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

6 Contextualization in Causal Reasoning and Scientific Thinking

As discussed in chapter 4, one major difference between causal reasoning tasks and scientific thinking tasks is the extent to which these tasks contain real-world scientific content. Although we might be able to conceptualize a kind of "pure" scientific thinking ability that operates independently of any particular content, that is rarely if ever how reasoning proceeds in the real world. Typically, we use our existing knowledge of how the world works to form a hypothesis space in which to reason scientifically (Goodman et al., 2011; Griffiths et al., 2011; Sobel et al., 2004; Ullman et al., 2012; see also Cheng & Novick, 1992, for another version of this hypothesis as applied to adult cognition). We refer to this as the *contextualization* of a problem: the extent to which a problem requires (or seems to require) us to draw on our prior knowledge.

Importantly, contextualization lies on a continuum, such that none of the tasks that we have described sit at either extreme. Nevertheless, the two families of tasks that we have been considering do differ in important ways. Causal reasoning tasks, such as ones involving the blicket detector, are less contextualized—they require minimal prior knowledge about the pieces of the system (i.e., the blocks and the machine) in order to make inferences. The diagnostic reasoning task using the blicket detector that we described in chapter 5 is presented as such. However, this task (and most causal reasoning tasks in general) still relies on some degree of general knowledge about how causal systems in the world work, for example, that causes precede their effects or that the machine works on the basis of contact.

In contrast, scientific thinking tasks are more contextualized; they present information about particular physical or biological systems in the world, like earthquakes, gravity, friction, or disease transmission. Although they do

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

contain specific scientific information, such tasks do not capture the full richness and complexity of real-world causal systems, for which it is virtually impossible to construct a fully controlled experiment. Indeed, when we teach experimental design to undergraduate students, they often get lost when they begin to consider just how many variables they might have to control.

These examples illustrate that causal reasoning tasks and scientific thinking tasks differ starkly in their degree of contextualization. In this chapter, we review a set of studies designed to determine the extent to which contextualization matters to children's reasoning. One possibility is that the inclusion of real-world content helps children's reasoning abilities by instantiating an otherwise abstract system in a familiar context. Being able to tap into their existing knowledge could jump-start children's thinking about the causal system. But a different possibility is that the inclusion of real-world content hurts children's reasoning abilities by distracting them away from the underlying causal structure and toward irrelevant features of the scientific information. To explore these two possibilities, we first describe some prior work on the effects of contextualization on adults' and children's reasoning. Then, we review a series of studies designed to investigate directly how different levels and types of contextualization affect performance on the blicket detector task we described in chapter 5 (previously published in Weisberg, Choi & Sobel, 2020).

The Role of Contextualization in Adult Reasoning

Work in adult cognitive psychology finds that contextualization often matters for reasoning, particularly for causal or logical inferences. The trouble is that a problem's context sometimes facilitates people's reasoning and sometimes interferes with it (see e.g., Stanovich, 2004, for a review).

One example of how the inclusion of more real-world content can help adults' reasoning is the Wason card selection task (e.g., Wason, 1960, 1966, 1968). Adult participants in this task are shown a set of cards. They are told that each card has a letter on one side and a number on the other. For example, they can see four cards: one with the letter E, one with the letter F, one with the number 4 and one with the number 7. Participants are then told a rule: If there is a vowel on one side, then there is an even number on the

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

other side. Participants are asked to turn over all and only the cards that are necessary for proving this rule true or false.

This is a problem of logical reasoning. Participants are given an “if-then” statement (called a *conditional statement*), and they need to determine whether that statement is true. Importantly, conditional statements are only false in one situation—when the antecedent (the part of the statement that starts with “if”) is true, but the consequent (the part of the statement that starts with “then”) is false. So this rule is only false if there’s a card with a vowel on one side and an odd number on the other.

That is a big hint to figuring out which card or cards one has to turn over to verify the truth of the rule (although participants are not given that hint). To be sure that the conditional statement is true, one has to do two things. First, one has to turn over the vowel (E) to make sure that there is an even number on the other side. Most participants do this, and many then stop there, assuming that they have verified the rule. Those that do not stop tend to additionally turn over the even number (4), to make sure that there is a vowel on the other side.

But this is a logical error. Why? Because the rule only goes in one direction: If there is a vowel, then there’s an even number. The rule never says that even numbers have to have vowels on the other side; this is the difference between a conditional and a biconditional, or “if and only if.” In simple conditional reasoning, it does not matter whether there is a consonant on the back of the even number because the rule only specifies the relations for vowels.

The error that tends to occur on this task is that the majority of participants do not check the other part of the conditional; they do not turn over the card with the odd number (7). Turning this card over is necessary because one has to make sure that there is not a vowel on the other side. If there is a vowel on the other side, then the rule is obviously wrong, because the rule states that a vowel on one side means that there is an even number on the other, not the odd number 7.

This abstract—relatively decontextualized—version of the problem might seem counterintuitive and difficult to follow. But there is a much easier version of this problem, which presents the same logical inference to participants. Let’s substitute the letters and numbers on the cards for ages on one side and beverages on the other (following Griggs & Cox, 1982, Experiment

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

3). For example, say the cards show two ages (17 and 22) and two drinks (beer and soda). Now, consider the legal drinking age in the United States: If you are under the age of 21, then you cannot (legally) drink alcohol. Which cards should you turn over to verify that the four people (represented by the cards) are following the rule? Griggs and Cox gave participants a version of the task in this context, further emphasizing the rule by having participants imagine that they were police officers trying to enforce the law. The answer in this case should be obvious. You turn over the card with the 17 on it, because you want to know what the underage person is drinking. And you also turn over the beer card to make sure that there is someone 21 or older on the other side.

Note that the structure of the problem is the same as the one described above. Participants are presented with conditional statements that they have to validate, and four possible values for the antecedent and the consequent that represent the possible worlds that could exist in this conditional. In the less-contextualized case—the original Wason selection task with letters and numbers—participants were not great at answering these questions. In fact, none of the participants in Griggs and Cox's experiment who were given this version of the problem got it right (i.e., turned over the E and the 7). But in the more contextualized case—the one with the rule about the legal drinking age—participants were pretty good; 71% of Griggs and Cox's participants succeeded at turning over both cards in this version. There are a number of follow-ups to this task, all of which point to the idea that adult reasoning obeys certain pragmatic constraints or has certain modes of thinking based on our prior knowledge (e.g., Cheng & Holyoak, 1985; Johnson-Laird et al., 1972). Moreover, adults tend to search for explanations of less contextualized problems (or strangely contextualized ones) in order to make the conditional (or other causal construct) make sense to them (e.g., Korman & Malle, 2016).

These findings suggest that, if we can place a reasoning problem into a familiar context, then reasoning about that problem becomes easier. However, as noted above, there are also cases in which the presence of a familiar context makes a problem harder. For example, Kuhn (2007a) described an experiment where she presented a set of adults with the task of determining which combination of entertainment options (e.g., door prizes, comedian) would be most effective at boosting the success of a fundraiser. Although clear evidence demonstrated that some factors were causal (because they

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

were always related to the desired outcome of high fundraising) and some factors were not (because they occurred in all scenarios, regardless of outcome), participants had trouble judging which factors were related to the outcome. Further, participants used inconsistent logic in justifying their responses, seeming to be swayed by their preexisting knowledge of door prizes and comedians and ignoring the underlying logical structure of the task. She writes, "[the participants'] responses revealed that their judgments were in fact influenced by their own ideas about how effective these features ought to be" (p. 47).

Another example comes from the logical reasoning literature. Consider a simple (hypothetical) example of syllogistic reasoning. One can be told that all things that are smoked are good for one's health, and then that cigarettes are smoked. In this case, adults are less likely to draw the correct logical conclusion (in this example, not in real life, of course) that cigarettes are good for one's health. Because this conclusion is false in the real world, participants have a difficult time saying that it is correct, even though it is supported by the premises (Evans et al., 1983; see also Markovits & Nantel, 1989).

These lines of research on adults' logical reasoning show that context matters for adults' inferences. Sometimes adding more context facilitates reasoning, but in other cases it can impair adults' ability to come to a logical conclusion. What remains unclear is whether including a real-world context in a reasoning problem changes how children respond to reasoning tasks.

The Role of Contextualization in Development (or, What's Fantasy Got to Do with It?)

Research in developmental psychology has also looked at the question of how contextualization impacts learning. Interestingly, much of this work has compared children's reasoning within typical, everyday contexts to their performance when presented with fantastical situations. As a whole, these studies tend to find that taking a problem out of its everyday context and placing it into a fantasy context improves children's performance.

One example comes from work by Dias and Harris (1988, 1990). These researchers presented logical syllogisms problems to 4- to 6-year-olds that were contrary to fact, such as, "All cats bark. Rex is a cat. Does Rex bark?" Preschoolers were not that good at reasoning about these kinds of problems, although the older children were a bit better than the younger children.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

In these studies, children tended to answer on the basis of their real-world knowledge, saying that Rex would meow and not bark. (Note that this response tendency is similar to what adults did in the study about smoking, described above.)

Rather than concluding that young children could not reason about such syllogisms, Dias and Harris had the insight to put these reasoning problems into a pretense or a fantasy context. For instance, in one case, they introduced children to a game that was about “another planet” where everything is different. They then gave the children the same types of syllogisms, but with respect to this other planet: “On this planet, all cats bark. Rex is a cat. Does Rex bark?” Children’s performance in this study improved markedly under these circumstances; they were able to draw the correct (logical) conclusion even though it was contrary to their real-world knowledge (see also similar findings by Hawkins et al., 1984). This work suggests that children are capable of syllogistic reasoning but get tripped up by their real-world knowledge. They reason differently where there is less real-world content or where they are better able to understand that they need to suppress their real-world knowledge.

Research on children’s self-control points to a similar conclusion. White and Carlson (2016) showed that 5-year-olds’ performance on tasks that require a high degree of self-control improved when they adopted a third-person perspective on themselves or when they pretended to be an exemplary fictional character, like Batman (see also White et al., 2017). Arguably, because children in these conditions were able to feel more distant from themselves, or because they were able to take on the positive characteristics of the fictional character, they felt more able to control their impulses on the tasks. Four-year-olds also had an easier time understanding the role of mental states in pretending (e.g., someone has to know what X is in order to pretend to be X) if the pretend scenario involved fictional characters or fantastical actions (see Lillard & Sobel, 1999; Sobel & Lillard, 2001). We consider this particular case in more detail in chapter 10.

The same “fantasy advantage” effect can be seen in some cases when children learn from fictional stories. The vast majority of media aimed at children is fantastical, with a recent content analysis finding that 78.8% of popular children’s books, 97.8% of popular children’s movies, and 100% of popular children’s television shows contain at least one fantasy element (e.g., magic spells, anthropomorphized animals; Goldstein & Alperson, 2020; see also

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

Chlebuch et al., 2021; Taggart et al., 2019). Many of these pieces of media are also explicitly educational, designed to teach children new languages or cultural practices (e.g., *Dora the Explorer*), new scientific information (e.g., *Sid the Science Kid*), or new vocabulary (e.g., *Word Girl*). Given the findings reviewed above, it is possible that the fantastical contexts of these stories may facilitate children's learning.

There is some evidence for this possibility. One of our studies (Weisberg, Ilgaz, Hirsh-Pasek, Golinkoff, Nicolopoulou & Dickinson, 2015) taught new vocabulary words to preschoolers in Head Start programs using some books that had realistic themes (like farming) and some that had fantastical themes (like dragons). Children in this study tended to learn the words better when they were embedded in the books with the fantastical themes; they specifically performed better on a measure asking them to provide definitions for these words. Other work from our labs corroborates this finding, showing that preschoolers are more likely to transfer some types of new science content from fantastical stories than from realistic stories (Hopkins & Weisberg, 2021; Weisberg & Hopkins, 2020).

Despite cases where children learn and reason better in unrealistic contexts, it is important to stress that the vast majority of research on children's learning from fantasy has found better learning from contexts that are more realistic. The logic of this conclusion is simple: Learning involves transferring information from the original context where one first encountered it to different contexts where it might apply. This transfer is made vastly easier if a learner can see the similarities between the learning context and the application context (e.g., Daehler & Chen, 1993; Gentner, 1983; Holyoak et al., 1984). Realistic contexts, because they are more similar to reality, provide children with more support for this transfer process.

To take one example, preschoolers who heard about a new causal relation ("popple flowers cause hiccups") within the context of a realistic story were more likely to believe that this relation also held true in reality than children who heard about this causal relation within the context of a fantastical story (Walker et al., 2015). Similarly, a series of studies on children's transfer of problem solutions out of fictional stories shows a clear advantage for realistic stories (Richert et al., 2009; Richert & Smith, 2011). In these studies, children heard a story in which a character comes up with a solution for a problem, like hiding behind a robot (fantastical story) or a babysitter (realistic story) to avoid being seen. Children were more likely

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

to transfer this solution to an analogous real-world problem (figuring out where a doll should go to avoid being in a second doll's photograph of their room) if they had heard the realistic story. In these cases, where children are being asked to apply the rules of the story's world to reality, their transfer is facilitated by story contexts that more closely resemble the real world.

In general, the literature is mixed about whether (or under what circumstances) fantastical contexts can aid children's learning or other aspects of their cognitive functioning (see Hopkins & Weisberg, 2017, for a review). The majority of the evidence suggests that realistic contexts are better, particularly for learning, and possibly for various kinds of reasoning abilities as well. But studies like the ones reviewed here suggest that it can sometimes be beneficial to decrease the similarity between the reasoning problem and the real world because fantasy contexts—or other contexts that are more removed from children's everyday experiences—seem to benefit children's thinking under certain circumstances.

However, our primary goal in the current work is not to investigate children's thinking about fantasy, but about science (although, to misquote Arthur C. Clarke, fantasy is just science we don't understand yet[1]).

Context in Causal Reasoning and Scientific Thinking

The previous two sections illustrate the difficulty of drawing general conclusions about the role of contextualization in reasoning, especially in development. We aimed to investigate this issue directly by examining how scientific contexts might help or hinder children's reasoning.

In these studies, we conceptualized scientific contexts as being those aspects of a reasoning problem that reflect some aspect of the real world. Although these aspects could take many forms, for the studies on young children that are our primary concern in the current work, they tend look like the kind of content that would standardly be found in science classrooms or in museum exhibits aimed at children, such as earthquakes, volcanos, sinking and floating, or dinosaurs. This means that the children engaging in these tasks should understand that they are being asked to think about some aspect of the natural world.

In terms of whether these contexts help or hinder children's reasoning, as noted above, past research provides little guidance. In the case of fantastical

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

contexts, there is some suggestion such situations help children's thinking in certain cases, although more realistic situations have also been shown to be beneficial in others. Further, although some of those studies have tested aspects of children's reasoning, none have presented a complex diagnostic reasoning task like the one that we introduced in chapter 5. For such cases, it is possible that including some type of real-world science content could jump-start children's thinking about the causal system by tapping into their existing knowledge and providing a supportive context for their thinking. It might also lead children to explicitly realize that they are in a task that requires scientific thinking, which could itself boost their performance, as we argued in chapter 1 (and will investigate in more detail in chapter 8).

But it is also possible that such content could hinder children's thinking by introducing additional task demands, such as having to remember a set of unfamiliar or complex labels for categories or to access prior category knowledge, which could lead them to draw an erroneous conclusion. For example, to understand that different kinds of rocks could be potential causes for earthquake risk, one has to understand that the different labels (e.g., igneous, sedimentary) represent different categories of rocks, and to remember those categories as potentially having different properties, which in turn could lead to different causal relations. This might introduce task demands that are not inherent in less contextualized causal systems.

Our approach to addressing this issue began with the diagnostic reasoning measure described in chapter 5, which uses a blicket detector. We then began by taking the smallest possible step toward greater contextualization, keeping as much of the original task intact as possible.

Blickets to Butterflies

As described in chapter 5, the diagnostic reasoning measure we used involved putting combinations of four colored blocks (for example, yellow, black, blue, and orange) onto a machine. There were three possible outcomes: the machine turned green and played music, or it turned red and played a different piece of music, or it did not activate. This version of the task was relatively decontextualized, in that it presented no explicitly scientific content. Further, participants did not need to know anything about the system before starting the task in order to solve it correctly. Apart from a few

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

general assumptions about how causes and effects operate, as noted above, all the information required to answer the test question was provided.

As we shifted toward a greater degree of contextualization, we wanted to maintain as many aspects of this structure as possible, to test whether contextualization alone could affect children's performance. This meant that we had to choose a real-world causal system in which colors could plausibly be considered causal factors, to give the task some degree of realism. We also wanted to be sure that any real-world domain knowledge that children possessed would not automatically put them at a disadvantage in terms of solving the task (as the snake activity in the Earthquake Forecaster procedure might do).

We thus initially decided to focus our more contextualized version of this task on butterflies and flowers. Specifically, instead of discovering which combination of colored blocks would make the machine light up green, we asked children to discover which combination of colored flowers would bring a certain kind of butterfly to a field. This allowed us to introduce some real-world scientific content, specifically about a biological system in which color could plausibly play a role. Many of the children tested in this procedure were recruited at the Academy of Natural Sciences in Philadelphia, which houses a butterfly exhibit, making this context even more relevant to them. Otherwise, however, the task itself was exactly the same as the diagnostic reasoning task with the blicket detector, described in chapter 5.

To create the butterfly version of the task, the colored blocks were replaced by a set of four identical silk flowers: white, black, orange, and blue. We also constructed a "flower pot," which was made out of a rectangular block of Styrofoam. It had four holes punched in it, into which we could "plant" the flowers; this was the equivalent of the plastic container into which we placed the blocks in the blicket version of the task (see figure 6.1).

To illustrate the effects of flowers in the system, we used red and green plastic butterflies. We glued each set of butterflies onto wooden sticks, which were painted sky blue. To display these stimuli, we constructed a box out of foam board, which was rectangular with no top and with one side shorter than the other so that participants could see the butterflies peek out over the shorter side (see figure 6.2). We also recorded a musical cue to accompany the appearance of the green butterflies, just like the blicket detector played music when it lit up green. Finally, as in the blicket version

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

Figure 6.1
Flowers and field block used in the butterfly version of the diagnostic reasoning task.

of the task, we made laminated pictures of the combinations of flowers displayed in the task, along with green and red dots.

This apparatus allowed us to present the same causal structure as the blicket task: Two of the potential causes on their own led to one effect (the orange flower and the blue flower separately brought red butterflies), while their combination led to a different effect (the orange and the blue flower together brought green butterflies).

We presented this butterfly version of the task to 126 children in the same age range as we had previously tested with the blicket detector task: 4- through 10-year-olds (see full report in Weisberg, Choi & Sobel, 2020, Experiment 1). We found that, overall, participants who saw the butterfly task were 41% correct, which is marginally better than chance performance. Additionally, we observed roughly the same developmental progression as we had for the blicket version of the task, reported in chapter 5. Performance improved with age, but only by about 9 years old

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

did they choose the correct combination at significantly above-chance levels (see table 6.1).

In addition to asking our main test question, we asked children to justify their responses. We were interested primarily in whether children's justifications referred to some aspect of the system's causal structure that they had observed in the demonstration phase, implying that children understood that their task was to use that information to determine how this system worked. For example, "because in this picture [with all four

Figure 6.2
Butterfly apparatus used in the butterfly version of the diagnostic reasoning task.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

Table 6.1
Children's average percent correct on the blicket and butterfly tasks across age groups, as reported in Weisberg et al., 2020 (standard deviation in parentheses)

	Youngest third	Middle third	Oldest third
Blicket task	38.46 (49.29) $n=39$ mean age = 58.84 months	41.03 (49.83) $n=39$ mean age = 84.65 months	55.29 (50.39) $n=38$ mean age = 108.17 months
Butterfly task	26.19 (44.50) $n=42$ mean age = 58.76 months	38.10 (49.15) $n=42$ mean age = 80.44 months	59.52 (49.68) $n=42$ mean age = 108.44 months

flowers] the green butterflies came out and it has the orange and blue." We categorized these justifications as "data-based," because they referenced the data about the system that the child had seen. All other justifications were categorized as "other." These tended to refer to some irrelevant aspect of the task or to its surface features, for example, "because they're brighter" or "because these two are next to each other." We found that about 24% of children provided a data-based justification. Although we had initially thought that this kind of justification would indicate that children were thinking about the task at a more explicit, mature level, providing a data-based justification was not related to better performance on the task. Interestingly, this implies that children might be able to successfully solve the task without thinking it through explicitly, and only later come to an explicit understanding of their thought processes about the data they had observed.

Comparing Contexts

More important than responses to this task, though, is the question of how this performance compares to the blicket version: Do children perform better with the butterflies than the blickets, or worse, or about the same? To answer this question, we compared the children who had engaged with our butterfly task to a group of 116 children who engaged with the blicket version of the same task (we previously reported these data in chapter 5).

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

To make this comparison as close as possible, both groups of children were recruited from and tested at the same science museum. Interestingly, our comparisons across these two groups of subjects reveal no differences, either overall or when considering performance within each age group (see table 6.1 above; full analyses available in Weisberg et al., 2020, Experiment 1). This suggests that this contextualization, minimal though it may be, does not help or hinder children's reasoning abilities.

Interestingly, although there was no difference in performance, there was a difference in justifications: Children who saw the butterfly version of the task were somewhat more likely to justify their response with respect to the data that they had observed (24%) than children who saw the blicket version of the task (18%). This difference was not statistically significant. There was, however, an error in our data collection—only about a third of the children in the blicket version of the task were asked to provide justifications.

A follow-up study corrected this error. We replicated the entire procedure in a second sample of 103 children who each saw both the blicket version of the task and the butterfly version of the task, rather than just one or the other (order counterbalanced; full report in Weisberg et al., 2020, Experiment 2). This within-subjects design allowed us to test connections between the more and less contextualized versions of the task even more strongly. Here again, performance improved with age for both versions of the task, with only the oldest group (8- to 10-year-olds) performing significantly above chance: 71% for the blicket version and 59% for the butterfly version. Importantly, we again found no differences between children's performance on the two versions of the task. We also found a similar difference in the percentage of children providing data-based justifications in this version of the study (14% for the blicket version and 25% for the butterfly version).[2] This suggests that something about the more real-world scientific context presented in the butterfly condition may help children focus more on the causal structure of the task, or that the more abstract version of the task may hinder children's abilities to use the relevance of the evidence we provided in the demonstration phase of the task in their explanation.

In general, these results confirm the conclusion from the between-subjects version of the task: At least this minimal way of including real-world science

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

content in this causal structure neither benefited nor detracted from children's reasoning abilities. However, the real-world context may have helped children justify their responses in a more mature way.

Blicket-saurus

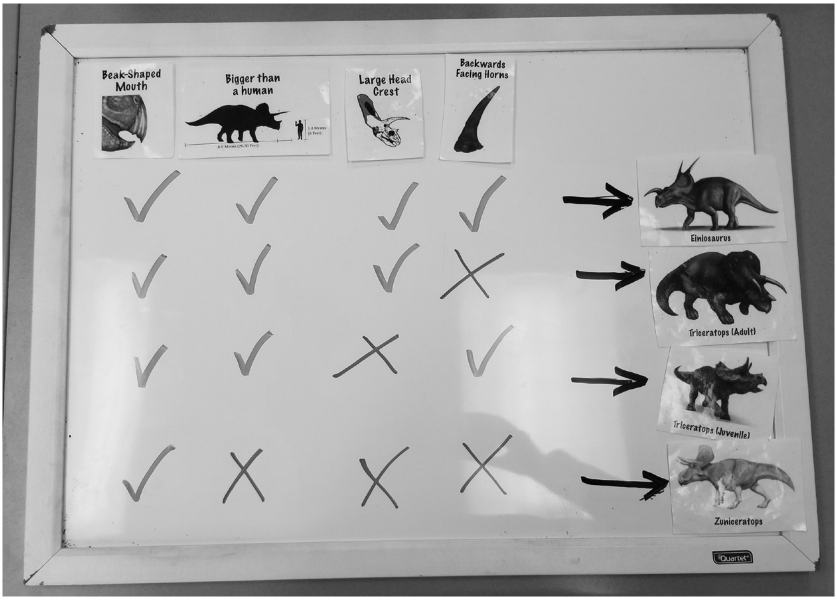
The butterfly version of the task presents one way to provide more real-world context to the same causal structure presented with the blicket detector. Although we found no differences in reasoning between these two versions of the task, it remains to be seen whether presenting children with a task that integrates even more real-world content would help or hinder their reasoning. To address that issue, we constructed a third version of the same task, this time focused on dinosaurs. Again, this context was chosen because of our partnership with the Academy of Natural Sciences in Philadelphia, which has an extensive dinosaur exhibit. This choice of context allowed us to take advantage of the fact that many children are interested in dinosaurs and even know quite a bit about them, which would test more strongly the role of prior knowledge in solving this task. Like the blicket and butterfly versions, though, no actual knowledge of dinosaurs was necessary for reaching the correct answer; thinking about the system's structure alone was enough.

This version of the task told children about different features that were characteristic of certain kinds of dinosaurs. Specifically, we selected dinosaurs from the family of ceratopsians (horned-headed dinosaurs). We told children about four different features that these dinosaurs could have or lack, in parallel to blocks being on or off the machine and to flowers being planted in the field or not: having a beak-shaped mouth (or not), being larger (or smaller) than a human when fully grown, having a large (or small) a head crest, and having backward-facing (or forward-facing) horns (see figure 6.3).

We first[3] introduced Einiosaurus, which has all four features. It has a beak-shaped mouth, is larger than a human when fully grown, has a large head crest, and has backward-facing horns. This dinosaur is thus the parallel of placing all four blocks on the machine or planting all four flowers in the field. The blocks or flowers are now dinosaur features, and the effect of combining all four of these variables (features) is that the dinosaur is an Einiosaurus.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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



	Beak-Shaped Mouth	Bigger than a human	Large Head Crest	Backwards Facing Horns
Einiosaurus	✓	✓	✓	✓
Triceratops (Adult)	✓	✓	✓	✗
Triceratops (Juvenile)	✓	✓	✗	✓
Zuniceratops	✓	✗	✗	✗

Figure 6.3
Whiteboard chart of the data presented to participants in the dinosaur version of the diagnostic reasoning task.

The next dinosaur has three of these four features. It has a beak-shaped mouth, is larger than a human when fully grown, and has a large head crest, but it does not have backward-facing horns; its horns face forward. This combination of features is an adult Triceratops. The next dinosaur has a different combination of three features: It has a beak-shaped mouth, is larger than a human when fully grown, and has backward-facing horns, but it has a small head crest. This combination of features is also a Triceratops, just a baby one. We chose Triceratops for this study because this species of dinosaur only grows a head crest in maturity, and its horns shift direction as it ages, from backward-facing to forward-facing (Horner & Goodwin, 2006). This allowed us to have two different combinations of causes (i.e., features) lead to the same effect (i.e., the same species of dinosaur), just as two different combinations of three blocks (or three flowers) both led to the effect of a red light (or red butterflies) in the other versions of this task.

Finally, we showed a dinosaur that had only one of the four features: It has a beak-shaped mouth, but is smaller than a human when fully grown, lacks a head crest, and has forward-facing horns. This dinosaur is a Zuniceratops.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

We presented these four dinosaurs and their combinations of features using a chart that we drew on a whiteboard (figure 6.3). As we described each feature, we stuck a picture of this feature on the title row of the chart. As we described each dinosaur, we put a green check mark or a red X in each column of the chart and stuck a picture of that dinosaur next to this row, so that children could see the full set of feature combinations by the end of our presentation.

Participants were then asked the test question: Which combination of two features defines an Einiosaurus? Specifically, we told them that paleontologists have found a new set of fossils that they know is an Einiosaurus. But they used only two clues to figure that out. Children were asked to choose which combination of the two clues the paleontologists had used: (a) being bigger than a human and having backward-facing horns, (b) being bigger than a human and having a large head crest, or (c) having backward-facing horns and having a large head crest. In parallel to the other two versions of this task, we consider the correct answer to be (c); only the Einiosaurus has this combination of two features.

We presented this version of the task to a new set of 110 participants between the ages of 4 and 10, as in the previous version of the task (Weisberg et al., 2020, Experiment 3). We found that children responded correctly on this version of the task only 27% of the time, which was not significantly different than chance. Unlike the blicket and butterfly tasks, there was no relation with age; even the oldest children did not respond differently from chance levels.

This comparison gives us a better picture of how introducing one kind of scientific contextualization affects children's thinking in this diagnostic reasoning task. The minimal contextualization of the butterfly task did not disrupt children's performance, while the richer and more realistic context presented in the blicket-saurus task did—even though the underlying causal structure of all three tasks was identical.

Granted, the dinosaur version of the task differed from the blicket and butterfly versions in a number of ways: It presented different features (e.g., size and shape) as relevant for causal structure, as opposed to different levels of a single feature (color). It included names of different and potentially novel dinosaurs. And it couched the task in terms of categorization rather than in terms of causal relations. These differences were introduced deliberately in order to present the same system in a way that more closely resembled

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

other scientific thinking tasks like Earthquake Forecaster. This combination of changes does not allow us to draw firm conclusions about which aspect of this contextualization may have affected children's performance, but it does clearly illustrate that the amount of prior knowledge that children must bring to bear (or feel that they must bring to bear) on constructing a hypothesis space for a problem impacts their reasoning abilities.

At least in the case of scientific contexts, then, the more real-world knowledge a task requires (or appears to require), the less children might be able to demonstrate their reasoning capacities. It thus does not necessarily make sense to draw general conclusions about the impact of "contextualized" and "decontextualized" systems. Contextualization occurs on a continuum, and the amount and type of contextualization appears to be critical. Further, the amount and type of contextualization might have different impacts in different contexts; there may be some domains where more contextualization helps children's reasoning and others where it hurts.

In turn, the results we have discussed in this chapter should encourage us to take a closer look at past work with older children and adults. It may be the case that some aspects of the realistic scientific contexts used in those tasks are indeed interfering with these participants' abilities to demonstrate their reasoning abilities. Or, put differently, it may be the case that both children and adults are capable of scientific thinking only in cases that bear little resemblance to real-world scientific contexts. If participants possess those reasoning skills, but are unable to demonstrate them in some tasks that have real-world contexts, then we must reconsider the extent to which we wish to conclude that young children possess genuine and sophisticated scientific thinking abilities.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

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

Constructing Science

Connecting Causal Reasoning to Scientific Thinking in Young Children

By: Deena Skolnick Weisberg, David M. Sobel

Citation:

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

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

The MIT Press

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024

MIT Press Direct

© 2022 Massachusetts Institute of Technology

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

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

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

Library of Congress Cataloging-in-Publication Data

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

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2253990/c004500_9780262370615.pdf by guest on 11 September 2024