfactual, causal, and scientific reasoning are closely related constructs. For our purposes, we consider causal reasoning as the ability to reason about both specific causes and generic categories of causes—sometimes referred to as causal selection and causal reasoning, respectively (e.g., Woodward, 2011). We refer to scientific reasoning as we have in the previous sections as something that is applied in more formal settings involving experimentation and evidence evaluation, carried out in pursuit of scientific understanding. Causal reasoning is thus an essential component of scientific reasoning, and a primary goal of scientific inquiry is to uncover the causal fabric of the world. In this section, we first discuss accounts of the relation between counterfactual and causal reasoning, drawing on previous theoretical proposals by several researchers. We reserve discussion of the relation between counterfactual reasoning and more formal scientific reasoning for the following section, *Counterfactual thought experiments*. However, we note that many of the proposals we outline in this section on causal reasoning also apply to scientific reasoning.

Researchers have proposed several relations between causal and counterfactual reasoning. These theories are described in more detail below, followed by relevant empirical evidence. We focus on three broad groups of theories: (1) *causal primacy view*: causal knowledge is an essential component of counterfactual reasoning, but counterfactual reasoning is not required for causal reasoning, (2) *psychological process view*[a]: computing a counterfactual is a *pre-requisite* for establishing a causal inference, (3) *psychological relatedness view:* counterfactual dependence is an essential commitment of causal inference. This is not an exhaustive list of theories on the relation between causal and counterfactual reasoning, and we have selected those that are most relevant to developmental considerations. There are different sub-theories under each of these groups of theories, and the groups may at times share overlapping elements (e.g., a need to simulate a counterfactual alternative).

[a] The terms "psychological process" and "psychological relatedness" were put forward by McCormack, Frosch, and Burns (2011).

Some of these claims are about the nature of causality itself, but all of those outlined here have been taken to inform how causal and counterfactual cognition.

Although substantial debate exists over the exact nature of the relation between causal and counterfactual reasoning, most accounts agree that the two are intimately connected. Reasoning about counterfactual scenarios depends on correctly representing causes in each model (Nyhout & Ganea, 2019a; Nyhout, Henke, & Ganea, 2019; Nyhout, Iannuzziello, et al., 2019) and importing domain-specific causal knowledge is often required for drawing counterfactual inferences (Sobel, 2011). Theoretical proposals differ in the primacy they give to causal versus counterfactual inferences.

Causal primacy accounts argue that causal knowledge is essential to counterfactual reasoning, but counterfactual reasoning is not an essential component of causal reasoning (e.g., Edgington, 2011). Counterfactual reasoning may be used to other ends (e.g., to make better future decisions; Roese & Epstude, 2017) but is neither a necessary nor common factor in causal inferencing. Thus, causal knowledge is required for counterfactual inferences, but not the other way around.

Similarly, Byrne (2007) argues that the relation between counterfactuals and causal judgments is unidirectional, whereby representing a counterfactual activates its real-world causal counterpart, but the converse is not true. Representing a causal relation does not influence the accessibility of a counterfactual alternative. Therefore, counterfactual reasoning is not an essential component of causal reasoning, but considering a counterfactual can and will make relevant causal relations more accessible (Roese & Olson, 1997; Tetlock & Lebow, 2001).

Those subscribing to a *psychological process view* assign an important role to counterfactual reasoning in causal inferencing. On one view, making causal inferences requires one to first consider a counterfactual alternative by comparing an actual sequence of events to an imagined alternative (e.g., Harris et al., 1996; Lewis, 1973; Mackie & Mackie, 1974). For instance, a reasoner may make the inference "the wind through the open window caused the candle to extinguish" by first considering that if the window had been closed, the flame would not have gone out.

Finally, on a *psychological relatedness view*, counterfactual inferences are a necessary commitment of but not necessarily a precursor to causal inferencing (e.g., Gopnik & Schulz, 2007; Pearl, 2000; Schulz, Gopnik, & Glymour, 2007; Woodward, 2005). These theories are often referred to as counterfactual or interventionist views of causation. Proponents argue that counterfactual dependence is a defining feature of causation. That is, when an individual makes a causal inference (X causes Y), they commit to its counterfactual (a change to X would lead to a change to Y).

But what is the evidence for these different theories? Current findings are mixed. On balance, however, most studies appear to show a lack of an essential correspondence between individuals' causal and counterfactual judgments. We review some of the main developmental findings here.

4.1 Developmental evidence

Researchers have investigated the correspondence between individuals' causal and counterfactual judgments about *specific* events. For instance, Harris et al. (1996) found that children's judgments were influenced by the availability of alternatives. Three- and 4-year-olds heard vignettes involving minor mishaps in which the character made a choice between two options that would lead to the same outcome or two options that would lead to different outcomes. For example, children heard a story about a character who decides between either a black pen and blue pen *or* a black pen and pencil to make a drawing. In both cases, the character chooses the black pen and gets ink all over her hands. Children answered questions about why the outcome happened (e.g., "Why did her fingers get all inky?") and how it could have been prevented (e.g., "What should she have done instead so her fingers wouldn't get all inky?"). Children answered differently depending on the alternative option (i.e., another pen or a pencil), identifying the character's choice of the black pen in response to causal questions and the rejected option in response to prevention questions more often when the alternative was the pencil than when it was the blue pen. These results suggest that children considered counterfactual alternatives when arriving at causal judgments.

However, German (1999) argued that Harris et al.'s (1996) results did not conclusively show that counterfactuals feed into children's causal judgments, and may instead have been due to the interleaving of causal and prevention questions, and the focus of the stories on negative events. This argument fits with findings from the adult literature showing that counterfactual thinking is triggered more by negative events (Roese & Olson, 1997). Indeed, children's responses to causal questions (before answering any prevention questions) in Harris et al.'s (1996) study generally did not reference counterfactual alternatives. To further examine the role of counterfactual reasoning in children's causal reasoning, German (1999) presented 5-year-olds with stories in which a character made a choice (e.g., between a cardigan and a woolly jacket) which resulted in a negative (e.g., chooses cardigan and is cold) or positive outcome (e.g., chooses jacket and is warm). As in prior findings, children referenced rejected alternatives for the negative outcomes. However, when considering *positive* outcomes, 5-year-olds in German's study rarely referenced the character's choice in undoing the outcome. These results cast doubt on the view that counterfactual thinking is involved when reaching causal judgments.

Another study found some correspondence between preschoolers' ability to answer causal explanatory questions and counterfactual questions about *different* scenarios (Sobel, 2004). Their responses to explanatory questions predicted their performance on a counterfactual task, though because these two measures were based on different vignettes, it is difficult to draw conclusions about the symmetry between children's causal and counterfactual judgments.

In studies with adults, researchers have investigated the relation between the contents of counterfactual and causal thoughts, generally finding that adults attribute the cause of an event to a *strong cause*, but invoke an *enabling condition* when generating a counterfactual or when thinking about how the event could have been prevented (Byrne, 2007; Mandel & Lehman, 1996; McEleney & Byrne, 2006). Consider the following scenario from a classic study by Kahneman and Tversky (1982). Mr. Jones is driving home from work and decides to take a different route than usual. On his drive, a drunk driver runs a red light and hits Mr. Jones's car. When asked *why* the accident occurred, adults reference the drunk driver. When asked to generate a counterfactual, however, adults reference Mr. Jones's decision to take an unusual route home ("if only he had taken his normal route…"). Several studies have found that the contents of adults' causal and counterfactual thoughts diverge (Mandel & Lehman, 1996; McEleney & Byrne, 2006; N'gbala & Branscombe, 1995, but see Spellman & Ndiaye, 2007; Wells & Gavanski, 1989). Typically, counterfactual and prevention judgments align with enabling conditions that are *controllable* compared to causal judgments that are *uncontrollable*, either because they are outside the focal character's control (e.g., another driver) or outside anyone's control (e.g., weather) (Mandel & Lehman, 1996).

We found a similar pattern of results in a recent study with preschool- and school-age children (aged 3.5–8 years). Children were shown a series of short vignettes in which a minor mishap occurred. In each case, a character's action *enabled* a force of nature to cause the mishap. For instance, a character leaves a window open and the rain comes in and damages her doll. When asked *why* the outcome occurred (causal question), children most often responded that the force of nature was the cause (e.g., "The rain ruined it."). When asked how the outcome could have been prevented, children most often answered that the character should have acted differently (e.g., "She shouldn't have left the window open."). Interestingly, approximately 30% of preschool-age children answered by undoing the force of nature (e.g., "It shouldn't have rained), though over all age groups, children's causal and counterfactual judgments tended to diverge, in line with the results in several past studies with adults (e.g., Mandel & Lehman, 1996).

These findings on the symmetry (or lack thereof) between causal and counterfactual judgments have been taken as evidence to resolve debates about the nature of causal thought. This lack of symmetry has led some researchers to suggest that counterfactual judgments do not influence causal ones (Byrne, 2007; Mandel & Lehman, 1996; McEleney & Byrne, 2006; N'gbala & Branscombe, 1995), and therefore a *psychological process* view must not be correct. Others have argued that this need not be the case. For instance, Woodward (2011) cautions that just because the counterfactual that is most readily brought to mind when participants are asked certain questions does not align with their causal judgments, this does not mean that other counterfactuals have not or will not enter their causal ascriptions. Proponents of counterfactual theories of causation have also argued that the counterfactuals that act as input to causal reasoning need not be past-focused, and could instead be future or timeless conditionals (e.g., Woodward, 2011).

In support of a *psychological relatedness view,* or counterfactual theories of causation, Schulz et al. (2007) found that preschoolers used information from interventions to infer the causal structure of a gear toy (e.g., common cause, causal chain). Children were shown a set of two gears that spun at the same time, but each gear could be removed to investigate the effect of this removal on the remaining gear. For instance, if Gear A caused Gear B to spin (causal chain, A→B), then removing Gear A would cause Gear B to stop spinning but removing Gear B would not cause Gear A to stop spinning. However, if the two gears were each under the control of a common cause (e.g., a switch), then removing either gear should not affect the other. Children also used their knowledge of causal structure, shown to them by the experimenter with a picture, to make *conditional inferences* about hypothetical interventions to the toy ("If I put this gear down and turn on the switch, will the gear spin or will the gear stay still?"). Note that the wording of this question would meet broader definitions, but not narrower definitions of counterfactual reasoning. Regardless, this study suggests a symmetry between children's causal and hypothetical inferences.

More recent studies have not found this same symmetry. For instance, 5- to 7-year-olds children in Frosch, McCormack, Lagnado, and Burns' (2012) study inferred different causal models based on the temporal sequence of a series of spinning objects on a board (see also Burns & McCormack, 2009). However, their counterfactual inferences did not comply with these different causal models. McCormack et al. (2011) suggested this discrepancy with Schulz et al.'s (2007) findings may have to do with how children's representations of causal structure were derived. In Schulz et al.'s (2007) study with preschoolers, children were *told* the correct causal model, whereas in Frosch et al.'s (2012) study, children had to infer causal structure from the temporal relation between events, which may not support reasoning about counterfactual interventions.

However, in our own studies with preschoolers using a similar design, we found that children's counterfactual inferences did not respect causal structure even when they were both shown *and* told the causal structure of different physical apparatuses (Nyhout & Ganea, 2021a). Across two studies ($n=144$), children learned about different causal models (e.g., a red, blue, and green windmill) that were related to one another in either a causal chain (red → blue → green), common cause (blue ← red → green), or common effect (overdetermined) model (red → green ← blue). The stimuli were displayed on a screen, and each causal model had accompanying narration (e.g., "The blue one makes the green one start spinning.") and each had a distinct spatial layout and temporal signature. Children received one trial for each model (causal chain, common cause, common effect) crossed with three different physical systems (spinning windmills, bouncing balls on platforms, flashing lights) that were connected within each system by electric cords. Children were then asked to operate each system themselves by use of a touch screen. Following this, children were asked to describe the physical system to the experimenter (causal question), and were corrected if they described it incorrectly. Children were then asked two counterfactual (Studies 1 and 2; "If someone had taken the red one away, would the green one still have started spinning?") or future hypothetical questions (Study 2; "If someone takes the red one away, will the green one still start spinning?"). Children's responses to the counterfactual and hypothetical questions did not surpass chance levels on any questions except those for the *common effect* model. This is surprising in light of the difficulties children have had reasoning about overdetermined models in some past studies (e.g., Nyhout, Henke, & Ganea, 2019; Nyhout, Iannuzziello, et al., 2019; Rafetseder et al., 2013). However, children may have succeeded on these trials because they were able to "rewind" to a previous state of the world in which the counterfactual proposition was true (see also Gautam, Suddendorf, Henry, & Redshaw, 2019). To display the causal effect model, we showed children each individual cause operating on its own (e.g., red making green spin, then blue making green spin) before showing them together. Therefore, children did not have to generate a novel counterfactual representation but could access one from recent memory. Overall, these more recent findings cast doubt on the claim that young children can make counterfactual inferences that respect different causal models.

In the following section, we will set aside this debate and consider how counterfactual reasoning may influence both the process and the product of scientific inquiry. One need not assume that counterfactual reasoning is a necessary component of causal or scientific reasoning in order to recognize that it may nevertheless influence the inferences one draws (e.g., Byrne, 2007).

5. Counterfactual thought experiments

The merits of the "child as scientist" metaphor are the subject of substantial debate. Do children engage in inquiry that is qualitatively the same as that of mature scientists, just on a smaller scale? One place where this metaphor may hold especially true is when it comes to *counterfactual thought experiments*.[b] Although several studies have documented limitations in children's scientific reasoning (e.g., Zimmerman, 2007), children's counterfactual reasoning appears to respect many of the principles that children eschew during experimentation in the real world. Indeed, in our own work, we have found that first thinking counterfactually can make otherwise-flawed experimenters more optimal. In this section, we discuss the educational potential of counterfactual thought experiments. We propose that counterfactual reasoning can influence both the inquiry *process* and the *product* of scientific inquiry (i.e., concept learning/belief revision).

[b] For recent discussions of a broader category of thought experiments in learning, see Bascandziev and Harris (2020) and Harris (2021).

5.1 Counterfactuals and the inquiry process

Why might one expect counterfactual thinking to affect the process of scientific inquiry? Several researchers have highlighted structural commonalities between counterfactuals and scientific experiments (e.g., Gopnik, 2009; Gopnik & Walker, 2013; Hagmayer, Sloman, Lagnado, & Waldmann, 2007; Rafetseder & Perner, 2014; Sloman, 2005). When reasoning counterfactually, one takes a representation of an event or causal model, stipulates a change (e.g., by subtracting a cause), and follows the causal implications of the change. Likewise, when conducting a controlled experiment, one typically isolates a single variable to explore its influence on an outcome of interest. In this way, a real event and its counterfactual counterpart are structurally analogous to a treatment and control condition. Indeed, in the parlance of experimental design, the control condition is sometimes referred to as a counterfactual.

Given the structural commonalities between counterfactuals and controlled experiments, one may expect counterfactuals to "prime" valid experimentation. Evidence from our research supports this view.

Children aged 7–10 years were first pre-tested on their ability to execute the control-of-variables strategy using a ramp apparatus (Nyhout, Iannuzziello, et al., 2019). An experimenter asked children to show her how they would find out whether a variable (e.g., ramp surface) mattered for how far a ball travels down the ramp. Two variables were available at a time, and to conduct a valid test, children had to vary the variable of interest, while holding the other constant. Only those who failed pre-test took part in subsequent phases of the study. Fig. 1 provides an outline of the study design.

In the *scaffolding phase*, children then watched a video of an actor setting up the ramps to test the two variables children had used at pre-test (e.g., ramp surface and run length). The video experimenter did not explicitly comment on her strategy, but always conducted a controlled test. Following each of two tests the actor set up, the experimenter paused the video to ask children one of two prompts. Children in the *counterfactual* condition were asked what would have happened differently if the actor had conducted a confounded test (e.g., "Let's imagine that she set Ramp 1 to rough. Would the ball have travelled down the ramp farther on Ramp 1, farther on Ramp 2, or you can't be sure?" when the actor had set Ramps 1 and 2 both to smooth, and used a long run on one and a short run on the other). Children in the *control* condition were asked to recall what had happened (e.g., "Let's imagine again what happened to the ball on

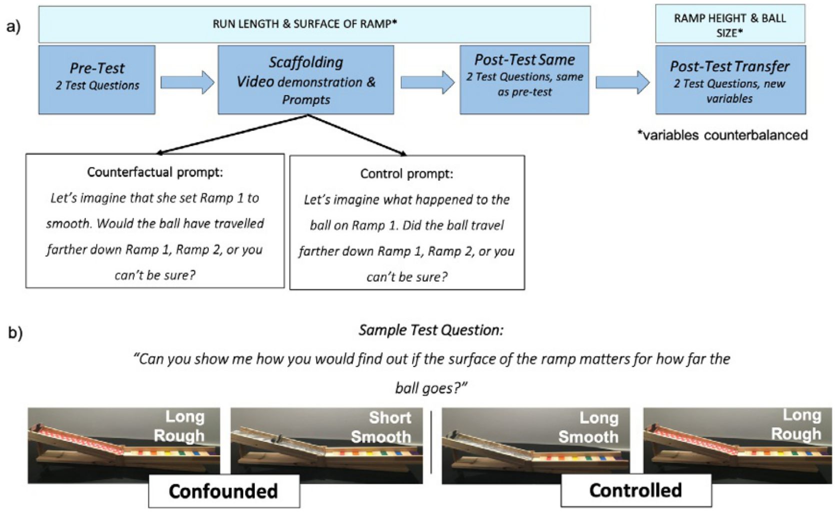

Fig. 1 Schematic of study design and materials used in Nyhout, Iannuzziello, et al. (2019). (A) Study Design: the study was divided into four phases. The same variables were used in the first three phases, whereas a new set of variables was introduced to children in the Post-test Transfer phase. The critical difference between conditions was the nature of the prompts used in the Scaffolding phase (Counterfactual or Control). (B) Study Materials and Test Question: Children were asked to set up two identical ramps that could be varied along two dimensions at a time, depicted here with the variables: ramp surface (smooth = aluminum foil, rough = bubble wrap) and run length (long or short). The set of ramps on the left depict a confounded test, whereas those on the right depict a controlled test.

Ramp 1? Did the ball travel farther on Ramp 1, farther on Ramp 2, or you can't be sure?"). "Can't be sure" was included as an option because the counterfactual scenario always created a confounded test, and therefore it was not possible to predict whether the ball would go farther on one or the other ramp.

Following the demonstration phase, children completed two post-test phases. In the first, *post-test same variables*, they were asked to conduct two tests using the ramps and the two variables (e.g., ramp surface and run length) they had been familiarized with during pre-test and demonstration. In a *post-test transfer* phase, the experimenter "fixed" the two variables that were previously used (e.g., by taking away the two surfaces) and introduced two new variables (e.g., ramp height, ball size). Children in both conditions improved from pre-test to post-test same variables. However, children in the counterfactual condition had the edge when conducting tests using new variables on *post-test transfer*. Whereas half of children in the counterfactual condition conducted two controlled tests, only a quarter of children in the control condition did so. These findings are surprising in light of the difficulty children have with conducting controlled tests (e.g., Klahr & Nigam, 2004; Schwichow, Croker, Zimmerman, Höffler, & Härtig, 2016).

Note that our manipulation was quite subtle; children in both conditions saw the same demonstration, and the only difference was the nature of the two prompts given to them by the experimenter. Interestingly, the answers children gave to the prompts were independent of their performance at post-test, suggesting the process of thinking counterfactually in itself may scaffold performance, regardless of the actual products of the thought process (Walker & Nyhout, 2020). Overall, these findings provide support for the assertion that counterfactuals, as imagined controlled experiments, may scaffold the learners' ability to conduct a controlled test in the real world.

An alternate, but not incompatible view is that considering counterfactuals may make the learner more open to alternative possibilities. In a series of experiments with adults, Galinsky and Moskowitz (2000) found that priming individuals with counterfactuals debiased reasoning on a range of tasks. Most relevant to the current proposal, however, was the finding that adults who had read a counterfactual-inducing scenario (e.g., about an individual who narrowly missed winning a large prize) were subsequently more likely to engage in disconfirmatory hypothesis-testing on an unrelated task than were those who read a control scenario. Participants in both conditions were asked to test a hypothesis that another person was an extrovert by sampling questionnaire items to ask the person. Those who had read the counterfactual-inducing scenario were more likely to select hypothesis-disconfirming items that tested an alternate hypothesis (e.g., that the individual was an introvert), whereas those in the control condition selected more hypothesis-confirming items. The authors argued that thinking counterfactually makes individuals more open-minded to alternative possibilities, causing them to adopt a "simulation mind-set" (Galinsky & Moskowitz, 2000).

Although the proposed mechanisms are different—imagined controlled experiments and a simulation mind-set—the two sets of findings we reviewed here suggest that counterfactuals ready the learner to engage in more optimal scientific inquiry by scaffolding the ability to conduct a controlled test and a disconfirmatory test of a hypothesis. In addition to conducting valid experiments, a core component of scientific inquiry is evidence evaluation. Does counterfactual reasoning similarly benefit evidence evaluation?

In a recent study, Engle and Walker (2021) investigated whether counterfactuals would influence children's selection between competing hypotheses on an evidence evaluation task. Five-year-olds saw an experimenter place blocks varying along two dimensions (color and size) on top of a box. Some blocks made the box light up, and others were inert. Block color covaried with the outcome 100% of the time, whereas block size co-varied with the outcome 75% of the time. After each observation, the experimenter asked children either to *explain* the outcome ("Why did this one make my toy light up?") or think about a *counterfactual alternative* ("What if this one had been this (picture of a different block)? Would my toy have lit up or not lit up?"). Previous research had established that children have a strong prior belief that size is more likely than color to be a causal variable on this task (Walker, Smallman, Summerville, & Deska, 2016). Therefore, this task pitted the evidence at hand with children's prior beliefs. On a test trial, children were asked to select between two blocks—one consistent with the 100% color observation, and one consistent with the 75% size observation (and their prior beliefs about causality). Children in the *counterfactual* group more often selected the block consistent with the color hypothesis, whereas those in the *explain* group more often selected the block consistent with the size hypothesis. These results suggest that getting children to think of counterfactual possibilities helps them to resist the pull of their prior beliefs and select a candidate hypothesis that is more consistent with the available evidence.

In another study, McCormack, Simms, McGourty, and Beckers (2013) investigated whether thinking counterfactually influences the *specific* causal inferences children make. An experimenter showed 5- to 7-year-old children a toy robot that lit up and made noise when given certain "foods" (causal blocks). The causal blocks had an additive effect, such that when the robot was given a single block, there was a small effect, but when given two foods, there was a larger effect (louder sound and brighter light). Other blocks—non-foods—had no causal effect. Across different trials, children observed the experimenter feeding the robot, and were asked either to *describe* what had happened (factual condition) or to *imagine* what would have happened if a non-causal block had been causal (counterfactual condition). In a test phase, 5-year-olds who answered counterfactual questions were more likely to make correct causal inferences than those who had answered factual questions, boosting their performance to a level similar to older children. If they observed a causal block (A) given to the robot along with a block of unknown causal status (B), they correctly reasoned that B was not causal if the additive causal effect was not observed. The authors reasoned that this conclusion is based on the blocking inference "if B had been causal, the effect would have been stronger," which itself is a counterfactual. Thus, the available evidence suggests that thinking counterfactually can optimize children's ability to select between candidate hypotheses for more general categories of events (causal reasoning; Engle & Walker, 2021) and specific events (causal selection; McCormack et al., 2013).

Thus, a growing body of research indicates that prompting children to think of counterfactual alternatives positively benefits their scientific inquiry. Children's performance on both experimentation (Nyhout, Henke, & Ganea, 2019; Nyhout, Iannuzziello, et al., 2019) and evidence evaluation tasks (Engle & Walker, 2021; McCormack et al., 2013), by way of counterfactuals, shows sophistication at a younger age than previous research has found.

If counterfactuals function as imagined experiments, then one may expect them not only to influence the inquiry process, but also the *products* of this process. In the following section, we consider evidence consistent with the view that engaging in counterfactual thought experiments can help the reasoner to update and enhance their understanding of the world.

5.2 Counterfactuals and scientific knowledge

> *I would think about these questions as I was walking … thinking about what it would be like to be in this situation with all this material collapsing around you and what would happen.*
>
> ***—Sir Roger Penrose, Novel Laureate in Physics (Science News, October 6, 2020)***

Roger Penrose, along with Reinhard Genzel and Andrea Ghez, were awarded the 2020 Nobel Prize in physics for their work illuminating the existence of a black hole at the center of the Milky Way. Their insights were the product of mathematical models, observations of stars' orbits, and, as the quote illustrates, counterfactuals. Counterfactual thought experiments have played a central role in several scientific advances (Brown, 2011).

Where some fields of inquiry are ripe for laboratory-based experimentation (e.g., botany), others are intangible, not only to the youngest learners, but also to professional scientists (e.g., astronomy). Scientists, both young and old, must therefore rely on inferential processes that allow them to tinker with their causal models of the world to reach new understanding (along with, perhaps, mathematical calculations, and piecemeal observations).

Can new understanding and learning result from counterfactual thought experiments in development? Our recent work, aimed at teaching children principles of astronomy, suggests they may (Nyhout and Ganea, 2021b). Before considering this question, however, we should first ask whether children can think counterfactually about different types of causal systems that are relevant to scientific understanding. Most developmental studies, as reviewed in the section *Counterfactual reasoning in development* have used fictional narratives. Those that have used non-narrative causal systems have produced mixed results (e.g., Frosch et al., 2012; Nyhout & Ganea, 2021a; Schulz et al., 2007). In a recent study, we tested children's ability to reason about counterfactual interventions to complex and novel biological systems.

In Nyhout et al. (2021), we showed 5-, 6- and 7-year-old children ($n = 144$) novel food chains involving three animals, and then asked children about the effects of the removal of one animal on others in the ecosystem (e.g., "Let's imagine there were no more palas in the pond. If there were no more palas, would there be more, less, or the same amount of mingos?"). Five-year-olds' performance was chance-like on all questions. Six- and seven-year-olds were better able to reason about the effects of one species' removal on the others but showed more success when reasoning about a species' *direct* predator or prey, rather than an indirectly connected one (e.g., its predator's predator). These results indicate that by 6 years of age, children can engage in counterfactual thought experiments about novel 3-part causal chains. Do such thought experiments contribute to knowledge and transfer?

In Nyhout and Ganea (2021b), we sought to teach 6- and 7-year-olds ($n = 109$) the principle of *planetary habitability*. Using a short, illustrated tutorial, we taught children the causal chain underlying this concept: distance of a planet from its star affects temperature; temperature affects the state of water; water affects opportunity for life. Children have not formally learned this concept in school but possess many of the relevant building blocks (e.g., water is required for life). Following this, we asked children either to think *counterfactually* about Earth being closer to or farther from the Sun, or to think about *examples* of different planets (Venus and Neptune). We then tested children's *comprehension* of the concept, asking why Earth was a planet we could live on. Both groups performed equally well in this comprehension phase, and outperformed a control group who had not been exposed to the concept. Next, we tested children's *transfer*, asking them to predict and explain whether novel planets in a pretend planetary system could support life. Children who had thought counterfactually about Earth significantly outperformed children who thought of examples, who surprisingly did not outperform children in the control group. These results suggest that counterfactual thought experiments (e.g., imagining Earth closer to the Sun) support abstract, generalizable understanding of a concept—much in the same way real-life experiments do.

In sum, we have reviewed arguments that counterfactual thinking is structurally analogous to real-world experiments (e.g., Gopnik, 2009; Gopnik & Walker, 2013; Hagmayer et al., 2005; Rafetseder & Perner, 2014; Sloman, 2005). We have provided evidence that thinking counterfactually positively contributes to scientific inquiry. Its engagement leads to controlled, disconfirmatory hypothesis-testing (Galinsky & Moskowitz, 2000; Nyhout, Henke, & Ganea, 2019; Nyhout, Iannuzziello, et al., 2019), and more ideal evidence evaluation (Engle & Walker, 2021; McCormack et al., 2013). Counterfactual thought experiments may also enhance learning and transfer of science concepts (Nyhout & Ganea, 2021b). Because this area of research is in its infancy, several open questions remain, including whether the exact nature of the counterfactual prompt has an influence and how long-lasting any effects may be. We turn to these considerations next, along with a discussion of the educational potential of counterfactual thought experiments.

6. Future directions for counterfactual thought experiments in research and education

What would the world be like with no bees? Do we need the moon? How has human action driven climate change? When asking scientific questions, counterfactual considerations often naturally arise. Indeed, a quick web search will turn-up several educational books and films dedicated to the counterfactual of what the world would be like with no bees. An open question is whether considering these counterfactual alternatives changes beliefs and behavior in meaningful ways. Better understanding the types of thought experiments children can consider and how they impact learning will be an important focus for future research.

As we highlighted in Section *Counterfactual reasoning in development*, counterfactual reasoning is not an all-or-none ability. With development, children are capable of reasoning counterfactually about increasingly complex causal models. Where the youngest preschooler may only be able to reason about a counterfactual alternative to a single cause-effect relation, a school-age child can reason about longer causal chains. This, of course, will influence the types of counterfactual thought experiments researchers and educators can ask children to consider. Children's domain-specific causal knowledge will also contribute to the types of thought experiments they can reason through (Sobel, 2011).

In addition to the issue of *complexity*, there are several other relevant features of counterfactual thought experiments to consider. We know from past research that across individuals and situations the mind tends to center on a small set of the infinitely many counterfactual possibilities (e.g., Byrne, 2007). How does the specific counterfactual alternative the child is asked to think about influence their learning and reasoning? In many developmental studies using fictional narratives, counterfactual prompts highlight close variations on reality. In Nyhout, Henke, and Ganea (2019), Nyhout, Iannuzziello, et al. (2019), we asked children to consider a minor change to how the actor set up her ramp experiment. In other studies, however, we have asked for more drastic changes. For instance, in Nyhout and Ganea (2021b), we asked children to consider what Earth would be like if it were very close to or very far from the Sun. Thus, past studies suggest that counterfactuals benefit inquiry and learning both when they are minor and more drastic variations on reality.

In all the developmental studies we have reviewed, however, the counterfactual prompts have pertained to the main casual relations of interest. What happens when the counterfactual is unrelated to the target outcome (e.g., asking children to consider a counterfactual about the color of the walls in the room)? To our knowledge, developmental studies have not focused on this question. Perhaps, as in Galinsky and Moskowitz's (2000) study with adults, *any* counterfactual serves to debias reasoning by opening the mind to alternative possibilities.

Moreover, we do not know how long-lasting the effects of counterfactual prompts will be. In all studies to our knowledge, test phases have immediately followed counterfactual prompts, and therefore it is unknown how short-lived such effects may be. Future work is needed to understand the extent and duration of the effects of counterfactuals. Doing so will also help to elucidate the mechanisms underlying any such effects.

Another important area of future study will be to investigate the use of counterfactual prompts outside the laboratory, in children's homes, classrooms, and museums. Open questions remain over how counterfactuals compare to other types of pedagogical prompts, and whether parents and educators spontaneously prompt children to consider counterfactual alternatives.

7. Conclusions

We have reviewed findings that children often struggle with aspects of scientific reasoning and inquiry. Their scientific misconceptions can cause them to overlook incompatible evidence that would help them to arrive at a more accurate understanding of the world. When asked to test hypotheses, they often conduct invalid tests by confounding variables or seeking only confirmatory evidence. We have argued that these limitations may be addressed using counterfactual thought experiments. The development of counterfactual reasoning is an area of substantial debate, with researchers disagreeing over its basic definition and when it develops. We argue that counterfactual thinking develops in the preschool years and children become able to think counterfactually about increasingly complex information with development. We reviewed theoretical arguments that counterfactuals are structurally analogous to real-world experiments, and presented evidence that counterfactual reasoning can benefit both the inquiry process and the outcomes of this process. Many open questions remain over when and why counterfactuals are useful, and how the possibilities one considers play a role.

References

Bascandziev, I., & Harris, P. L. (2010). The role of testimony in young children's solution of a gravity-driven invisible displacement task. *Cognitive Development*, *25*, 233–246.

Bascandziev, I., & Harris, P. L. (2020). Can children benefit from thought experiments? In A. Levy, & P. Godfrey-Smith (Eds.), *The scientific imagination* (pp. 262–279). New York, NY: Oxford University Press.

Beck, S. R. (2016). Why what is counterfactual really matters: A response to Weisberg and Gopnik. *Cognitive Science*, *40*, 253–256.

Beck, S. R., & Riggs, K. J. (2014). Developing thoughts about what might have been. *Child Development Perspectives*, *8*(3), 175–179.

Beck, S. R., Riggs, K. J., & Gorniak, S. L. (2010). The effect of causal chain length on counterfactual conditional reasoning. *British Journal of Developmental Psychology*, *28*(3), 505–521.

Beck, S. R., Robinson, E. J., Carroll, D. J., & Apperly, I. A. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development*, *77*(2), 413–426.

Brown, J. R. (2011). *The laboratory of the mind: Thought experiments in the natural sciences*. London: Routledge.

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

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). New York: Academic Press.
Burns, P., & McCormack, T. (2009). Temporal information and children's and adults' causal inferences. *Thinking & Reasoning*, *15*(2), 167–196. https://doi.org/10.1080/13546780902743609.
Byrne, R. M. (2007). *The rational imagination: How people create alternatives to reality*. Cambridge: MIT 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*, *63*, 1–49.
Coley, J. D., & Tanner, K. D. (2012). Common origins of diverse misconceptions: Cognitive principles and the development of biology thinking. *CBE—Life Sciences Education*, *11*, 209–215.
Dias, M. D. G., & Harris, P. L. (1988). The effect of make-believe play on deductive reasoning. *British Journal of Developmental Psychology*, *6*, 207–221.
Edgington, D. (2011). Causation first: Why causation is prior to counterfactuals. In C. Hoerl, T. McCormack, & S. R. Beck (Eds.), *Understanding counterfactuals, understanding causation: Issues in philosophy and psychology*. Oxford, UK: Oxford University Press.
Engle, J., & Walker, C. M. (2021). Thinking counterfactually supports children's evidence evaluation in causal learning. *Child Development*.
Frosch, C. A., McCormack, T., Lagnado, D. A., & Burns, P. (2012). Are causal structure and intervention judgments inextricably linked? A developmental study. *Cognitive Science*, *36*, 261–285.
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*, 384–409.
Ganea, P. A., Larsen, N. E., & Venkadasalam, V. P. (2021). The role of alternative theories and anomalous evidence in children's scientific belief revision. *Child Development*..
Gautam, S., Suddendorf, T., Henry, J. D., & Redshaw, J. (2019). A taxonomy of mental time travel and counterfactual thought: Insights from cognitive development. *Behavioural Brain Research*, *374*, 112108.
Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition*, *38*, 213–244.
German, T. P. (1999). Children's causal reasoning: Counterfactual thinking occurs for 'negative' outcomes only. *Developmental Science*, *2*, 442–457.
German, T. P., & Nichols, S. (2003). Children's counterfactual inferences about long and short causal chains. *Developmental Science*, *6*, 514–523.
Gopnik, A. (2009). *The philosophical baby: What children's minds tell us about truth, love & the meaning of life*. Random House.
Gopnik, A., & Meltzoff, A. N. (1998). *Words, thoughts, and theories*. Cambridge: MIT Press.
Gopnik, A., & Schulz, L. (Eds.). (2007). *Causal learning: Psychology, philosophy, and computation*. Oxford: Oxford University Press.
Gopnik, A., & Walker, C. M. (2013). Considering counterfactuals: The relationship between causal learning and pretend play. *American Journal of Play*, *6*, 15–28.
Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin*, *138*, 1085.
Guajardo, N. R., & Turley-Ames, K. J. (2004). Preschoolers' generation of different types of counterfactual statements and theory of mind understanding. *Cognitive Development*, *19*(1), 53–80.
Hagmayer, Y., Sloman, S. A., Lagnado, D. A., & Waldmann, M. R. (2007). Causal reasoning through intervention. In A. Gopnik, & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 86–100). Oxford University Press.

Hardy, I., Jonen, A., Möller, K., & Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students' understanding of "floating and sinking". *Journal of Educational Psychology*, *98*(2), 307.

Harris, P. L. (2021). Early constraints on the imagination: The realism of young children. *Child Development* (Online Version of Record).

Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, *61*(3), 233–259.

Hast, M., & Howe, C. (2012). Understanding the beliefs informing children's commonsense theories of motion: The role of everyday object variables in dynamic event predictions. *Research in Science & Technological Education*, *30*(1), 3–15.

Hood, B. M. (1995). Gravity rules for 2-to 4-year olds? *Cognitive Development*, *10*(4), 577–598.

Joh, A. S., Jaswal, V. K., & Keen, R. (2011). Imagining a way out of the gravity bias: Preschoolers can visualize the solution to a spatial problem. *Child Development*, *82*(3), 744–750.

Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American*, *246*(1), 160–173.

Kavanagh, C., & Sneider, C. (2007). Learning about gravity I. Free fall: A guide for teachers and curriculum developers. *Astronomy Education Review*, *5*(2), 21–52.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, *15*(10), 661–667.

Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. MIT Press.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, *96*(4), 674.

Kuhn, D., Garcia-Mila, M., Zohar, A., Andersen, C., White, S. H., Klahr, D., et al. (1995). Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development*. i-157.

Lewis, D. (1973). Counterfactuals and comparative possibility. In *Ifs* (pp. 57–85). Dordrecht: Springer.

Mackie, J. L., & Mackie, J. L. (1974). *The cement of the universe: A study of causation*. Oxford: Clarendon Press.

Mandel, D. R., & Lehman, D. R. (1996). Counterfactual thinking and ascriptions of cause and preventability. *Journal of Personality and Social Psychology*, *71*(3), 450.

McCloskey, M. (1983). Intuitive physics. *Scientific American*, *248*(4), 122–131.

McCormack, T., Frosch, C., & Burns, P. (2011). Children's causal and counterfactual judgements. In C. Hoerl, T. McCormack, & S. R. Beck (Eds.), *Understanding counterfactuals, understanding causation: Issues in philosophy and psychology* (p. 54). Oxford, UK: Oxford University Press.

McCormack, T., Ho, M., Gribben, C., O'Connor, E., & Hoerl, C. (2018). The development of counterfactual reasoning about doubly-determined events. *Cognitive Development*, *45*, 1–9.

McCormack, T., Simms, V., McGourty, J., & Beckers, T. (2013). Blocking in children's causal learning depends on working memory and reasoning abilities. *Journal of Experimental Child Psychology*, *115*(3), 562–569.

McEleney, A., & Byrne, R. M. (2006). Spontaneous counterfactual thoughts and causal explanations. *Thinking & Reasoning*, *12*(2), 235–255.

National Research Council. (2000). *How people learn: Brain, mind, experience, and school: Expanded edition*. National Academies Press.

N'gbala, A., & Branscombe, N. R. (1995). Mental simulation and causal attribution: When simulating an event does not affect fault assignment. *Journal of Experimental Social Psychology*, *31*(2), 139–162.

Nyhout, A., & Ganea, P. A. (2019a). Mature counterfactual reasoning in 4-and 5-year-olds. *Cognition*, *183*, 57–66.

Nyhout, A., & Ganea, P. A. (2019b). The development of the counterfactual imagination. *Child Development Perspectives*, *13*(4), 254–259.
Nyhout, A., & Ganea, P. A. (2021a). *Children's reasoning about counterfactual and future hypothetical interventions to different causal structures*. (in preparation).
Nyhout, A., & Ganea, P. A. (2021b). *Learning science concepts through examples and counterfactuals*. (under revision).
Nyhout, A., Henke, L., & Ganea, P. A. (2019). Children's counterfactual reasoning about causally overdetermined events. *Child Development*, *90*(2), 610–622.
Nyhout, A., Iannuzziello, A., Walker, C. M., & Ganea, P. A. (2019). Thinking counterfactually supports children's ability to conduct a controlled test of a hypothesis. In A. Goel, C. Seifert, & C. Freska (Eds.), *Proceedings of the 41st annual meeting of the cognitive science society*. Cognitive Science Society: Montreal, Canada.
Nyhout, A., Sweatman, H., & Ganea, P. A. (2021). *Children's reasoning about hypothetical interventions to complex and dynamic systems*. (under revision).
Palmquist, C. M., Keen, R., & Jaswal, V. K. (2018). Visualization instructions enhance preschoolers' spatial problem-solving. *British Journal of Developmental Psychology*, *36*(1), 37–46.
Pearl, J. (2000). *Causality*. New York: Oxford University Press.
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.
Perner, J. (1991). *Understanding the representational mind*. Cambridge: MIT Press.
Perner, J., Sprung, M., & Steinkogler, B. (2004). Counterfactual conditionals and false belief: A developmental dissociation. *Cognitive Development*, *19*(2), 179–201.
Pine, K., Messer, D., & St. John, K. (2001). Children's misconceptions in primary science: A survey of teachers' views. *Research in Science & Technological Education*, *19*(1), 79–96.
Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, *66*(2), 211–227.
Potvin, P., & Cyr, G. (2017). Toward a durable prevalence of scientific conceptions: Tracking the effects of two interfering misconceptions about buoyancy from preschoolers to science teachers. *Journal of Research in Science Teaching*, *54*(9), 1121–1142.
Rafetseder, E., Cristi-Vargas, R., & Perner, J. (2010). Counterfactual reasoning: Developing a sense of "nearest possible world". *Child Development*, *81*(1), 376–389.
Rafetseder, E., & Perner, J. (2014). Counterfactual reasoning: Sharpening conceptual distinctions in developmental studies. *Child Development Perspectives*, *8*(1), 54–58.
Rafetseder, E., Schwitalla, M., & Perner, J. (2013). Counterfactual reasoning: From childhood to adulthood. *Journal of Experimental Child Psychology*, *114*(3), 389–404.
Renken, M. D., & Nunez, N. (2010). Evidence for improved conclusion accuracy after reading about rather than conducting a belief-inconsistent simple physics experiment. *Applied Cognitive Psychology*, *24*(6), 792–811.
Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development*, *13*(1), 73–90.
Robinson, E. J., & Beck, S. (2000). What is difficult about counterfactual reasoning? In P. Mitchell, & K. Riggs (Eds.), *Children's reasoning and the mind* (pp. 101–120). Hove: Psychology Press.
Roese, N. J., & Epstude, K. (2017). The functional theory of counterfactual thinking: New evidence, new challenges, new insights. In J. M. Olson (Ed.), *Advances in experimental social psychology: Vol. 56. Advances in experimental social psychology* (pp. 1–79). Elsevier Academic Press.

Roese, N. J., & Olson, J. M. (1997). Counterfactual thinking: The intersection of affect and function. *Advances in Experimental Social Psychology*, *29*, 1–59.
Schulz, L. E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science*, *10*(3), 322–332.
Schwichow, M., Croker, S., Zimmerman, C., Höffler, T., & Härtig, H. (2016). Teaching the control-of-variables strategy: A meta-analysis. *Developmental Review*, *39*, 37–63.
Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition*, *124*(2), 209–215.
Sloman, S. (2005). *Causal models: How people think about the world and its alternatives*. Oxford, UK: Oxford University Press.
Sobel, D. M. (2004). Exploring the coherence of young children's explanatory abilities: Evidence from generating counterfactuals. *British Journal of Developmental Psychology*, *22*(1), 37–58.
Sobel, D. M. (2011). Knowledge and children's reasoning about possibility. In C. Hoerl, T. McCormack, & S. R. Beck (Eds.), *Understanding counterfactuals, understanding causation: Issues in philosophy and psychology* (pp. 123–146). Oxford, UK: Oxford University Press.
Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*(4), 605.
Spellman, B. A., & Ndiaye, D. G. (2007). On the relation between counterfactual and causal reasoning. *Behavioral and Brain Sciences*, *30*(5–6), 466.
Tetlock, P. E., & Lebow, R. N. (2001). Poking counterfactual holes in covering laws: Cognitive styles and historical reasoning. *American Political Science Review*, *95*(4), 829–843.
Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, *24*(4), 535–585.
Walker, C. M., & Nyhout, A. (2020). Asking "why?" and "what if?": The influence of questions on children's inferences. In L. P. Butler, S. Ronfard, & K. H. Corriveau (Eds.), *The questioning child: Insights from psychology and education* (pp. 252–280). Cambridge, UK: Cambridge University Press.
Walker, R. J., Smallman, R., Summerville, A., & Deska, J. C. (2016). Motivated by us but not by them: Group membership influences the impact of counterfactual thinking on behavioral intentions. *Social Cognition*, *34*(4), 286–305.
Weisberg, D. S., & Gopnik, A. (2013). Pretense, counterfactuals, and Bayesian causal models: Why what is not real really matters. *Cognitive Science*, *37*(7), 1368–1381.
Weisberg, D. S., & Gopnik, A. (2016). Which counterfactuals matter? A response to Beck. *Cognitive Science*, *40*(1), 257–259.
Wells, G. L., & Gavanski, I. (1989). Mental simulation of causality. *Journal of Personality and Social Psychology*, *56*(2), 161–169. https://doi.org/10.1037/0022-3514.56.2.161.
Woodward, J. (2005). *Making things happen: A theory of causal explanation*. Oxford, UK: Oxford University Press.
Woodward, J. (2011). Psychological studies of causal and counterfactual reasoning. In C. Hoerl, T. McCormack, & S. R. Beck (Eds.), *Understanding counterfactuals, understanding causation: Issues in philosophy and psychology* (pp. 16–53). Oxford, UK: Oxford University Press.
Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, *27*(2), 172–223.

This page intentionally left blank