8 Children's Definitions of "Science"

Part II (chapters 5 to 7) of this book described a series of case studies aimed at charting how children's causal reasoning abilities develop into more mature scientific thinking abilities, with particular focus on some of the variables that might affect this transition. Here in part III (chapters 8 to 10) we turn our attention to children's explicit understanding of concepts related to scientific thinking. Specifically, we ask how children's thinking about science, learning, and pretending (among other concepts) develops across the preschool and early elementary school years. Our goal in this work is to try to understand more about the gap between children's early abilities with causal reasoning, which tend to be implicit, and their later abilities with scientific thinking, which involve some degree of explicit reflection.

Intensions and Extensions

There is a long philosophical literature on the role of *intensions* in the field of philosophy of language (e.g., Kripke, 1972; Putnam, 1975). This literature mostly examines the relation between propositions described by linguistic utterances and their truth values. Roughly speaking, the intension of a word is the function that maps its usage to its meaning. Putnam (1975) called it the "something else" that represents the meaning of a word inherent to the concept itself, as opposed to the entities that are members of the category represented by the concept.

If that previous paragraph is confusing to you, don't worry; it is to us as well. Talking about intensions is confusing, and we suspect this is why Quine (quoted at the beginning of this part of the book) wanted them exorcised. From the point of view of the psychological literature, this term has mostly been used to describe the difference between our internal representations of the meanings of words (our concepts) and the external objects or actions or events that are denoted by those words (their referents). The idea, as we understand it, is that linguistic utterances have meaning beyond simply the objects or actions or events in the world that they refer to, because they are embedded in our psychological representations of these objects, actions, or events.

Understanding concepts—particularly in the philosophical literature—has often been characterized as constructing definitions for concepts. This approach determines a set of necessary and sufficient features for a particular thing to be an instance of a particular concept. To take a classic example, the necessary and sufficient features of being a bachelor are being unmarried and being a man. Although this seems sensible at first glance, this approach has been widely criticized in both the philosophical and psychological literatures, primarily because so many of our concepts are not able to be adequately defined in this way (see, e.g., Keil, 1989; Lakoff, 1987; Murphy & Medin, 1985; Wittgenstein, 1958).

Putting aside these philosophical foundations, within the field of cognitive development, researchers have tended to use "intension" to just mean "definition." Work on children's category-based intensions has generally taken two forms. First, there's a long literature in which children are asked to provide definitions for specific concepts. For example, Piaget (1929) asked children questions such as, "Do you know what it means to think of something?" This probed children's understanding of what he called "the notion of thought" (p. 37). From children's definitions and their reflection on their thinking, he suggested that children do not initially understand the differences between mental and physical entities, and that this understanding must be acquired (see also Johnson & Wellman, 1982). These conclusions were part of his theory of egocentricity, in which younger children had difficulty differentiating between the self and the other. The development of that differentiation marked an important transition in development, specifically forming one facet of the emergence of concrete operations.

Second, there is another long literature relating children's intensions (their definitions of categories) to their *extensions* (which in cognitive development mostly refers to their judgments about category membership). In this research, children are asked whether particular objects or entities are members of a category given descriptions of the objects or entities that emphasize different features. These studies tend to show that, between the ages of 4 and

10, children's conceptual representations transition from being based more on superficial, perceptual, or nonessential features of objects (e.g., taxis are yellow) to ones that are more central to their meaning (e.g., taxis give people rides for money). This is known as the "characteristic-to-defining shift" (Keil, 1989; Keil & Batterman, 1984) or the shift between concrete and abstract definitions (Anglin, 1977).

Despite these interesting findings, Keil (1989) warned researchers against relying too heavily on the method of asking children to define words explicitly (i.e., asking about intentions). His concern was that, "because of the wide range of definition types children give" (p. 69), responses from children would be difficult to code systematically. He adds, "The task of giving definitions places unusual demands on children, and the younger ones may have little experience with it. . . . The definition task is at least partly a metalinguistic skill" (p. 69). His concern was that asking children for category-based intensions is an artificial task, which might capture children's understanding about pragmatics or about social interactions and not their understanding of the target concepts or categories. More generally, the concern is that asking children to define categories probes their linguistic capacities, not their cognitive capacities.

We want to point out three reasons why investigating children's category-based intensions might be more informative than Keil suggests. First, there are several instances in which even preschoolers' intensions and extensions show a great deal of coherence (Caplan & Barr, 1989; Gentner, 2003; Maguire et al., 2008). For example, Caplan and Barr (1989) showed that the ways in which children defined categories related to the ways in which they made judgments about category membership. This implies that asking children to articulate their definitions of a concept might tell us something about their understanding of that specific concept, particularly what should and should not be categorized as a member of the category.

Second, the majority of this research has investigated children's definitions for artifacts or natural kinds. Asking children to define "taxi" or "island" may provide insight into how they represent those specific concepts, as well as others that have salient characteristic features. But work in this area has rarely investigated how children define concepts that are more abstract or concepts that are related to an activity (as opposed to an object or entity). Our approach in this part of the book is to query precisely these kinds of concepts (specifically, science, learning, teaching, play, and pretending) in an effort to understand more about how those concepts develop and about how children's understanding of those concepts develop. This approach fits with one of the major tenets of rational constructivism, namely that children are actively responsible for constructing theories about the world based on input they receive, whether via testimony from others or based on the results of their own exploration. As part of this process, children not only develop intuitive theories about how the world works, they also develop the ability to reflect explicitly on the content of these theories. The conceptual structures that children construct should be coherent, and the relation between children's intensions and extensions of their conceptual structure is an example of this coherence. It is even possible that children's reflection on this coherence could promote further conceptual change (see also Sobel, 2004b).

Third, as noted in chapter 1, we suspect that there may be relations between children's understanding of what science is and their abilities to think scientifically. To review briefly, children who understand that science is a set of abilities or a way of learning about the world may approach scientific thinking problems with a more productive mindset for solving such problems than children who believe that science is merely a set of content areas. This may be another case in which children's metacognitive abilities to reflect on their own thought processes or to understand their own actions can positively influence their reasoning abilities in scientific contexts (see Kuhn, 1989, 2002).

Children's Conceptions of "Science"

Although asking young children what they think the word "science" means is a fairly simple matter, few researchers have actually done so. One related body of prior work examines children's responses to the Draw a Scientist task (Chambers, 1983), in which (as the name suggests) children are asked to draw a scientist. Participants are sometimes additionally asked to label parts of their drawing or to provide the scientist with a name. Their drawings are scored on various dimensions, including how many stereotypical elements are included (e.g., being male, wearing a lab coat, having glasses) and whether the child exhibits self-efficacy as a scientist by including elements that resemble themselves or by naming the scientist after themselves.

In a large-scale study of this test, Chambers (1983) asked more than 4,800 children from Canada, the United States, and Australia to draw a scientist.

His qualitative analysis of aspects of the drawings showed that drawings by children in kindergarten and first grade had few or none of the elements of our culture's stereotypical image (i.e., a white male chemist working alone in a lab, wearing glasses and a lab coat). These elements emerged starting in the second grade, and their inclusion continued to increase through the elementary years. Students from lower-socioeconomic status (SES) backgrounds in this study showed an even later emergence of this stereotypical image, and their drawings were also less likely to include varied equipment. Chambers took these findings to indicate that these students had a less well-developed understanding of science instruments than students of higher-SES backgrounds. Further, it was extremely rare for any child to draw a woman in response to the task prompt, and all female scientists in the sample were drawn by girls. In terms of other elements, almost no child drew naturalists or explorers, tending instead to focus on labs and chemistry. Chambers also found some rare but consistent images of a stereotypical mad scientist, including underground labs, high secrecy, explosions, and so on. More recent treatments of this task tend to find similar results, with the stereotypical view of scientists as male chemists dominating children's drawings across age, gender, and national boundaries (Finson, 2002).[1]

Although the Draw a Scientist measure primarily focuses on when and to what extent children internalize a culture's stereotypical view of a scientist, it also provides insight into how children conceptualize science in general. To the extent that children internalize a view of scientists as male chemists working alone in labs, this internalization shapes the kinds of activities they consider to be scientific, as we will see from their responses to our questions (described below).

Aside from this drawing task, most of the work done on children's explicit conceptions of science and scientists has been done with children in late elementary school or later, and these measures usually do not ask directly about these children's definitions of "science" (e.g., Aikenhead & Ryan, 1992; Bourdeau & Arnold, 2009; Halloun, 2001; Kahle et al., 2000; Klopfer & Cooley, 1961; Lederman et al., 2002; Osterhaus et al., 2015; Rubba & Anderson, 1978; Schwartz et al., 2008; Tsai & Liu, 2005). Conversely, while some explicit measures of scientific understanding are aimed at younger children, these tend to examine their conception of themselves as able to learn science (e.g., Mantzicopoulos et al., 2008), which may not be directly related to their understanding of what science itself is.

The one instrument that we know of that does ask younger children to define the word “science,” as we do, is the VNOS-E (Views of Nature of Science Questionnaire—Elementary School; Lederman et al., 2002; Lederman & Lederman, 2004). This question is typically administered as part of a larger interview probing students’ understanding of the nature of science. For example, Walls (2012) presented a version of this instrument to a group of 23 African American third graders. These children said that science was for learning about the natural world, making inventions and other creative products, and making discoveries. Most students mentioned experiments as being important for science, and they explained experiments in terms of finding something out or testing something or answering a question one might have. Specifically, most of the students in this study mentioned topics within the natural world (91%) as well as experimentation (74%) in their responses. Additionally, and in line with results from the Draw a Scientist task, a large minority (39%) mentioned potions or mixing substances together.

Studies using the VNOS-E provide a good baseline for understanding how children view the practice of science. However, the approach taken with this instrument differs from the approach we describe below in two main ways. First, children’s responses to this question are usually analyzed qualitatively, rather than the more quantitative analysis we present below. Second, the children who are interviewed tend to be in third grade or higher. We believe that this is past the point when they would have begun to form conceptions about what science is.

Our sample thus included children between the ages of 3 and 11, to track children’s early experiences with science and to probe how their conceptualizations of science develop over time. In support of this approach, we note that most of young children’s knowledge about science is learned from exposure to formal and informal environments and from natural experiences, such as parent-child interaction, peer-to-peer interaction, or their own observations (Callanan & Jipson, 2001; Callanan & Oakes, 1992; Evans, 2000; Evans et al., 2016; Fender & Crowley, 2007; Jant et al., 2014; National Research Council, 2009; Szechter & Carey, 2009), all of which are available to children well before third grade.

Documenting what children understand about science has potentially important implications for early science education, which aims to impart a sense of what science is and how it works in addition to knowledge of science content (e.g., National Research Council, 2012; NGSS Lead States, 2013). While we agree that we should remain skeptical about the extent to which young children can reflect on the content of their own concepts, querying children's definitions of "science" can serve to illustrate the processes that connect an implicit conceptual understanding to a more explicit awareness of one's beliefs. Understanding more about children's definitions of "science" can allow us to trace the connections between the development of children's abilities to engage in explicit reflection and the development of their scientific thinking in general.

Asking children about the definition of science can also help elucidate what Osterhaus et al. (2017) refer to as an "epistemological understanding of science"—the understanding of how data relate to theories (following Kuhn, 2002). Osterhaus and colleagues demonstrated that 8- to 10-year-olds showed relations between their performance on a battery of advanced theory of mind measures (described in chapter 7) and their epistemological understanding. This epistemological understanding of science was also related to a third capacity, children's understanding of experimental design. These researchers measured this understanding by stepping children through confounded or unconfounded experimental designs and their resulting conclusions to see whether children registered that unconfounded procedures would result in legitimate causal conclusions while confounded procedures would not (see also Osterhaus et al., 2015).

That study indicates that there is a relation between children's understanding of the scientific process and their ability to determine whether certain experimental designs would result in causal conclusions. Our goal in having children define science, which we discuss in the next section, was to investigate a similar relation. We wanted to explore what children believe science is and how these definitions relate to their understanding of scientific methods and conclusions, as well as to their reasoning and exploration more generally.

"What Is Science?"

Our approach in this study was simple. We asked our participants an open-ended question: "What is science?" If they said that they did not know or if they refused to answer, the question was rephrased: "What do you think the word 'science' means?" If they still said that they did not know, the interview ended. If they did provide a response, they were given further open-ended prompts to elaborate on their response, for example, "Can you tell me anything more about that?" or "What else can you tell me about science?" These prompts were repeated until they indicated that they had provided all the information that they could, following a procedure for eliciting children's understanding of the meaning of a word from Blewitt et al. (2009). It usually took about a minute for each child to complete the interview.

We administered this interview to many of the participants that we tested in the other investigations reported in this book. Although their experiences in our studies differed in other ways, here we report just their responses to this interview. This complete data set includes 940 children ranging between 3 and 11 years of age (mean age = 84.7 months; age range 37.3–143.6 months), with 481 girls and 457 boys (the gender of two children was not reported). Some of these children were tested in the lab, some at preschools or elementary schools, and some at museums or playgrounds.

We also tested a sample of 113 adults (85 female), some of whom were undergraduate students recruited from psychology department participant pools and some of whom were the parents of children that we tested. Our analyses here focus primarily on children's responses, but these adults' responses provide a useful developmental comparison.

Unsurprisingly, children's responses to our question were wide-ranging (see figure 8.1). Many participants talked about a specific domain of science or a particular science-related activity (e.g., "nature, like a frog or outside"; "mix stuff together"). Other participants focused more on science as a general process of inquiry (e.g., "an experiment; when you test something," "discover something, like what the moon is made of; like about learning"). Still others responded that they did not know what the word meant or provided responses that were off-topic or repetitive (e.g., "artwork," "science museum"). As in previous work, a significant minority of participants gave responses that conformed with social stereotypes of scientists, referencing potions,[2] mad scientists, or science-fiction events (e.g., "It means a lot like math, except you do potions and stuff"; "It's really dangerous because most scientists want to make dinosaurs come back alive, meat eaters too").

Given the wide variety of responses to this question and the fact that many of the answers included multiple themes, we adopted a quantitative

Figure 8.1

Word cloud representation of children's responses to the "what is science" interview. The larger the word in the cloud, the more frequently it appeared in children's responses.

coding scheme that allowed us to apply multiple nonexclusive codes to each answer. There were five coding categories: Specificity, Science Word, Learning, Other Process, and Personal Experience.

A response received a Specificity code if it mentioned a specific area or aspect of science or related topic, such as "bones," "math," "earth," or "fossils." The point of this code was to consider whether children's definition of "science" referenced a particular activity or topic.

A response received a Science Word code if it contained a science-oriented keyword such as "experiment," "observe," "hypothesis," or "research." The goal here was to determine whether children's definitions were just key words that they might have been exposed to, whether in school or through media. For example, children's television shows often use words like "hypothesis" when teaching science; not all of them, however, use it correctly.

A response received a Learning code if it referred to a process of knowledge change. These responses often contained the word "learning" but also could refer to other processes, such as figuring things out or making discoveries. The goal of this code was to consider whether children recognized the relation between science and knowledge change and recognized that scientific investigations could be used for the purpose of acquiring knowledge or changing beliefs. In contrast, a response received an Other Process code if it referred to an activity other than learning, such as "mixing things together," "inventing," or "making stuff." Unlike the other categories, the Learning and Other Process categories were treated as mutually exclusive. If a child mentioned an active process that involved learning, or if they mentioned a definition with both a learning process and another process, the Learning code took precedence.

The final category, Personal Experience, was applied if the response either described an experience that the child had with science or expressed the child's personal opinion. Responses such as "I go to science camp" or "I like science" received this code. We included this category because some prior work suggests that talk about personal connections might be related to children's engagement, as well as their understanding of parents' causal explanations (Callanan et al., 2017).

Six independent coders, blind to participants' ages and genders and performance on other measures, scored the responses. Reliability was established by asking each coder to categorize 25 responses. Agreement was measured and the coding scheme was iteratively refined twice until agreement on each code was 85% or higher. At that point, each coder independently coded a section of the remainder of the sample.

We employed this coding scheme to try to rein in the complexity of this data set, though we acknowledge that it glosses over much of the richness of children's and adults' responses. It is also important to note that we did not conduct these analyses with the goal of determining whether children or adults provided a correct answer to the question. The issue of how science should be defined is a complex one, and there likely is not a simple, single correct answer (see Godfrey-Smith, 2003). Rather, we aimed to discover developmental trends in children's responses and to examine how and when they used the different coding categories in their responses.

Distribution of Coding Categories

We first examined the overall usage of each of the five coding categories across all children and adults (table 8.1). For the children, we found that about 30% to 40% of the sample used the Specificity, Science Word, Learning, and Other Process categories. Science Word was significantly less common than the other three categories,[3] which did not differ from each other.[4] The Personal Experience category was significantly less likely to be used than the other four.[5] Adults, in contrast, used the Science Word and Learning categories most, followed by the Specificity and Other Process categories, followed by Personal Experience.[6]

Age differences. We compared the use of each of the five codes by age to test for developmental differences. Within the child sample, we found no significant relations with age in the rates of referring to specific topics,[7] non-learning processes,[8] or personal experiences.[9] Comparing children to adults, we found no differences in children's overall appeal to specific topics[10] and non-learning processes.[11] These findings are somewhat surprising; they indicate that even young children share adults' views about what science is with respect to these features. However, children were significantly more likely to refer to personal experiences than adults.[12]

In contrast, the use of science words increased significantly with age across the child sample,[13] and adults were even more likely than children to use these words (figure 8.2).[14] It is somewhat unsurprising that older children and adults were significantly more likely than younger children to use

Table 8.1

Use of the five coding categories for the "what is science" interview by age

Coding category	Children aged 3–11 (*n* = 940)	Adults (*n* = 113)	Total (*n* = 1053)
Specificity	377 (40.5%)	38 (33.6%)	415 (39.8%)
Science Word	250 (26.9%)	69 (61.1%)	319 (30.6%)
Learning	349 (37.5%)	72 (63.7%)	421 (40.3%)
Other Process	354 (38.0%)	36 (31.9%)	390 (37.4%)
Personal Experience	76 (8.2%)	2 (1.8%)	78 (7.5%)

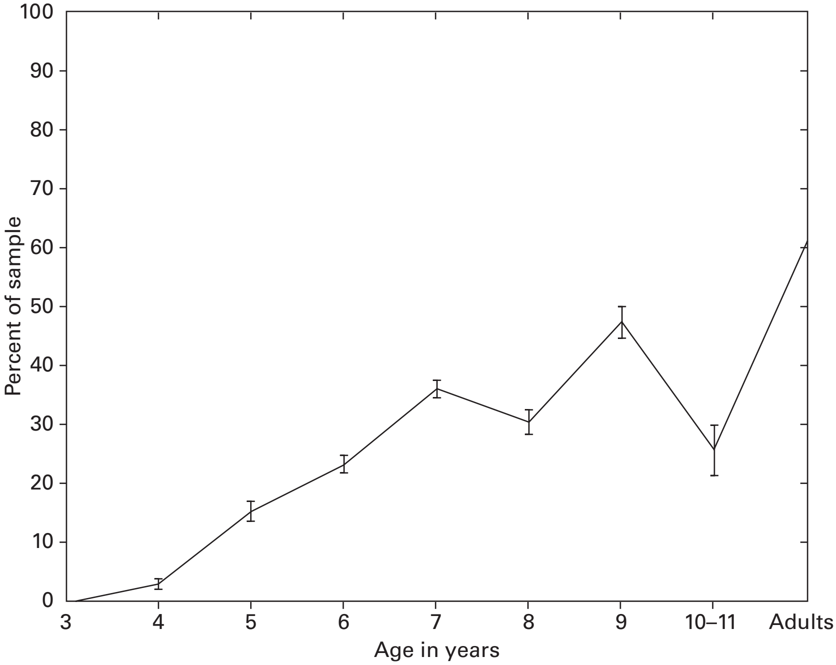

Figure 8.2
Proportion of sample using a science word, by age.

science-specific vocabulary. This indicates that science education already includes these terms even in early elementary school or in informal educational materials aimed at this age group.

We find it more intriguing that older children were significantly more likely than younger children to refer to science as a process of learning,[15] and adults were yet more likely to do so[16] (figure 8.3). This more general and flexible way of understanding science could be seen as a hallmark of developing a more mature understanding of science. For this reason, we had initially expected references to specific topics to decline with age while references to general learning processes rose. But this was not the case. Rather, it seems that both adults and children retain an idea of science as being about particular topics or involving particular kinds of activities. What develops in parallel to this understanding is the idea that science is a general way of learning about the world.

Gender differences. In addition to age, we examined whether there were any gender differences in children's definitions (table 8.2). Looking at the

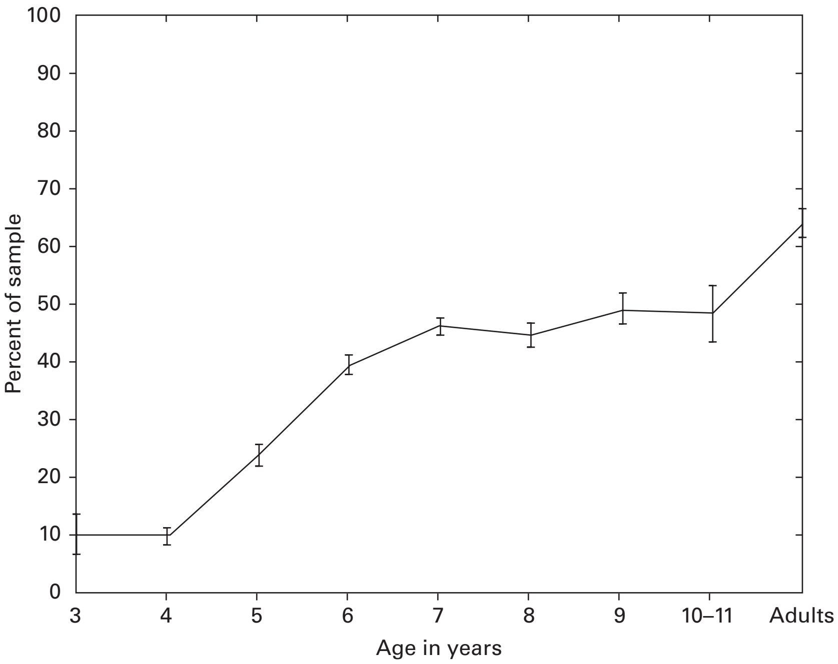

Figure 8.3
Proportion of sample referring to learning, by age.

entire child sample, we found that girls were marginally less likely to name a specific topic than boys[17] and significantly less likely to refer to learning.[18] Girls, however, were marginally more likely to use a science word than boys,[19] and girls were also significantly more likely to refer to a personal experience than boys.[20] The pattern for these four coding categories remained consistent across children of all ages in our sample.

For the Other Process category, we found no gender differences in the overall sample.[21] However, when we examined the effect of both gender and age, we found an interaction effect: As children got older, girls became more likely to refer to an active process while boys became less likely to do so.[22] No other coding category showed this pattern.

Overall, then, we see some evidence of gender's effects on how children talk about science, with girls and boys choosing to refer to different aspects of science in their definitions. These differences might be related to the persistent gender bias in science, whereby male scientists are disproportionately more common than female scientists (see e.g. Huang et al., 2020; Larivière et al., 2013). Additionally, in some studies, adults perceive men as being more competent in science (Nosek et al., 2009), and men often receive preferential treatment in science labs (Moss-Racusin et al., 2012). Even children are aware of these gender stereotypes: By the age of 6, they report that men are more likely to be “really, really good” at science than women (Bian et al., 2017), and their drawings of scientists are disproportionately male, as reviewed above (e.g., Finson, 2002). Girls’ persistence in a challenging science task are also disproportionately negatively affected by instructions to “be a scientist” rather than to “do science,” as described in chapter 1 (Rhodes et al., 2019). These findings suggest that early exposure to science and early socialization of children with respect to these gender biases could have influenced children’s responses to our questions.

What is not clear from our data, however, is whether these differences in children’s definitions have consequences, either for children’s performance on related tasks or for their attitudes toward science more generally. Looking just at the subset of children from this sample who participated in our blicket task in one form or another ($n = 590$), we found no relation between gender and performance: 42.9% of girls and 44.6% of boys responded to the task in a manner consistent with understanding the interactive causality. For the moment, then, we cannot say with any certainty whether the gender differences we discovered in response to our interview might be symptomatic of girls’ attitudes toward science or diagnostic of their performance on science, technology, engineering, and mathematics (STEM) tasks. Indeed, the fact that girls in our sample were proportionally more likely to refer to a personal experience when asked about science could reveal that they feel particularly connected to science or competent at doing science, contrary to the research cited above.

Table 8.2
Use of the five coding categories for the “what is science” interview by children’s gender

Coding category	Girls ($n = 481$)	Boys ($n = 457$)
Specificity	178 (37.4%)	198 (43.7%)
Science Word	141 (29.6%)	109 (24.1%)
Learning	163 (34.2%)	186 (41.1%)
Other Process	188 (39.5%)	165 (36.4%)
Personal Experience	48 (10.1%)	27 (6.0%)

There were no gender differences in the adults' uses of these categories in our data. This suggests that the patterns we observed in children may not be maintained throughout their lives. Later science instruction, caregiver-child or peer-to-peer interaction, or other experiences in high school or in college may negate any earlier gender differences in our sample.

This conclusion was corroborated by data from a different sample: parents who participated in a set of studies about parent-child interactions, particularly in museum settings. These data were taken from several samples we collected (Callanan, Legare, Sobel et al., 2020; Letourneau et al., 2021; Sobel et al., 2021), which include 600 parents of children participating in research at a set of museums around the United States. The adults completed the Attitudes Towards Science questionnaire (Szetcher & Carey, 2009), which consists of fifteen questions that look at how adults think about science and scientists. Responses to these questions are scored on a scale from 1 to 7, with higher scores indicating that participants had more positive attitudes toward science. The overall questionnaire is divided into three subscales. The first subscale is Personal Interest in Science, which measures how much participants are interested in science and scientific topics. This subscale includes items like, "I would enjoy being a scientist." The second subscale is Views of Science and Scientists, which measures how much participants think science is an important topic and scientists are engaged in important work. This subscale includes items like, "Scientists are among the most successful people." The third subscale is Utility of Science, which measures the extent to which participants think science is useful for society. This subscale includes items like, "Thinking like a scientist is only useful when taking a test in science class."

The goal of administering this measure in this parallel project was to see whether children's causal reasoning, learning, and engagement with STEM activities at museum settings were related to parents' attitudes toward science. The short answer was that they were not, at least in this set of studies (although other facets of parents' background in science did relate to the ways in which children explored the museum exhibits). However, we can use this set of data to examine whether there are any gender differences in parents' attitudes to try to probe more deeply into the patterns that we found in the children's responses to our "what is science" interview.

Of the 600 parents, 157 self-identified as male and 436 self-identified as female (the other 7 left this question blank). Overall, the group of parents who identified as male had significantly higher scores on this questionnaire (mean = 5.41 out of 7, SD = 0.77) than the group who identified as female (mean = 5.12 out of 7, SD = 0.73).[23] Although that result might suggest that there are gender differences with respect to parents' attitudes toward science, the picture becomes more interesting when we examine the subscales. The entirety of the difference between these two groups was in the Personal Interest subscale, where males scored higher than females (5.43 vs. 4.94).[24] Otherwise, males and females in the adult sample did not significantly differ in their views about science and scientists (scores of 5.07 vs. 4.81)[25] or their beliefs about the utility of science (scores of 5.59 vs. 5.52).[26] Contrary to typical findings about gender biases in science, but in support of our results with children, these results suggest that adults may not fully endorse the stereotype of men being more competent at science. However, this sample may be somewhat biased because these parents chose to participate in a scientific study.

Relations with Language

Unsurprisingly, adults' definitions of "science" (mean word count = 28.14, SD = 22.68) tended to be longer than children's (mean word count = 17.47, SD = 17.12).[27] The length of children's definitions increased linearly with age.[28] Further, longer definitions tended to be coded into more categories for both children and adults.[29]

Although it makes sense that longer responses received more codes, we were particularly interested in examining whether longer responses were also more likely to be coded as indicating a learning process, which seems to be a hallmark of more mature responses. This code was more common in older children and adults, so children who gave longer responses (i.e., children who might be more linguistically advanced) might also have been more likely to mention learning regardless of their age. This was indeed the case: Older children, as well as those who give longer responses, were more likely to receive a Learning code.[30] Interestingly, though, these two effects were independent from each other; there was no interaction between age and response length here.[31]

Because we administered the "what is science" interview to all the participants in our Springfield school sample (see chapter 7), we could also examine relations between children's responses to this question and their scores on standardized tests of language ability. For the first graders who were tested in 2017 ($n = 112$), students who used a science word were marginally more likely to score higher on the MAP test of reading.[32] This makes sense, because both the Science Word code and the reading test capture aspects of children's vocabulary. The relations for the third graders ($n = 116$) were more interesting: These students received significantly higher scores on the MAP test of reading when their definitions of "science" mentioned learning,[33] but significantly *lower* scores when their definitions mentioned another active process.[34]

For the statewide PSSA test that these students went on to take in fourth grade, again the students who received a Learning code had higher scores on the language section of the test and the students who received an Other Process code had lower scores, though these relations were only marginally significant.[35] Even more interestingly, there were significant positive relations between these students' use of the Learning code and these students' scores on the math[36] and science[37] sections of the PSSA test. Children who have more mature language, math, and science skills, as measured by these standardized tests, were thus more likely to view science as a way of learning about the world.

Relations among Coding Categories

Our five coding categories were designed to capture orthogonal concepts in our participants' definitions. However, with the exception of the Learning and Other Process categories, they were not mutually exclusive. We were thus interested to determine whether there were relations among these codes. When examining responses from all children together, we found that their use of the five categories was generally not related to each other; children whose definitions were coded into one category (e.g., Specificity) were not necessarily more or less likely to have their definitions also coded into a second category (e.g., Learning).[38] This indicates that these codes captured independent features of children's definitions of "science," as intended.

There were two exceptions to this general pattern. First, children whose definitions used a science word were significantly less likely to also talk about a non-learning process.[39] Potentially, children who chose to use scientific vocabulary in response to our prompt may not have felt the need to also talk about a particular scientific activity. Second, children whose definitions *did not* mention an active process were significantly less likely to also talk about a personal experience.[40] Given that so few children referred to personal experiences, though, we do not believe that this finding reflects a particularly important pattern in children's thinking about science.

In terms of adults, as with the children, there were generally no relations among the coding categories. Here, however, the exception to this general conclusion was with their use of science-specific vocabulary: When adults used a science word, they were also significantly more likely to refer to learning and to mention a different active process.[41] That is, adults who used science-specific vocabulary in their responses also tended to conceptualize science as a process of learning, potentially reflecting a more complete understanding of science.

Summary of Findings from the "What Is Science" Interview

This investigation was designed to uncover how children and adults talk about science. Our question here was intentionally broad and open-ended, as we wanted to find out what subjects would say about science with little guidance. We applied a five-category coding scheme to narrow down the scope of these responses and to investigate how other variables affected what children and adults said. In general, both children and adults saw science as a type of activity involving specific content. Older participants, as well as children in our sample with more advanced linguistic and mathematical skills, were additionally more likely to define "science" as a process of learning.

Testing the Relation between Definitions of "Science" and Measures of Scientific Thinking

Diagnostic Reasoning Task

Having obtained a view of how children define "science," we wanted to test whether there were any relations between these definitions and children's scientific thinking. We began this investigation with the first sample of children we tested using our diagnostic reasoning task with the blicket detector (introduced in chapter 5). This sample included thirty-six 5- and 6-year-olds and thirty-six 7- and 8-year-olds (previously reported in Sobel et al., 2017). We measured whether children responded correctly to the diagnostic reasoning task and whether they defined science as an active process rather than as a particular type of content. In this sample, as predicted, we found a significant relation between responding correctly on the blicket task and children's definitions of "science."[42] Children who mentioned some kind of active process in their definition of "science" were significantly more likely to succeed at the blicket task (choosing the correct response 83% of the time) than children who did not (choosing the correct response only 39% of the time). Even more importantly, this correlation remained significant when age was factored out.[43] That is, children who had a conception of science as a process also performed well on our diagnostic reasoning task, whereas those who provided other types of definitions did not—regardless of their age.

These results imply that children who believe science to be a learning activity or another active process might be better able to engage in diagnostic reasoning, as we hypothesized. We also have some support for this conclusion from our study in the Springfield School District, described in chapter 7, in which students who correctly solved our blicket task in third grade performed significantly better on a standardized assessment of their science knowledge in fourth grade.

We next attempted to replicate these findings in a larger sample of children, which included 590 children who both completed our blicket task (or the butterfly version of this task) and provided a definition of "science" (303 girls, 285 boys, 2 unreported gender; mean age = 85.9 months; age range 46.2–132.9 months). These children were drawn from our studies of the role of contextualization (previously reported in chapter 6) and from our partnership with the Springfield School District (previously reported in chapter 7). In this sample, children who defined science as a way of learning about the world were marginally more likely to succeed on the diagnostic reasoning task that they were administered (blicket or butterfly).[44] Although these findings were not as strong as those from the earlier sample, they still support our argument that something about understanding science as a learning process may benefit children's reasoning abilities.

We also found a relation between these definitions and children's justifications for their responses on the diagnostic reasoning task. Children whose definitions of "science" referred to learning were more likely to justify their responses by referring back to the data they had observed in the blicket task (24%), relative to children whose definitions did not refer to learning (15%).[45] This could suggest that understanding science as a process of learning may help children explicitly understand the structure of our diagnostic reasoning task, which does require an examination of the data from the demonstration phase to solve. But we are hesitant to draw strong conclusions from this relation, primarily because children's justifications of their choices in the blicket (or butterfly) task did not generally relate to their performance in this task.

We also attempted a more direct test of the hypothesized connection between definitions and task performance in a subset of 89 participants (48 girls, 41 boys; mean age = 102.0 months; age range 84.2–129.7 months).[46] These children were all administered the blicket detector task. Half of these children heard the task described as "a science game" and the other half heard this task described as "a puzzle game." Given the arguments above, we had anticipated that participants who heard the task described as "a science game" would perform better, because this framing would encourage them to bring their scientific thinking abilities to bear on answering our questions. However, there was no difference in performance between these two groups.[47]

Children's Exploratory Behavior

We also looked at the relation between children's definitions of science and their exploratory behavior at museum exhibits. One such project is reported in Callanan, Legare, Sobel et al. (2020), previously described in chapter 3. In this study, we asked 325 children between the ages of 3 and 6 to define "science" after they had played with a parent at a gear exhibit in a children's museum. These children also played with an experimenter in a set of tasks that were designed to test their understanding of gears. As reported in that monograph, children's definitions of science did not relate to their causal knowledge about gears.[48]

Here, we consider a different facet of these data: how children's definitions of science related to the way they played at the exhibit with their parent. One aspect of children's play at the exhibit was what we called *systematic exploration*, in which they constructed a gear machine by connecting one gear to a set of gears and then tested how that machine functioned by spinning the machine. As reported in chapter 3, the proportion of children's systematic exploration in their free play related to their understanding of the causal system on a set of follow-up measures. The proportion of systematic exploration also significantly positively correlated with whether children used a science word in their definition and whether they defined science in terms of learning. This proportion also correlated with whether they defined science in terms of a specific activity or topic, although this correlation was negative,[49] indicating that the more children thought science was a specific activity, the *less* systematic exploration they engaged in during their play. The relations between systematic exploration and generating a specific definition[50] or one that involved learning[51] were also still significant after controlling for children's age. So although responses to the "what is science" interview did not have much predictive power on the way children responded to questions about gears when tested in a laboratory setting, there were direct relations between how children defined science and their exploratory behaviors. How children defined "science" thus seems to have some bearing on how they chose to explore a STEM activity in a museum setting.

In a more recent study (Sobel, Letourneau, Legare & Callanan, 2021), we investigated parent-child interaction at a different museum exhibit, this time about electric circuits. Parents and 4- to 7-year-olds played together at the exhibit. We then presented children with a set of progressively more difficult circuit-construction challenges. After each challenge, children were asked whether they wanted to continue, and we coded the number of challenges they participated in as a measure of their engagement with the activity. We also coded the proportion of challenges that they solved as a measure of their understanding of the causal complexity of the circuits.

These same children were also given the "what is science" interview. Both of our measures—the number of challenges that children engaged in and the proportion of those challenges they could solve on their own—correlated negatively with whether they described science as a specific activity.[52] These correlations, however, did not remain significant when age was factored out.[53] Older children in this sample were more engaged by the challenges and performed better on them, and older children were also less likely to generate a definition of science that appealed to a specific activity.[54]

In this study, we also looked at the circuits that children and parents built during free play and coded when particular circuits were completed. We examined the 30 seconds prior to the completion of each circuit and counted the average number of actions that parents and children engaged in before completing the circuit. This provides a measure of who is completing the circuits. As reported in Sobel et al. (2021), parents' actions during this time frame negatively related to children's performance on the challenges: In general, the more that parents involved themselves in their children's play, the less children learned about the circuits when tested on their own. Interestingly, one facet of children's responses to the "what is science" interview did relate to the number of actions that both parents and children engaged in: If children generated a personal connection to science in their definition, they were more likely to engage in more actions in the 30 seconds prior to completing a circuit during their free play, and their parents were less likely to act.[55]

Do Definitions of "Science" Relate to Scientific Thinking?

Taken together, these data point to potential relations between children's definitions of "science" and their scientific thinking, including solving diagnostic reasoning tasks and engaging in exploratory actions. Some of our studies find strong relations, some find weak relations, and some find no relations. The museum studies with the gears and circuits are illuminating here, as children's definitions of "science" sometimes relate to the way they actively explore an environment and to whether this exploration leads to learning. Interestingly, we sometimes saw negative relations between children's performance in these museum tasks and their tendencies to refer to specific topics in their definitions, suggesting that this way of thinking about science may be somewhat less mature or less helpful to their exploration. And that makes sense: Thinking that science is just mixing things together, for example, ignores the process of scientific thinking that is the focus of these investigations.

Although we have some evidence for our hypothesized relation between explicit definitions and children's scientific thinking, we cannot conclude definitively that we have found strong relations because of the inconsistencies in our results. This might mean that our argument for the connection between explicit definitions of "science" and scientific thinking might be flawed. It is also possible that the relation is just weak from the perspective of statistical power, but is still applicable in certain circumstances. We take up these arguments in more detail in chapter 11, but for right now, what is important to take away from this discussion is that there are some relations between children's definitions of science and their reasoning behaviors, although the relations are not clear and require more investigation.

Finally, in agreement with the arguments from Keil that we reviewed at the beginning of this chapter, it is worth noting that our definition task is a bit odd. We are committed to the value of exploratory measures like this one to give us a sense of children's thinking, but the broad range of possible answers might be both a feature and a bug. Children may never before have had to explicitly reflect on how they conceptualize science, and their responses to this question may not have captured the full range of their thinking. Given this, the fact that any of our tasks found positive relations with performance suggests that there may still be some reason to believe that our hypothesis about this connection holds. A more focused definition task or a more highly structured interview could uncover more precisely how children think about science, and these responses might show stronger relations with their abilities to engage in scientific thinking on some tasks. These studies take the first step in this direction, and future work should tease out these potential connections in more detail.

"Is That Science?"

One of the major strengths of our open-ended interview is that it allows children to talk about anything that they believe about science, providing a rich view of the development of this concept. But this open-endedness is also one of its weaknesses: Children generated many different kinds of responses, and the conversation sometimes ran off course, particularly when the children we tested had little understanding about what the word means. We thus conducted two additional investigations, which asked closed-ended questions to probe children's beliefs about what science is in more detail.

The first investigation provided children with a list of actions and asked them to report whether each one was an example of science. These questions were asked to a subset of 88 children from the main sample, ranging from 3 to 10 years of age (44 female, 44 male; mean age=81.5 months; age range 37–131 months), as well as to 24 adult participants (22 female; 2 male). Each participant was presented with a list of 12 specific actions and asked whether each action was science (shown in table 8.3). Some of these actions were general processes used in scientific endeavors but that lacked specific scientific content (e.g., asking questions), some were specific scientific actions (e.g., finding out what happened to the dinosaurs), and some were not canonically scientific (e.g., doing jigsaw puzzles). We chose these items to reflect topics studied by several scientific disciplines (e.g., biology, chemistry, psychology) and other nonscience disciplines covered in elementary school (e.g., history, mathematics). These items were presented in a random order for each subject. For each item, participants were asked, “Is that science?” and were asked to respond “yes” or “no.”

In looking at patterns of agreement for individual items, we uncovered a few interesting patterns. First, children and adults did not differ in their rates of agreeing that the following activities are science: trying to figure out what happened to the dinosaurs, mixing things together, and learning about the world.[56] These activities, from the earliest age that we tested, seem to fall squarely into the category of science for all participants. It is not particularly surprising that something involving dinosaurs and a chemistry-related action were seen as scientific. It is reassuring, however, that the general activity of finding out about the world was also seen as scientific at all ages. Similarly, children and adults did not differ in their rate of agreeing that doing jigsaw puzzles is not science.[57]

Second, there were several items on which children and adults differed because adults or older children were more likely to claim that an activity is scientific than younger children. This was the case for trying to figure out why babies cry, asking questions, playing in the dirt, and finding out the names of different birds.[58] Figuring out why babies cry and playing in the dirt were not seen as scientific until adulthood. This latter item may have gained a high level of agreement from the adults because this study often recruited the parents of the children we were testing, who may have been particularly likely to see open-ended exploratory activities like this one as scientific. Only the 3- and 4-year-olds were ambivalent about whether finding out the names of different birds is scientific; children at all other ages, as well as adults, agreed that this is a scientific activity.

The item about asking questions showed an unexpected U-shaped pattern, with 3- and 4-year-olds as well as 9- and 10-year-olds and adults claiming that this is a scientific activity. Children in the middle of our age range, from 5 to 8 years old, tended to deny that this is science. We are not sure of the source of this U-shaped curve, but we suggest three speculative explanations. First, participants at the ends of our age spectrum may genuinely have had a different understanding of science, such that they agree that asking questions is science while children in the middle do not. Second, it is possible that participants across our entire age spectrum believed that asking questions is science, but the children in the middle of the age range might have been more influenced by their educational experiences, which tend to be highly didactic at these ages and which hence might have led

Table 8.3
Closed-ended categorization questions and percent agreement by age group

"Is that science?"	3- and 4-year-olds ($n=22$)	5- and 6-year-olds ($n=24$)	7- and 8-year-olds ($n=23$)	9- and 10-year-olds ($n=19$)	All children ($n=88$)	Adults ($n=24$)
Trying to figure out what happened to the dinosaurs[a]	90.9%	91.7%	91.3%	50%	93.2%	100%
Trying to figure out why babies cry[a,b]	40.9%	33.3%	52.2%	73.7%	48.9%	100%
Mixing things together[a]	77.3%	83.3%	91.3%	89.5%	85.2%	95.8%
Learning about the world[a]	86.4%	75.0%	91.3%	94.7%	86.4%	95.8%
Asking questions[a,b]	63.6%	58.3%	39.1%	68.4%	56.8%	87.5%
Playing in the dirt[a,b]	45.5%	29.2%	26.1%	31.6%	33.0%	83.3%
Finding out the names of different birds[a,b]	50.0%	83.3%	73.9%	89.5%	73.9%	79.2%
Studying what happened a really long time ago[a,c]	81.8%	79.2%	87.0%	52.6%	76.1%	70.8%
Doing jigsaw puzzles	50.0%	45.8%	26.1%	26.3%	37.5%	54.2%
Trying to add numbers together[b]	86.4%	66.7%	17.4%	31.6%	51.1%	37.5%
Reading stories[b]	50.0%	62.5%	17.4%	31.6%	40.9%	37.5%
Learning to write[b]	77.3%	75.0%	17.4%	10.5%	46.6%	29.2%

Note: [a]Indicates that more than 70% of the adult sample categorized the item as being science. [b]Indicates a significant difference among the age groups in rates of categorizing the item as being science. [c]Indicates a marginally significant difference in the age groups in rates of categorizing the item as being science.

them to respond that asking questions is not a scientific activity. Finally, and relatedly, children in the middle age range may have seen question-asking as an activity that applies across multiple disciplines and is not specifically characteristic of science, so they may have been reluctant to say that it was scientific because it is not limited to science.

A third interesting pattern that we uncovered in these data is that there were several items on which children and adults differ because children were more likely to claim that an activity is scientific than adults. This was the case for studying what happened a long time ago (marginally), trying to add numbers together, reading stories, and learning to write.[59] For studying what happened a long time ago, 9- and 10-year-olds were marginally less likely to claim that this is science, possibly because they are beginning to study history and social studies in a more in-depth way or in classes that are separate from their science classes, leading them to see this subject as different from science. For the other three activities, children younger than 7 tended to claim that these activities are scientific, while older children and adults tend to claim that they are not. This may reflect a general trend on the part of these children to see any kind of school-like activity as science.

Composite Scores

To analyze the overall development of children's understanding of scientific activities, we coded their responses to this set of closed-ended questions into a composite score. For all the items in which more than 70% of the adults categorized that activity as science, we gave children a score of 1 if they said "yes" for that item and a score of 0 if they said "no." If less than 70% of the adults categorized the item as science, we reversed this scoring, giving children a score of 1 for the item if they said "no" and a score of 0 if they said "yes." We then summed these scores into a scale from 0 to 12 (see table 8.4). Importantly, these scores reflected only the extent to which children's responses agreed with adults' modal responses; we remain neutral about whether these are the correct answers to the question of what science is.

These composite scores significantly correlated with age.[60] That is, older children were more likely to provide an adult-like pattern of responses about which activities were scientific.[61] More specifically, the 3- and 4-year-olds and the 5- and 6-year-olds did not differ from one another,[62] and the 7- and 8-year-olds and the 9- and 10-year-olds also did not differ from one another.[63]

Table 8.4
Children's average composite scores

	3- and 4-year-olds (*n* = 22)	5- and 6-year-olds (*n* = 24)	7- and 8-year-olds (*n* = 23)	9- and 10-year-olds (*n* = 19)
Mean score (standard deviation)	6.73 (2.31)	6.83 (1.93)	8.74 (1.19)	9.00 (1.15)

However, scores for the 7- and 8-year-olds were significantly higher than scores for the 5- and 6-year-olds.[64] This age range might thus be an important time for the development of children's ideas about science. That conclusion aligns nicely with the work reported in part II of this book, which also finds that this time period is important for the growth of children's scientific thinking, such as the diagnostic reasoning tested by our blicket detector task, and for many of the age-related changes that we documented on children's metacognitive capacities. We return to the question of whether these two developmental trends might relate to each other in chapter 11.

Relations between Closed- and Open-Ended Questions

We also considered the relation between children's categorization of items in this closed-ended measure and their responses to our open-ended question asking them to define "science." We found that children who used a science word in their definitions had higher composite scores ($M = 8.81$ out of a possible 12, $SD = 1.52$) than children who did not ($M = 7.53$, $SD = 2.22$).[65] Additionally, children who referred to learning in their definitions had higher composite scores ($M = 8.86$, $SD = 1.42$) than children who did not ($M = 7.39$, $SD = 2.24$).[66] This confirms the trends noted above, because higher composite scores and definitions of "science" that mentioned learning both reflect more adult-like patterns.

These results imply that, before the age of about 7, children's understanding of science is somewhat undifferentiated, and it is around this age that children begin to appreciate science as a process of learning and discovery. This developing understanding is reflected not only in their explicit definitions, but also in their judgments of what activities are science, and possibly also in how they are able to demonstrate their capacity for scientific thinking.

Children's Understanding of What Makes an Investigation Scientific

In our second closed-ended investigation, we probed children's understanding of science by telling them brief stories about characters who were conducting different kinds of investigations. Our general approach was to introduce children to a character who had a question in a particular area of science (chemistry, biology, or psychology). We varied what answer the character received to their question and how the character went about obtaining that answer.

Specifically, in one set of conditions, characters obtained answers to their question that were either accurate or inaccurate, but no information was provided about how these answers were obtained. In another set of conditions, instead of being told the answer that the character found, children were told only the method by which the character went about learning the answer. These methods were either consistent or inconsistent with “systematic empiricism” (Stanovich, 2012, p. 9): experimental interventions designed to answer scientific questions. In all cases, children were asked whether the character had done science and to justify their answers.

In general, judgments about whether investigations into certain questions are scientific should be independent of topic and independent of the exact answer that one obtains. Chemistry is no more or less scientific than biology or psychology; thinking that certain disciplines are more scientific than others places undue emphasis on the content of science, rather than on its process (see chapter 1 of Stanovich, 2012; see also our discussion in chapter 1). If children understand this, they should judge all cases with appropriate methods as science, regardless of the topic that the character is investigating or the answer that they receive.

Recent work, however, suggests that both adults’ and children’s folk thinking about science is influenced by the topic under investigation (Fernandez-Duque et al., 2015; Hopkins et al., 2016; Keil et al., 2010). This research suggests that individuals believe the natural sciences are more scientific than the life sciences, which in turn are more scientific than the social sciences. Results from our own work are consistent with these ideas. In the closed-ended study described earlier in this chapter, children readily judged certain topics (dinosaurs, chemistry) as science, while they were less willing to judge psychological investigations (such as finding out why babies cry) as science. These results suggest that children might be biased by topic when making judgments in our tasks, focusing less on an investigation’s method or its outcome and more on the nature of the question itself.

General Materials and Methods

To parallel the work on children’s judgments of activities reported above, we chose three phenomena for these experiments, one to represent each of three scientific disciplines: chemistry (why salt disappears in water), biology (what happened to the dinosaurs), and psychology (why babies cry). For each phenomenon, we constructed four vignettes that described a character who wanted to find out about that phenomenon. There were 12 vignettes in total.

Half of these vignettes focused on what the character found out—that is, the outcome of the investigation. These six vignettes followed a 3 (Discipline: psychology, biology, chemistry) × 2 (Outcome: appropriate/inappropriate) design; both variables were within-subject such that each child received six vignettes. For these stories, children received no information about the methods used to obtain this result. For example, the vignette that presented the appropriate outcome for biology said, "Joanne/Robert [gender matched to the child participant] wants to know what happened to the dinosaurs. She/He finds out that they all died because a meteor crashed into the Earth." The matched vignette with the inappropriate outcome said, "Joanne/Robert wants to know what happened to the dinosaurs. She/He finds out that they all died because they ate too much cake at a party." We tested 76 children in this condition (38 girls, 38 boys; mean age 88.03 months; age range 61–136 months).

The other half of the vignettes focused on what the character did, that is, the methods of the investigation; no information was provided about the investigation's outcome. These 6 vignettes also used a 3 (Discipline: psychology, biology, chemistry) × 2 (Method: appropriate/inappropriate) design. For example, our biology/appropriate method vignette said, "Joanne/Robert [again, gender matched to the child participant] wants to know what happened to the dinosaurs. So s/he looks carefully at fossils." The paired inappropriate method vignette said, "Joanne/Robert wants to know what happened to the dinosaurs. So s/he takes a picture of his/her dog." No information was ever provided about the result of these investigations. We tested 101 children in this condition (55 girls, 46 boys; mean age = 77.33 months; age range 37–143 months).

All the vignettes were accompanied by two pictures. The first picture was the same across conditions and always showed the character thinking about the question. The second picture varied by condition and depicted either the character's idea about the answer or the character engaging in an investigation (figures 8.4 and 8.5). Following each vignette, children were asked whether the main character of the vignette was doing science (e.g., "Is Joanne/Robert doing science?"). After answering this yes/no question, children were asked to justify their response: "Why/why not?"

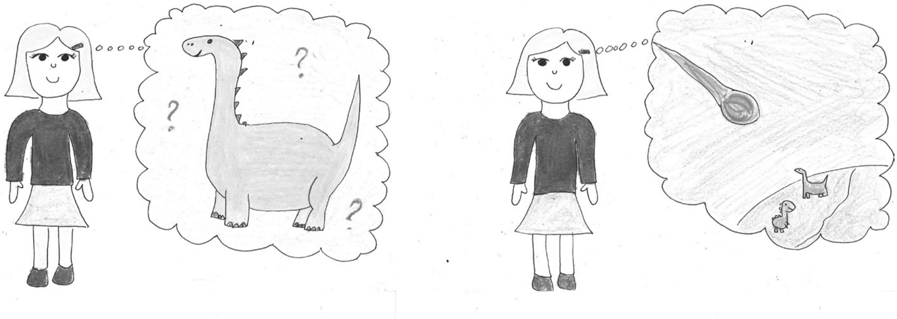

Figure 8.4
Example stimulus item from the study on outcomes. The character wants to find out what happens to the dinosaurs and discovers that a meteor hit the Earth and killed them.

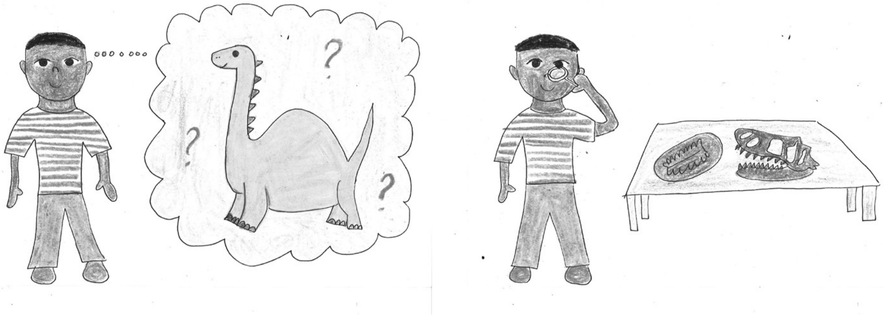

Figure 8.5
Example stimulus item from the study on methods. The character wants to find out about the dinosaurs and does so by looking at fossils.

What Kinds of Outcomes Are Scientific?

To test children's understanding of the role of outcomes, we counted the number of times that children accepted an appropriate outcome as being science or rejected an inappropriate outcome as not being science. The proportion of vignettes that children judged in this way are shown in table 8.5, with data separated by age group (via a median split). The average age of the younger group was about 6 years, and the average age of the older group was about 8 years.

Children responded similarly to all three of the vignettes with appropriate outcomes, correctly judging all of these as scientific.[67] But there were

Table 8.5

Proportions of correct responses across vignette types and age in the study on outcomes (standard deviations in parentheses)

	Chemistry Appropriate	Chemistry Inappropriate	Biology Appropriate	Biology Inappropriate	Psychology Appropriate	Psychology Inappropriate
All Children (*n*=76)	.82 (.39)	.53 (.50)	.70 (.46)	.68 (.47)	.74 (.44)	.61 (.49)
Younger Children (roughly 5–7-year-olds, *n*=38)	.76 (.43)	.34 (.48)	.68 (.47)	.50 (.51)	.84 (.37)	.34 (.48)
Older Children (roughly 7–11-year-olds, *n*=38)	.87 (.34)	.71 (.45)	.71 (.46)	.87 (.34)	.63 (.49)	.87 (.34)

differences in children's responses to the vignettes with inappropriate outcomes. Most importantly, only for the chemistry vignettes did responses differ significantly between the appropriate and inappropriate vignette.[68] That is, children only reliably distinguished between investigations with appropriate and inappropriate results for chemistry.

When we examined these results separately for the two age groups, however, different patterns emerged. The younger children tended to accept vignettes with appropriate outcomes as scientific,[69] but did not generally reject vignettes with inappropriate outcomes as not being science.[70] While this response pattern could reflect a simple bias to say "yes" to all items, a closer look at the data does not support this possibility. These children's responses did not differ between the two biology vignettes, which were judged as equally scientific,[71] but their responses did differ between the two psychology and two chemistry vignettes.[72] For these latter two topics, younger children were more likely to correctly distinguish between appropriate and inappropriate outcomes.

For the older children in the sample, responses were all significantly above chance,[73] except on the appropriate psychology vignette, which did not differ from chance.[74] Additionally, they were significantly more likely to reject the inappropriate psychology vignette as not being science than they were to accept the appropriate psychology vignette as being science.[75] This difference was not seen in the other two topics.[76]

Overall, then, there were marked differences between the younger and older children in our sample with respect to which features they considered in their judgments. Between the ages of 3 and 7, children correctly accepted investigations with appropriate results as scientific. But these younger children were much less likely than the older children to reject investigations with inappropriate results. For children between the ages of 7 and 11, topic played more of a role in their judgments: Older children were less likely to accept investigations with appropriate outcomes in psychology than for other two topics. In general, rejecting each of the inappropriate vignettes as not being science increased with age. However, accepting the appropriate vignettes as being science did not improve with age for the biology and chemistry vignettes; these were seen as scientific at all ages. Further, accepting the appropriate psychology vignette as being science actually decreased with age.

These results suggest that the outcome of an investigation is of primary importance in children's decisions about whether that investigation is scientific, but the investigation's topic (especially for psychology) and the age of the child also play some role. Again, we see important differences emerging around age 7.

Justifications of Responses. We categorized children's justifications for their responses into three main categories. *Metacognitive* responses referenced mental states or character's intentions (e.g., "because he's trying to figure out how dinosaurs died in the past"). *Topic* responses referred exclusively or primarily to the topic of investigation (e.g., "because babies is not a science"). *Outcome* responses referred exclusively or primarily to the result of the investigation (e.g., "because that's what [salt] does, the water melts it"). All remaining responses were coded as *Irrelevant*. Distributions of these justifications by age are shown in table 8.6.

For the most part, the distribution of justifications did not differ between the age groups (see last row of table 8.6). This suggests that both older and younger children in this sample understood the criteria they were using to make their judgments. What did differ is the number of Metacognitive and Outcome justifications between the appropriate and inappropriate vignettes for each topic. For each topic, children generated more justifications that referred to outcomes for the inappropriate vignette than for the appropriate one.[77] Conversely, children generated more justifications that referred to metacognition for the appropriate than for the inappropriate vignettes, though this pattern was only statistically significant for the psychology and chemistry vignettes.[78] This means that, when justifying their acceptance of an investigation with an appropriate outcome, children tended to refer to the character's beliefs and intentions. But when justifying their rejection of an investigation with an inappropriate outcome, children tended to refer to the outcome itself.

We also examined justifications in relation to their responses to the main test question. We first examined the number of children who justified a correct response to the test question for each vignette in terms of the outcome of the study. For the three appropriate vignettes, participants who judged that the character was doing science were more likely to justify that response in terms of the investigation's outcome than were participants who judged that the character was not doing science.[79] We found the same

Table 8.6

Distribution of justifications by age group (# of children) for the study on outcomes

Younger Group ($n=38$)						
	Chemistry Appropriate	Chemistry Inappropriate	Biology Appropriate	Biology Inappropriate	Psychology Appropriate	Psychology Inappropriate
Metacognitive	9	7	10	5	10	4
Outcome	12	20	13	20	13	22
Topic	2	1	2	2	2	2
Irrelevant	15	10	13	11	13	10
Older Group ($n=38$)						
Metacognitive	14	5	6	4	13	4
Outcome	16	31	21	28	13	26
Topic	2	1	3	2	4	2
Irrelevant	6	1	8	4	8	6
Difference in age groups	$\chi^2(3)=5.52$, $p=.14$	$\chi^2(3)=10.07$, $p=.02$	$\chi^2(3)=4.27$, $p=.23$	$\chi^2(3)=4.71$, $p=.19$	$\chi^2(3)=2.25$, $p=.52$	$\chi^2(3)=1.33$, $p=.72$

relation for the inappropriate vignettes, in which participants were more likely to justify their response that the character was not doing science by referring to outcomes.[80] This suggests that these children explicitly believe that appropriate outcomes are of primary importance in making an investigation scientific.

We also examined children's metacognitive justifications in this way and found generally the same pattern: Children were more likely to use metacognitive justifications when they accepted appropriate outcomes or rejected inappropriate outcomes.[81] The one exception to this pattern was for the psychology vignettes that had inappropriate outcomes. In this case, children's justifications did not differ depending on how they had responded to the main test question.[82]

Summary of Study on Scientific Outcomes. This study was designed to investigate the impact of an investigation's topic and its outcomes on children's explicit judgments of whether that investigation was scientific. Our results showed that children have a rudimentary understanding of scientific outcomes. They seemed to have a good understanding of what kinds of outcomes were scientific at the early end of the age range that we studied, and they seemed to be developing an understanding of what kinds of outcomes are not scientific over the early elementary years. However, they were also beginning to believe that different kinds of questions can be more or less scientific—in line with adults' intuitions that psychology is less scientific than fields like biology or chemistry (e.g., Keil et al., 2010).

What Kinds of Methods Are Scientific?

The second condition in this study considered whether children correctly accepted or rejected the use of a scientific method of investigation, and whether these judgments varied by topic (biology, chemistry, and psychology), by children's age, and by whether the method was appropriate or inappropriate. The proportion of correct responses are shown in table 8.7.

In general, children considered both an investigation's topic and its method in their judgments of whether that investigation is scientific. Specifically, when the method was appropriate, children responded correctly at above-chance levels to both the chemistry[83] and the biology[84] vignettes, but not to the psychology vignette.[85] When the method was inappropriate, children performed at chance on the psychology[86] and biology[87] vignettes and, unexpectedly, performed below chance on the chemistry vignette.[88] That

Table 8.7

Proportions of correct responses across vignette types and age in the study on methods (standard deviations in parentheses)

	Chemistry Appropriate	Chemistry Inappropriate	Biology Appropriate	Biology Inappropriate	Psychology Appropriate	Psychology Inappropriate
All Children (*n* =101)	.92 (.27)	.36 (.48)	.94 (.24)	.54 (.50)	.44 (.50)	.56 (.50)
Younger Children (roughly 3–7-year-olds, *n* = 50)	.88 (.33)	.28 (.46)	.88 (.33)	.51 (.50)	.48 (.50)	.46 (.50)
Older Children (roughly 8–11-year-olds, *n* = 51)	.96 (.20)	.44 (.50)	1.00 (0)	.58 (.51)	.39 (.49)	.67 (.48)

is, children often incorrectly said that the chemistry vignettes were science, even when the method used to investigate the question was inappropriate.

As can be seen in table 8.7, these same patterns were found independently in both age groups; there were few differences between the younger and the older children in this study. But when age was treated continuously rather than categorically, several relations emerged. For the chemistry vignettes, correct responses increased with age on both the appropriate and inappropriate cases.[89] For the biology vignettes, correct responses increased with age when the method was appropriate but not when the method was inappropriate.[90] For the psychology vignettes, correct responses increased with age when the method was *inappropriate* but not when the method was appropriate.[91] That is, older children were more likely to say that inappropriate psychological methods were not science, but they did not get better at saying that appropriate psychological methods were science.

These results show that children do not simply ignore the method of investigation; overall, they correctly judged investigations that used scientific methods as scientific more often than investigations that used nonscientific methods. However, this tendency varied significantly by topic. Children were much more likely to accept investigations using scientific methods as scientific and to reject investigations using nonscientific methods as not scientific when asked about a biology vignette. In contrast, children responded at chance for both types of psychology vignettes, regardless of method. And both chemistry vignettes were judged to be scientific, regardless of whether the method was appropriate. As children got older, they were more likely overall to reject nonscientific methods as not scientific. This was especially true for of the biological and psychological vignettes we presented.

Justifications of Responses. As in the previous study, we also coded children's open-ended justifications of their responses. We coded responses as *Metacognitive* if they contained some reference to mental states or intentions (e.g., "because he's learning about why babies cry"). *Exclusive* justifications involved children using a particle like "just" or "only," which indicated that the character's actions were inadequate to be scientific (e.g., "because he's just taking pictures").[92] All remaining responses were coded as *Irrelevant*. The distributions of these justifications by age group are shown in table 8.8.

Older and younger children generated different types of justifications for all six vignettes (see last row of table 8.8). Specifically, older children tended to generate more Metacognitive justifications while younger children tended

Table 8.8

Distribution of justifications by age group (# of children) for the study on methods

	Chemistry Appropriate	Chemistry Inappropriate	Biology Appropriate	Biology Inappropriate	Psychology Appropriate	Psychology Inappropriate
Younger Group (*n* = 50)						
Metacognitive	12	5	14	5	4	6
Exclusive	0	2	0	7	9	7
Irrelevant	38	42	36	37	36	36
Older Group (*n* = 51)						
Metacognitive	25	22	27	16	19	17
Exclusive	1	2	1	3	4	3
Irrelevant	24	26	22	32	26	30
Difference in age groups	$\chi^2(2) = 8.83$, $p = .02$	$\chi^2(2) = 15.02$, $p = .002$	$\chi^2(2) = 10.07$, $p = .02$	$\chi^2(2) = 8.81$, $p = .03$	$\chi^2(2) = 16.69$, $p < .001$	$\chi^2(2) = 7.62$, $p = .06$

to generate more Irrelevant justifications. We did not find differences in the pattern of justification types for the appropriate as opposed to the inappropriate methods for the chemistry[93] or psychology vignettes,[94] but there was a significant difference for the biology vignettes.[95] Specifically, participants tended to give more Exclusive justifications for the inappropriate methods and more Metacognitive justifications for the appropriate methods.

We also looked at the relation between responses to the test question and justifications. Here, we looked just at Metacognitive justifications, primarily because there were so few Exclusive justifications, but also because we believe that Metacognitive justifications are more likely to reflect a mature view of how science works.

We found that children were more likely to justify a correct acceptance of an appropriate method by referring to a metacognitive process, and less likely to justify an incorrect answer to these questions in this way.[96] However, correct rejections of inappropriate methods were only associated with metacognitive justifications for the chemistry vignettes; justifications for biology and psychology vignettes with inappropriate methods did not follow this pattern.[97]

General Summary of the Outcomes and Methods Studies

The results of these two studies reveal different patterns in how an investigation's topic, results, and methods affect children's judgments of whether that investigation is categorized as science. For vignettes that gave information only about the investigation's results, the topic of the investigation mattered little. Across the ages we investigated, most children were likely to accept investigations that came to appropriate outcomes as scientific and reject investigations that came to inappropriate outcomes. One exception to this general pattern was a negative correlation with age for psychology investigations that came to appropriate conclusions, suggesting that psychology was seen as increasingly *less* scientific as children grew older. This is in line with adults' views about psychology as being less scientific than other topics like physics, chemistry, or biology (e.g., Keil et al., 2010). In contrast, all three topics showed strong positive relations with age when the result was inappropriate: As children got older, they were more likely to reject inappropriate outcomes as not science. Moreover, on the vignettes where the character came to an inappropriate conclusion, children also provided more sophisticated justifications for those correct rejections as they got older. Children thus seem to be learning what are appropriate outcomes for each topic, leading them to judge that coming to such erroneous conclusions should not be counted as science.

Note that we do not necessarily endorse this view of science. Finding a correct answer may loom large in children's views, and it is likely to be emphasized in school. But true scientific investigations do not necessarily depend on their outcomes. Science aims at finding the truth, but finding an incorrect answer or getting a null result does not on its own make a particular investigation any less scientific.

In contrast, for vignettes that gave information only about the investigation's method, both the topic of investigation and the method influenced children's judgments. Children judged an investigation of a biological or chemical phenomenon as scientific more often than an investigation of a psychological phenomenon, regardless of whether the method was appropriate or inappropriate. At least during the preschool years, chemical phenomena are seen as so scientific that even investigations using nonscientific methods are judged as scientific the majority of the time.

But children in our studies did not simply ignore the method of investigation; they were generally more likely to judge investigations that used appropriate methods as scientific than they were to judge investigations that used inappropriate methods as scientific, though this tendency did not exceed chance levels. Here, what may be developing is an understanding that appropriate methods are scientifically valid, as the tendency to justify a correct answer with reference to the character's intentions and goals increased with age.

Combined with the less-nuanced pattern of responses found for vignettes that described only outcomes, these results indicate that, during the preschool and elementary-school years, children have an outcome-focused view of what science is, and they believe that investigations need to get the right answer in order to be scientific. This is potentially part of the more undifferentiated concept of science that children might have before the age of 7.

Interestingly, this focus on outcomes parallels developments within the domain of moral reasoning, in which young children of roughly the same age tend to initially privilege information about the outcome of an action, rather than an agent's intention, when deciding on issues of blame and punishment (see Cushman, 2008). The fact that children might be more influenced by the outcome of an investigation as opposed to its method suggests that science educators should focus on incorporating discussions of what makes a good scientific method into their pedagogy, and that such instruction should also take topic into account. These data can thus help point the way toward better science instruction in early elementary school, suggesting that rudimentary lessons about the philosophy of science—ones that focus on imparting an understanding of what methods are scientific and of how appropriate methods relate to doing science—might help even the youngest learners.

Developing an Understanding of What Science Is

In closing this chapter, we want to say two things. First, all the research we have described here has examined a relatively wide age range (3-year-olds to 11-year-olds). Given this, it is not particularly surprising that we found developmental changes in how children answered our questions. It is thus more important to consider what specifically is changing and why. We found that children at the younger end of this range were not particularly adept at understanding science as an abstract method of inquiry; their explicit definitions were more tied to specific topics or actions. The tendency to describe "science" as a series of domain-specific facts was maintained with age, but older children and adults were additionally more likely to define "science" as a process of learning than younger children were. Also, responses to our closed-ended questions about which activities are scientific reveal an important period in development between about 6 and 9, roughly corresponding to the ages in which we saw rapid development in children's scientific thinking abilities in the studies described in part II of the book. Further, we have some preliminary evidence in favor of our proposed connection between children's explicit understanding of science as a process of learning or knowledge change and their abilities to engage in various facets of scientific thinking.

Second, we want to repeat the point that our research did not aim to discover when children come to know the correct answer to the question about what science is. There may be no single answer to this question, and indeed there seems to be a tension even in adults' definitions between thinking of science as a set of topic areas (as captured by our Specificity code) and thinking of science as a process of inquiry or experimentation (as captured by our Learning and Other Process codes). This tension is reflected in the research in this area, with some studies examining how children come to acquire specific pieces of scientific knowledge, and others examining how children come to demonstrate particular kinds of scientific skills, as reviewed in chapter 1. Both the process of science and its specific disciplinary content are important for a full understanding of what science is; what we wish to highlight is that an explicit conception of what science is should also be part of this understanding.

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

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