THE INFLUENCE OF AN INHERENCE HEURISTIC ON SCIENTIFIC EXPLANATION

BY

ZACHARY S. HORNE

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Dissertation Committee:

Associate Professor Andrei Cimpian, Chair
Professor John Hummel
Professor Eva Pomerantz
Professor Emeritus William Brewer
Assistant Professor Daniel Hyde
Assistant Professor Jonathan Livengood

Abstract

Although scientific reasoning is usually deliberate, it may also be influenced by the intuitive processes involved in everyday reasoning. Here, I investigated whether explanatory heuristic processes biased towards inherence (that is, the tendency to oversample easily-accessible inherent facts; Cimpian & Salomon, 2014) influence scientific explanation. Across seven studies, children and adults (N = 1,455) explained outcomes of unfamiliar experiments in physics, biology, and chemistry predominantly in terms of inherent features. These inherence-biased explanations exhibited multiple signatures of heuristic reasoning (e.g., they decreased with age and time spent deliberating). Strikingly, I also found traces of this bias toward inherence in initial explanations of phenomena from the history of science (e.g., phlogiston explains combustion); these historical explanations were obtained via a survey of historians of science and were coded by hypothesis-blind researchers. These findings suggest that scientific explanation may be influenced by the same inherence heuristic that biases everyday explanation.

Table of Contents

Chapter 1: Introduction

In 1827, Scottish biologist Robert Brown was investigating the physical properties of flower pollen. During one series of studies in which he placed pollen suspended in water under a microscope, Brown observed that the pollen particles were rapidly zigzagging about. This was surprising to Brown, although similar events had been observed previously in other organic substances. The explanation for this phenomenon was thought to be well understood by other scientists – the motion of these molecules was due to a “vital force” present in living matter. Brown independently came to a similar conclusion: In Brown’s own words, “These motions … arose neither from current in the fluid nor from its gradual evaporation, but belonged to the particle itself.” However, as Brown pursued this line of reasoning further, he noticed inconsistencies between this explanation and the data. Brown noticed that when he tried the same experiment with inorganic particles he still observed this zigzagging motion. He realized that if the pollen’s motion was due to the presence of an inherent property of organic materials then inorganic materials—which were assumed to lack such a property—should not exhibit the same behavior. That is, upon reflection, Brown realized that the motion of the pollen was not due to something inherent to the pollen but instead to some other factor acting on the pollen. Brown attempted to rule out other alternative hypotheses that could explain the pollen’s motion, including checking whether his equipment was faulty, but it was not until the work of Einstein and Smoluchowski that it was discovered that the pollen’s motion was caused by the rapid motion of submicroscopic particles in a solvent pushing the pollen about.

Although Brown eventually realized his initial explanation of the motion of the pollen particles was incorrect, his first attempt at understanding their motion was closer to an explanatory guess than it was to a well-reasoned explanation. Brown is not alone in his reliance on explanatory guesses. There is considerable research that suggests that adults and children alike appear to rely on intuitive guesses to understand the world (see Driver et al., 2005 for a review). In an effort to understand the source of scientists' and laypeople's explanatory guesses, psychologists have begun to investigate the cognitive mechanisms and contours of explanatory reasoning (e.g., Gopnik, 1998; Keil, 2006; Khemlani & Johnson-Laird, 2011; Lombrozo, 2006; Murphy & Medin, 1985; Premack & Premack, 1996; Ross, 1977; Weiner, 1985).

Initial research on explanatory reasoning aimed to elucidate how people reconcile their intuitive understanding of the world with new evidence they acquire both inside and outside of the classroom (e.g., Brewer et al., 1998; Driver et al., 2005). More recently, however, research on explanatory reasoning has concerned how people evaluate explanations. For instance, researchers aimed to answer questions about the properties that good explanations have (e.g., Lombrozo, 2006; Lombrozo, 2007; Wilson & Keil, 1998), or what laypeople and practicing scientists think constitutes an intelligible explanation of a scientific observation (e.g., Waskan et al., 2014; Wilkenfeld et al., 2016).

While considerable advances have been made in understanding explanatory reasoning, researchers have now begun developing theoretical models aimed at describing the process of explanatory reasoning (e.g., Cimpian & Salomon, 2014; Khemlani & Johnson-Laird, 2013; also see Holyoak & Thagard, 1996; Rozenblit & Keil, 2002). Still, despite the centrality of explanatory reasoning to cognition, much remains unknown about its underlying cognitive mechanism.

In this dissertation, I aim to investigate and understand the cognitive processes underlying explanatory reasoning in the domain of science. Specifically, I investigate the hypothesis that there is a domain-general intuitive bias that affects scientific explanation. I will propose that that people's explanations of scientific phenomena will tend to exhibit an *inherence* (rather than *extrinsicness*) bias rooted in a heuristic cognitive process: I will argue that people's explanations will often invoke inherent facts about the observation to be explained rather than appealing to extrinsic factors operating on those entities. To return to the example of Brownian motion, I hypothesize that when confronted with a scientific phenomenon (e.g., the motion of pollen in a solvent) people will tend to explain the phenomenon in terms of the features of entities in the pattern (e.g., the inherent properties of the pollen) rather than facts extrinsic to the pattern (e.g., the motion of particles in the water acting on the pollen).

I motivate this prediction by considering several lines of research, including (1) findings that suggest general principles like heuristic processing, knowledge activation, and relational reasoning predict an inherence bias in explanation and (2) findings that people's explanations of everyday social patterns exhibit an inherence bias. After motivating this prediction, I discuss the results of nine studies that test the prediction that scientific explanation exhibits an inherence bias rooted in a heuristic cognitive process.

An inherence bias in explanation

Constructing explanations is an enormously complex computational task. Whether it be about the mysterious motion of pollen under a microscope or everyday social patterns, reasoners must access vast amounts of conceptual and relational knowledge encoded in semantic memory and combine relevant portions of that knowledge to produce a plausible account of why an event occurred. Yet, people often explain observations rapidly even in the face of anomalous experiences. But it is also clear that our initial explanatory guesses are often incorrect: even Robert Brown, a thoughtful botanist fell prey to his initial, but incorrect, explanatory guess.

However, the ways in which our explanatory hunches are incorrect shed light on the cognitive processes underlying how we construct explanations in a quick and effortless fashion.

Cimpian and Salomon (2014) recently articulated a theoretical account explaining the apparent effortlessness of explanatory reasoning. They argue that people produce initial explanations heuristically and that these explanations are characterized by their over-reliance on inherent facts (that is, facts about the way an object is in and of itself) rather than extrinsic facts (that is, facts that are dependent on other entities or broader interactions with the world).[1] Cimpian and Salomon argue that when reasoners attempt to make sense of some observation (e.g., that people drink orange juice for breakfast), they explain those patterns in terms of inherent facts about the entities involved in the pattern (e.g., the tanginess of orange juice) instead of extrinsic facts, like the context and history of the entities in the pattern (e.g., the promotion of oranges by crop growers).

Cimpian and Salomon posit that the overreliance of inherent features in explanation follows from general principles that govern heuristic processing (e.g., Evans, 2006; Kahneman, 2011; Stanovich & West, 2000), knowledge activation (e.g., Higgins, 1996), and relational reasoning (e.g., Halford et al., 1998). They hypothesize that when reasoners search for an explanation (e.g., Why do people drink orange juice for breakfast?), the main entities of the phenomenon to be explained (orange juice) become active in working memory. Active memories often provide relevant clues for the problem at hand (e.g., Higgins, 1996; Hummel & Holyoak, 2003). As a result, when people retrieve information relevant to generating an explanation, they may oversample information about these salient entities. As inherent properties are a part of the representation of these entities in semantic memory (e.g., McRae & Jones, 2013; Rosch &

[1] It should be noted, however, that inherent facts need not be centrally important to what an object is, but merely a property such that if it was changed, it would change the object.

Mervis, 1975), and are thus easily accessible to retrieval processes, Cimpian and Salomon hypothesize that these properties will often dominate what is retrieved to generate an explanation. Further, Cimpian and Salomon (2014) have argued that even when extrinsic facts are retrieved, they are more relationally complex (and thus cognitively complex) than inherent facts—extrinsic explanations entail an interaction between the environment and the pattern to be explained (e.g., Halford et al., 1998; Maybery et al., 1986; Waltz et al., 1999; Waltz et al., 2000). Thus, using extrinsic facts to produce an explanation is more cognitively demanding than using inherent facts to produce an explanation. Together, these general principles that govern heuristic processing, knowledge activation, and relational reasoning predict a bias towards inherence in explanatory reasoning.[2]

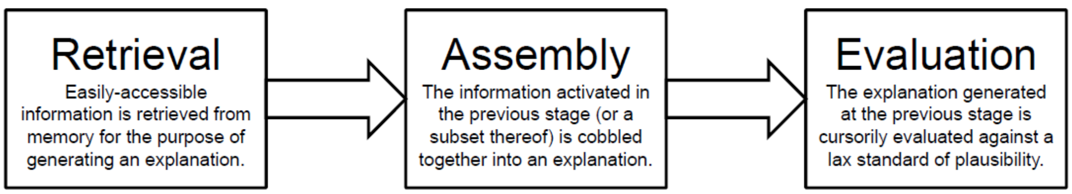

Figure 1. Schematic diagram of the inherence heuristic proposed in Cimpian and Salomon (2014).

The corresponding imbalance in the properties that are explanatorily relevant have important consequences for downstream processing: When the cognitive system goes on to assemble an explanation, its output will be correspondingly skewed towards explanatory intuitions that appeal to the inherent features of the relevant focal entities (e.g., the tanginess of orange juice). If the content retrieved initially is inherence-skewed, is part of the semantic representation of what is active in working-memory, and is less relationally complex (and thus

[2] This account is consistent with findings indicating that, for instance, children have early emerging skeletal frameworks that affect their understanding of physical phenomena. Indeed, the present proposal is derived from foundational findings of this sort rather than being in competition with them.

easier to use when producing an explanation) then the options of the assembly stage will be affected accordingly. Note that this claim does not entail that the cognitive architecture "prefers" inherent explanations *because* they are inherent. Instead, inherent facts are what *happen* to be accessible and this in turn leads to an inherence bias in explanatory reasoning; the bias towards inherence is a structural by-product rather than a preference of our mental architecture.

As noted, despite the obvious shortcuts built into the inherence heuristic process, the intuitions it generates may be generally adopted without much scrutiny, as is the case with the output of other heuristic processes: Extensive evidence suggests that humans often take the path of least cognitive effort, routinely "satisficing" rather than striving for accuracy (Evans, 2006; see also Simon, 1982). Thus, as long as the explanations generated by the inherence heuristic pass a threshold of plausibility, they may be readily accepted as correct and therefore shape one's general understanding of the world (beyond the particular circumstance or event that may have triggered the search for an explanation). In other words, the explanations produced using this heuristic process serve as a rough sketch of a complex causal mechanism whose answer is not already known, but this sketch suffices to satisfy reasoners' cognitive drive to explain.

However, the scope of the inherence heuristic is not all encompassing. Rather, it is limited to the domain where heuristics are necessary to efficiently navigate the world—that is, situations in which very complex problems require knowledge of difficult to access information (Kahneman & Frederick, 2002). For instance, the heuristic account of explanation is not aimed providing an account of situations in which one could simply retrieve answers from memory.

As a theoretical framework, the inherence heuristic proposal has proved extremely fruitful. In several sets of studies, researchers have found that people's explanations of, for instance, social patterns exhibit an inherence bias and this bias is linked to intuitive reasoning processes (e.g., Cimpian & Steinberg, 2014; Cimpian & Salomon, 2014; Hussak & Cimpian, 2015; Sutherland & Cimpian, 2015; Tworek & Cimpian, 2016).[3] However, although the inherence heuristic process has now been shown to be extensively involved in people's explanations of everyday phenomena, there has been notably little research on its potential influence in one major domain in which explanation plays a particularly significant role: scientific reasoning. Given the centrality of explanation to the history and philosophy of science, and the importance of science inside and outside the classroom, one open question is whether the cognitive mechanisms that influence everyday explanation, specifically those related to the inherence bias in explanation, likewise play a role in the formation of scientific explanations.

Do scientific explanations also exhibit an inherence bias and is this bias rooted in heuristic processes? From a normative perspective—that is, if we note how science ought to be conducted—the answer may seem obvious: Science relies on methods which are data driven, deliberative, and thus free of intuitions, heuristic processes, and the like. A priori, it seems unlikely that scientific explanation would be affected by inherence-biased heuristic processes. However, descriptively—that is, if we examine how science is actually conducted in the lab—the answer is less obvious: While science is often a deliberative enterprise conducted by people that spend many years of their lives attempting to make sense of complex phenomena, it isn't entirely divorced from the psychological constraints of its practitioners (e.g., Dunbar, 1997; Yu et al., 2014). Returning to the example of Brown, despite the centrality of deliberation to Brown's reasoning, his initial explanation—which appears to have been less well thought-out and more of

[3] These findings and the inherence heuristic proposal do not entail, however, that experience over the course of a person's life is irrelevant to the degree to which they might produce inherence biased explanations. Rather, the proposal is that overriding this prepotent bias towards inherent explanations requires effort, education, or cultural immersion. This claim is consistent with results from cross-cultural psychology showing that young children across cultures tend to endorse essentialist beliefs (a kind of inherent belief) but that, as they develop, the contours of these essentialist attitudes change in response to learning, education, and cultural input.

a gut explanatory guess—served as his theoretical starting place and informed much of his subsequent theorizing and experimentation.

My hypothesis is that science, as it is conducted, is likely to exhibit an inherence bias rooted in heuristic cognitive processes. In the next section of the paper, I discuss several empirical findings that suggest, under some circumstances, intuitive biases influence scientific explanation. In light of the evidence I will discuss, along with prior work showing an inherence bias in everyday explanation, I predict that scientific explanation will exhibit a pervasive inherence bias in situations in which scientists need to rely on intuitions to begin to make explanatory inroads as they attempt to explain novel complex scientific phenomena.

Intuitive biases in scientific explanation

> Many of the research studies [on science learning] refer to the way in which learned ideas are not applied by pupils when they reflect upon experience: they fall back upon their prior, lay ideas which have been retained alongside learned science. – Driver, Squires, Rushworth, and Woods-Robinson (2005).

Explanations of everyday phenomena are often shallow (Keil, 2006; Rozenblit & Keil, 2002), generated quickly (Cimpian & Salomon, 2014), and exhibit predictable biases in content (Sutherland & Cimpian, 2015; Tworek & Cimpian, 2016). However, much of the work documenting shortcuts and flaws in explanatory reasoning has not been performed in contexts that *necessarily* demand thoughtful deliberation, such as scientific inquiry. Thus, these findings do not directly indicate that intuitive biases will affect people's explanations of scientific phenomena. In fact, there are even reasons to believe that such biases will *not* be present in scientific reasoning, since a key part of what makes science so successful is its reliance on careful deliberation. There is a sizable literature on the mechanisms underlying the deliberative reasoning processes that go on in the lab, including how insight thinking and analogical inference lead to scientific breakthroughs (e.g., Holyoak & Thagard, 1996; Nersessian, 1992). For example, researchers have argued that analogical reasoning allowed scientists to begin to make sense of atomic structure: Predicating that the atom is like the solar system allowed scientists to begin to reason about a target domain (atoms) based on a more well understood source domain (the solar system). This relational mapping allowed scientists to make explanatory inroads into an unobservable and wholly mysterious target domain. Further, scientists typically describe scientific phenomena in the language of mathematics – an undoubtedly rational tool. Thus, one might be tempted to conclude that scientific reasoning is by and large, if not entirely, a rational, deliberative exercise unaffected by cognitive biases and psychological constraints.

However, even in the context of a deliberative exercise such as analogical reasoning, for example, psychological constraints unambiguously affect how practicing scientists explain phenomena. For example, Dunbar (1997) found that even in world-class biology labs, analogies tended to be "local", relying on featural similarity of a source and target, rather than abstract (such as the solar system – atomic structure analogy) (Chen & Klahr, 1999). Findings such as these suggest that even in a characteristically rational enterprise, cognitive limitations can impose constraints on the kinds of explanations scientists are likely to produce, at least on a first pass.

In fact, there is substantial evidence that intuitive processes shape both children's and adults' explanations of scientific phenomena. Although a full overview of this literature is beyond the scope of this dissertation (see Driver, Squires, Rushworth & Wood-Robinson, 2005, for a review), I will now discuss some influential findings that show how biases make their way into people's understanding of scientific phenomena. These findings suggest that when children and adults first begin to understand scientific phenomena, they bring to bear intuitive theories and rely on cognitive shortcuts to deal with the complexity of the phenomena they are attempting to understand. I begin by discussing several prominent findings in detail and then provide a brief summary of several other findings from this literature.

Some of the most influential work on how intuitive biases shape scientific understanding comes from McCloskey and Kohl (1983). McCloskey investigated how naïve theories shape high school and college students' explanations and predictions about the dynamics of objects. McCloskey showed that people explain dynamics in terms quite similar to the theory of impetus – a revised theory of Aristotelian dynamics according to which, when an object is thrown, for example, it acquires a kind of internal force from the thrower that ensures the object will continue to move. As the object soars through the air, the impetus within the object dissipates causing it to fall. Strikingly, high school and college students believe that objects have a kind of impetus instilled in them by the actions of an agent hurling or swinging those objects. For example, McCloskey had participants judge the motion of an object after it was swung around in a circle on a rope. He found that, rather than believing the object will proceed in a straight line once released, people expected that the object will continue to curve. When asked to explain this prediction, people appealed to an impetus-like idea that the circular motion had been imparted to the object (also see Driver & Erikson, 1983; Driver et al., 2005; Halloun & Hestenes, 1985).

Intuitive theories affect people's explanations and understanding of scientific phenomena not just in adulthood but also early in life. For example, Vosniadou and Brewer (1992) investigated children's conceptions of the Earth's shape and how they change over the course of elementary school (also see Vosniadou & Brewer, 1987). Vosniadou and Brewer proposed that children have naïve theories about the shape of the Earth – namely, that it is flat. However, when children learn that the Earth is in fact a sphere floating in space, they are left to reconcile their intuitive theory with the new information that they've acquired. In one particularly striking study of how children reconcile these conflicting ideas, Vosniadou and Brewer found that, for example, children would draw the Earth as a sphere with only the top "flat part" of the Earth containing the properties we associate with the Earth (viz. having the sun, moon, stars, oceans and so forth – see Figure 1, Panel C below).

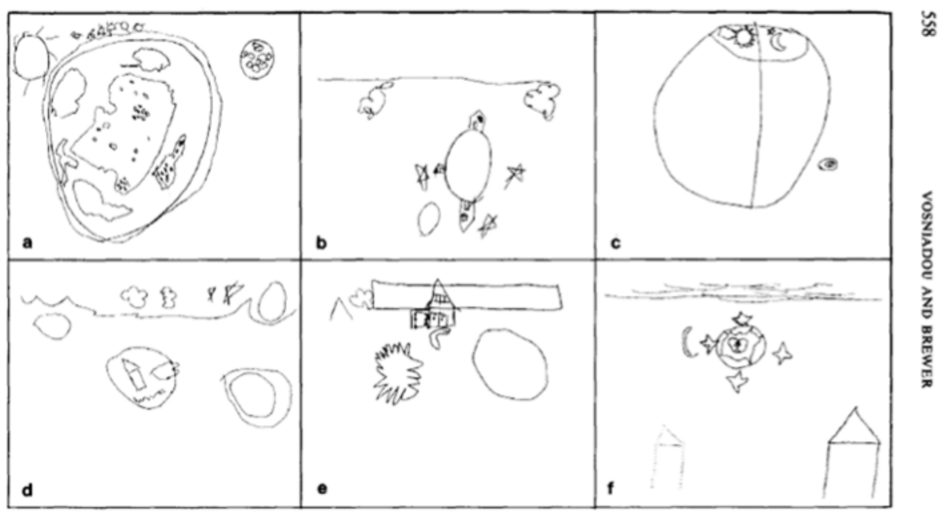

Figure 2. Drawings of the Earth, moon and stars from children in Vosniadou & Brewer (1992). The figure depicts children's attempt to reconcile their intuitive notions that the Earth is flat with new information they've acquired in school that the Earth is a sphere floating in space.

Beyond these specific examples in the domain of physics, researchers have shown that intuitive biases shape people's understanding of light, sound, and heat (Driver et al., 2005) among many other phenomena.

The idea that intuitive notions influence children's explanations of scientific phenomena extend beyond the domain of physics. Indeed, there has been a significant amount of research on intuitive but mistaken biological (e.g., Gelman, 2003; Hatano & Inagaki, 1994) and chemical (e.g., Leslie, 2013) conceptions. For example, Inagaki and Hatano demonstrated that children understand and interpret living entities in terms of vitalistic causality – a "vital power" attributed to biological entities which is conceptualized by children as an essential substance or energy (see Gelman, 2003).

In sum, there is considerable evidence explanations of scientific phenomena are influenced by intuitive understanding of the world among both adults (not just naïve adults but also practicing scientists) and children. Although much of this literature has focused on domain-specific intuitive biases, I hypothesize that there is a domain-general intuitive bias that affects scientific explanation. Although scientific explanation may often be a deliberative enterprise, vestiges of the intuitive processes underlying everyday explanation may systematically affect the beliefs people form when they begin to understand scientific phenomena. In the remainder of the paper, I directly test this proposal in a series of studies examining adults and children's explanations of scientific phenomena.

Predictions for Studies 1 - 8

I conducted 8 studies to examine the proposed influence of a domain-general heuristic process biased towards inherence on scientific explanation. In Studies 1 and 2, I examined whether scientific explanation among naïve adults exhibits an inherence bias. I predicted that participants would explain scientific observations in chemistry, physics, and biology in terms of the inherent features of the entities in the observation, even when plausible extrinsic facts are available.

In light of this initial evidence, I then sought to test the source of the inherence bias. Specifically, I hypothesized that naïve participants' inherent explanations of scientific phenomena are rooted in heuristic cognitive processes. I tested this hypothesis by examining whether the inherence bias exhibits several signatures of heuristic thought.

First, I examined the influence of salience on the inherence bias. If extrinsic facts are less likely to be used in generating explanations of scientific phenomena because they are less salient, then increasing the salience of extrinsic facts should increase their use (Study 3).

Next, I examined the time course of participants' decisions about what type of fact (i.e., inherent or extrinsic) was more likely to explain a phenomenon. Heuristics are comparatively fast, automatic, and intuitive; thus, if inherent explanations stem from a heuristic process, then participants who deliberate longer should be less likely to provide inherent explanations (Studies 4 and 5).

The final signature I examined was how the inherence bias in explanation changes over the course of development. Younger children have fewer cognitive resources and rely on heuristic processes more than older children (Stanovich & West, 2000). Consequently, I predicted that, if extrinsic explanations are generated less often because they are more cognitively complex, then older children ought to produce extrinsic explanations more often than younger children (Studies 6 and 7).

Finally, having provided converging evidence of an inherence bias from 7 studies with naïve participants, I examined how understanding this bias would shed light on theory change in the history of science by surveying historians of science (Study 8).

General Methods for Studies 1 - 7

Methods

Participants. For Studies 1 – 5 participants were recruited on Amazon Mechanical Turk. A HIT was posted on Mechanical Turk advertising payment for participation in an academic survey. People who chose to participate were then directed from Mechanical Turk to a page that confirmed they could participate in the study. People who had previously participated in similar studies were blocked from participation using the software Turkgate. Participants were then directed to an informed consent form and then finally to the instructions for a given study. For Studies 1 – 6 participants' responses were collected using Qualtrics survey software.

In Studies 6 and 7, children ages 6 – 9 were recruited in a Midwestern college town to participate in a study about concepts and categories. Their responses were collected using pen and paper questionnaires filled out by an experimenter.

Attention and memory checks. Several measures were taken to confirm that participants (both adults and children) were paying attention while they participated in the study. One of these measures checked participants' attention by asking them to select certain responses throughout the course of an experiment (e.g., "Select 'Agree' to confirm that you are paying attention"). At the end of each study, participants were also asked to report honestly if they took the study seriously and paid attention, making it clear that there would be no negative consequences if they said they did not. Participants who missed any of the attention checks or who indicated they were not paying attention were excluded from analyses. In our studies with children, we confirmed that they understood the story they were read by having them repeat central details of the story back to the experimenter.

Materials. For all studies with naïve participants, vignettes about scientific phenomena were developed to be both age- and skill-appropriate; I designed the materials so that participants could generate explanations for scientific phenomena even in the absence of having detailed knowledge about how focal entities in the pattern. Participants – both adults and children – had little trouble coming up with explanations for patterns they just learned about.

Coding. Participants responses were coded for inherence and extrinsicness in studies where participants provided open-ended responses (all studies except Study 4). Across studies with adult participants, inter-rater agreement between two-coders ranged from 86% to 91% for inherence and 89% to 92% for extrinsicness, suggesting that coders could reliably categorize an explanation as inherent or extrinsic. For Studies 6 and 7, a hypothesis-blind researcher coded children's responses for inherence and extrinsicness. A second coder then recoded one third of responses, agreeing on 95% of inherence codes and 95% of extrinsicness codes.

Analytic Strategy

Data analysis. Rather than performing null-hypothesis significance testing, I used Bayesian data analysis. Specifically, I computed Bayesian Credible Intervals (CIs) (Gelman et al., 2014) rather than Confidence Intervals. Additionally, I report Bayes factors (e.g., Dienes, 2014; Wagenmakers, 2007), which are computed using the Savage-Dickey method (Rouder et al., 2009). The motivation for using Bayesian data analysis is that it allows the researcher to directly infer the probability of a hypothesis given the data (i.e., $P(\mathrm{H}|\mathrm{D})$) rather than having to test whether the null hypothesis can be rejected to (indirectly) determine whether the proposed hypothesis holds. The probability of the hypothesis given the data is computed via the *likelihood* $P(\mathrm{D}|\mathrm{H})$, the *prior* $P(\mathrm{H})$, and the probability of the data $P(\mathrm{D})$. The *likelihood* is the probability of the data given the hypothesis and the *prior* is initial information about the hypothesis.

Bayesian estimation. I performed Bayesian estimation using the probabilistic programming language Stan (Gelman et al., 2014), which simulates the posterior using a Hamiltonian Monte Carlo method (Duane, Kennedy, Pendleton, & Roweth, 1987; Neal, 2011) and its extension, the No-U-Turn Sampler (NUTS) (Hoffman and Gelman, 2014). These methods use simulation techniques to take random draws from the posterior distribution to estimate the most credible values for a parameter θ. As I noted above, this technique yields Credible Intervals rather than Confidence Intervals. Credible Intervals are kernel density plots of the most credible values for a parameter θ given the prior probability of θ and the probability of the data given θ. Values closer to the mean of the Credible Interval are more credible than values at the tails of this interval (note that this is *not* the meaning of a Confidence Interval). In studies in which I performed Bayesian estimation (i.e., Studies 3 – 8), error bars and shaded regions represent 50% Credible Intervals (as recommended by Gelman, 2016) and plots display the marginal density of the predictor variable on the x-axis (known as a "rug"). For these studies, I also report 95% Credible Intervals in the text.[4]

Estimation diagnostics. When performing Bayesian estimation is important to take several precautions to ensure that the parameter estimates are stable. For all models reported henceforth, I performed diagnostics checks as recommended by Gelman and colleagues (2014).[5]

Bayes factors. Throughout the paper, I also report Bayesian hypothesis tests by computing Bayes Factors. Bayes Factors are the likelihood of the data under the alternative hypothesis divided by the likelihood of the data under the null hypothesis. Larger Bayes Factors indicate that the data are more likely under the alternative hypothesis than the null hypothesis, as those hypotheses are specified.[6] For example, a Bayes Factor of 3 indicates that the data are three times more likely under the alternative hypothesis than they are under the null hypothesis.

---

[4] Credible Intervals that span zero can still permit the inference that there is credible effect. This is because Credible Intervals still allow one to infer that the most credible value is at the center of the probability distribution. Of course, when these Intervals are incredibly wide it would be unwise to assume that specific values are particularly credible even if, by comparison, they are the most credible in the distribution.

[5] First, I confirmed that the steps in a chain were not autocorrelated after warm-up by examining the number of "Effective Steps" taken by the workers exploring the posterior distribution – larger number of effective steps (> 100) indicate that the subsequent step a worker takes is uncorrelated with the previous steps the worker had taken. When steps are highly correlated, workers will not explore the entire posterior distribution, biasing the parameter estimates of a model. Second, I confirmed that the chains exploring a posterior converged on similar parameter estimates (as indicated by the value $\hat{r}$) – this is to say, "workers" exploring the posterior distribution settled on similar parameter estimates. Third, I confirmed there were no or minimal numbers of divergent steps after warm up because large numbers of divergent steps can also lead to unstable parameter estimates. Finally, I confirmed that the posterior of the estimated data resembled the distribution of the original data.

[6] As a rule of thumb, Bayes Factors larger than 3 are considered "some" evidence for the alternative hypothesis.

Because Bayes Factors can be strongly influenced by selection of priors, I performed prior robustness checks to confirm my conclusions were not materially affected by my prior selection.

Chapter 2: Is there is an inherence bias in scientific explanation?

Prior work suggests that people over-rely on inherent facts when they explain everyday patterns in the world (Hussak & Cimpian, 2015; Sutherland & Cimpian, 2015; Tworek & Cimpian, 2016). We also see hints of the influence of intuitive biases on naïve adults and practicing scientists' explanations of novel scientific phenomena (Driver et al., 2005; Dunbar, 1997). On the basis of these results, I predicted that a heuristic-based inherence bias would emerge even in the context of scientific explanation. I tested this prediction by examining how naïve participants—that is, people who best mirror the epistemic position of early scientists unfamiliar with the detailed interactions of the entities in the pattern they hope to explain—explain fictitious scientific phenomena (Study 1) and unfamiliar phenomena culled from the history of science (Study 2).

Study 1: Naive participants' explanations of scientific phenomena

In Study 1, I investigated whether people's explanations of fictitious scientific phenomena in chemistry, biology, and physics overweigh inherent facts.

Method

Participants. Participants in Study 1 were 288 Amazon Mechanical Turk workers who were paid $0.50 for participating in the study.

Materials and Procedure. In Study 1, participants were told that they would read about a series of scientific discoveries and that their task was to come up with the most plausible explanation for why these events occurred. Participants read three vignettes describing fictitious scientific phenomena from one of three disciplines: chemistry, biology, or physics. The three vignettes were selected randomly from a larger pool of six vignettes constructed for each domain (see Table A of the Supplementary Materials for additional examples). After reading each vignette, participants were asked to explain why the event occurred ("Explain why this happened. Please do your best") by typing an explanation into a text box.

Below is a sample vignette:

> Chemists in a lab high in the Colorado Rockies were investigating the possibility of storing hydrogen atoms in lithium nitride crystals. In a strange turn of events, these chemists found that when they attempted to store the hydrogen atoms in the lithium nitrides crystalline structures, it led to a result that completely defied their expectations. Instead of the lithium nitride structure storing 15% of hydrogen atoms, the crystalline structures stored 2% of hydrogen atoms. This is despite the fact that these scientists had more than 40 years of experience between them working with lithium nitride.

The vignettes were designed to provide a conservative test of my hypothesis: First, the vignettes mentioned a plausible extrinsic factor that participants could use in their explanations. This feature of the vignettes ensures that, if participants explain a pattern inherently, it is not for lack of extrinsic explanation options. Second, the scientists' observations were described as anomalies (e.g., the result "completely defied their expectations"). This, again, makes it more difficult for participants to generate an inherent explanation: If an object is behaving anomalously it is unlikely that its behavior would be due to a reliable (inherent) property of the object—recall that inherent properties are properties of the object in and of itself. (see the Supplementary Materials for a study validating this assumption).

Within each discipline, the six vignettes were constructed using three consistent templates (3 templates × 2 tokens = 6 vignettes per domain; see Table 1 below) that varied along a number of dimensions (e.g., whether the extrinsic fact—namely, location—was mentioned before or after the phenomenon to be explained).

Table 1

The vignette templates used in Study 1.

Template 1	[Disciplinists] in [location] were investigating [hypothesis or phenomenon]. In a strange turn of events, these [disciplinists] found that when they attempted [perform action], it led to a result that completely defied their expectations. Instead of [expected outcome], [unexpected outcome]. This is despite the fact that these scientists had more than 40 years of experience between them [performing action].
Template 2	A group of experienced [disciplinists] were looking to [perform action]. These [disciplinists] were working in [location]. Typically, [performing action] leads to [expected outcome]. Surprisingly, when these [disciplinists] [performed action] an [unexpected outcome] occurred.
Template 3	[Disciplinists] have discovered something astonishing about [chemical, biological, or physical phenomena], which they observed/experimented with in [location]. [Disciplinists] observed [chemical, biological, or physical phenomena]. This [event] puzzled chemists because they initially theorized that [objects in event] would behave differently.

After completing this task, participants answered a series of demographic questions about their college education, religion, and gender.

Results and Discussion

My primary interest was in the descriptive question of whether people's explanations of scientific observations would exhibit an inherence bias. To test this hypothesis, participants' explanations were coded for inherence and extrinsicness. For example, an explanation that was inherent but not extrinsic, would receive a code of 1 for inherence and 0 for extrinsicness. These categories were not mutually exclusive: Participants could provide more than one type of explanation, in which case their response would receive a code of 1 for inherence and 1 for extrinsicness. I then averaged participants' explanations across vignettes into composite scores representing the proportion of explanations that were inherent and the proportion of explanations that were extrinsic.

I was interested in examining the presence of an inherence bias in scientific explanation. To examine this prediction, I compared the proportion of inherent explanations to the proportion of extrinsic explanations (see Figure 4). As I predicted, a Bayesian paired "*t* test" revealed that participants explained scientific observations more inherently (*M* = .56, *SD* = .29, *SE* = .02) than extrinsically (*M* = .41, *SD* = .29, *SE* = .02), BF = 5,892, Error = 5.88e -9. I observed this bias in chemistry (Inherent: *M* = .53, *SD* = .32; Extrinsic: *M* = .43, *SD* = .32), biology (Inherent: *M* = .62, *SD* = .28; Extrinsic: *M* = .35, *SD* = .26), and physics (Inherent: *M* = .54, *SD* = .31; Extrinsic: *M* = .45, *SD* = .29), all BFs > 3. Further, participants explained more inherently than we would expect by chance alone, BF = 45.34, Error = 9.775e -7. Of the 18 vignettes, 13 vignettes exhibited a bias in the predicted direction (i.e., towards inherence), BF > 3. Of the 284 participants, 169 participants' explanation averages exhibited a bias in the predicted direction, BF > 3.

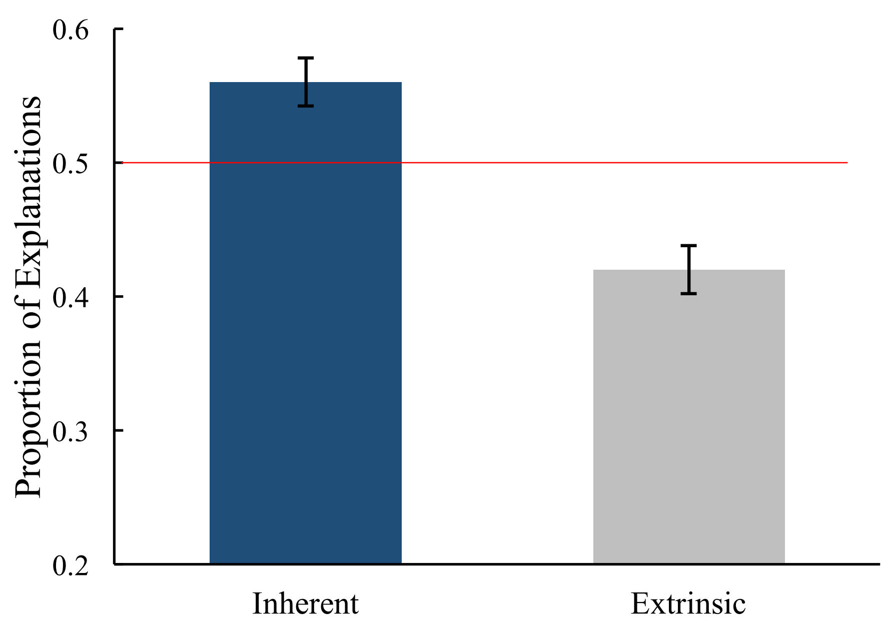

Figure 3. The mean proportion of inherent and extrinsic explanations for scientific observations in Study 1. Error bars represent ±1 standard error

In summary, the results of Study 1 are consistent with the prediction that naïve participants' scientific explanations exhibit an inherence bias. People's explanations of scientific phenomena in chemistry, biology, and physics relied on inherent facts more than extrinsic facts, even though a plausible extrinsic fact was available to be used in an explanation.

Study 2: Naive participants' explanations of phenomena from the history of science

Although describing fictitious scientific phenomena allowed me to exercise maximum control over the content of the stimuli, the vignettes used in Study 1 were relatively artificial. Thus, in Study 2 I asked participants to explain real scientific observations that played a role in the history of science. For example, the observation that pollen moved in solvents was thought to show that the pollen was still alive, that is, until it was revealed that submicroscopic particles were pushing the pollen around. I predict that naïve participants will gravitate toward inherent explanations much as Brown did when he first peered through his microscope.

Method

Participants. Participants in Study 2 were 252 Amazon Mechanical Turk workers who were paid $0.25 for participating in the study.

Materials and Procedure. The procedure of Study 2 was similar to that of Study 1. However, in Study 2, participants explained two scientific observations from one of three disciplines: chemistry, biology, or physics. Only six vignettes were constructed for this study (two per discipline; see Supplementary Materials Table A for additional examples). For example, participants read the following vignette taken from the history of chemistry in which Lavoisier was investigating Aristotle's hypothesis about how water when heated begins to transmute into soil:

> A chemist conducted the following experiment: They boiled distilled water in a sealed vessel for 100 days and found that solid sediment appeared at the bottom of the vessel. This result was quite striking to the chemist.

As in Study 1, after reading each vignette participants were asked to do their best to explain why the event occurred.

Results and Discussion

I again coded participants' responses for inherence and extrinsicness. Just as in Study 1, each response was coded on each of these dimensions (0 = dimension absent, 1 = present). I then averaged participants' explanations across vignettes into composite scores representing the proportion of explanations that were inherent and the proportion of explanations that were extrinsic.

As in Study 1, I performed a Bayesian paired *t* test which revealed that participants explained observations culled from the history of science more inherently (*M* =.64, *SD* = .36, *SE* = .02) than extrinsically (*M* = .38, *SD* = .37, *SE* = .02), BF = 1.043e +6, Error = 3.964e -12. I observed similar effects in chemistry (Inherent: *M* = .71, *SD* = .34; Extrinsic: *M* = .32, *SD* = .36), biology (Inherent: *M* = .62, *SD* = .39; Extrinsic: *M* = .38, *SD* = .38), and physics (Inherent: *M* = .59, *SD* = .36; Extrinsic: *M* = .44, *SD* = .38), all BFs > 3. Also, participants explained more inherently than we would expect by chance alone, BF = 5.705e +6, Error = 2.331e -13 (see Figure 4). Of the six vignettes, four exhibited a bias in the predicted direction (i.e., towards inherence), BF < 3. Of 254 participants, 210 participants' explanation averages exhibited a bias in the predicted direction, BF > 3.

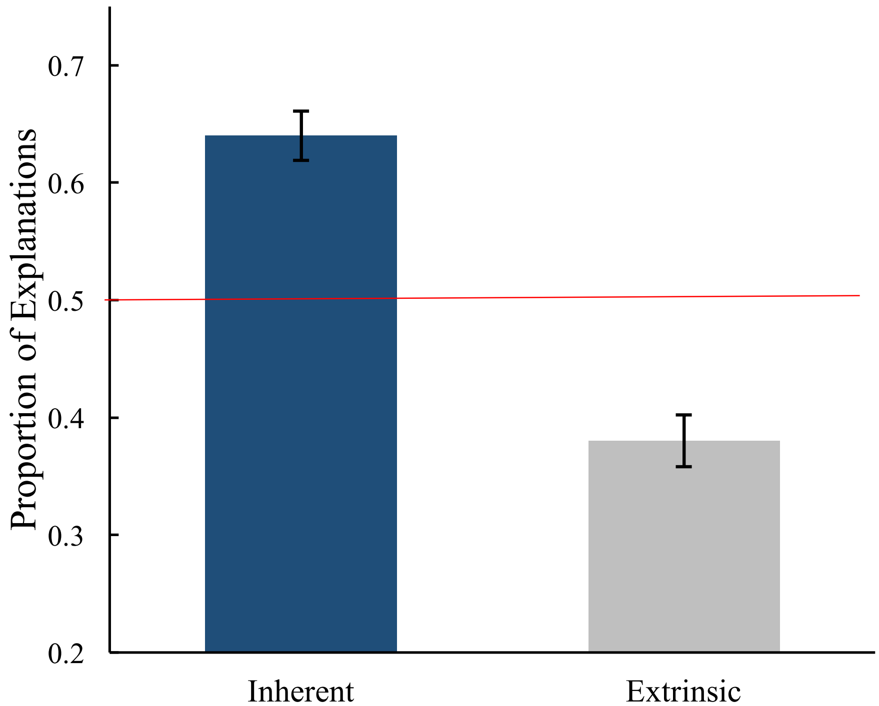

Figure 4. The mean proportion of inherent and extrinsic explanations for scientific observations culled from the history of science in Study 2. Error bars represent ±1 standard error

In summary, the results of Study 2 support my hypothesis that naïve participants' explanations of scientific observations culled from the history of science exhibit an inherence bias. This suggests that the inherence bias revealed in Study 1 can generalize real episodes from the history of science.

Interim Summary

The data from Studies 1 and 2 suggest that when people first confront and attempt to explain a novel scientific phenomenon, they do so more inherently than extrinsically. I observed this inherence bias in the domains of chemistry, biology, and physics. This finding, of course, does not entail that in *every* scenario what is salient is inherent nor that people will *always* explain more inherently than extrinsically. Rather, these results suggest that when people first begin to explain a complex novel observation that requires reliance on difficult-to-access information (Kahneman & Frederick, 2002), they do so more inherently than extrinsically. This finding supports the hypothesis that a domain-general inherence heuristic (Cimpian & Salomon, 2014) influences scientific explanation. This is striking in and of itself given that, as I have noted, scientific explanation is a paradigmatic example of a deliberative enterprise.

While the data thus far are consistent with the influence of an inherence heuristic on scientific explanation, I more directly tested the heuristic source of participants' inherence bias by examining whether it exhibits key signatures of heuristic thinking.

Chapter 3: Is the inherence bias in scientific explanation rooted in heuristic thinking?

Study 3: The influence of salience on inherent explanations

Prior research suggests that the inherence bias in everyday explanation is due to the fact that when people explain an observation, the main entities of the phenomenon to be explained (e.g., orange juice) become active in working memory. Because active memories are often assumed to be a cue for the problem at hand (e.g., Higgins, 1996; Hummel & Holyoak, 2003), inherent facts will dominate what is salient when generating an explanation. If inherent facts are more salient when participants first attempt to explain a pattern and this leads to an inherence bias then increasing the salience of extrinsic facts should lead to a decrease in the proportion of inherent explanations of a pattern.

Methods

Participants. Participants in Study 3 were 486 Amazon Mechanical Turk workers who were paid $0.50 for their participation.

Materials and Procedures. In Study 3, participants read eight vignettes describing scientific observations from one of three disciplines: chemistry, biology, or physics. These vignettes were selected randomly from a larger pool of 24 (18 from Study 1 + 6 from Study 2).

After each vignette, participants were asked to complete the following sentence to the best of their ability: “Something about the ______ explains the outcome of the experiment.” Participants were assigned to either the Forced-Choice condition or the Production condition (between subjects). In the Forced-Choice condition, participants were instructed to choose between two words representing an inherent and an extrinsic explanation. For example, if participants read about the chemist boiling water in a vessel, they were provided with the options “Water” and “Vessel” representing inherent and extrinsic explanations, respectively. In the

Production condition, participants were asked to type in a word or words that best completed the sentence. For example, participants might write “water” or “vessel.” Just as in Studies 1 and 2, the words participants produced were coded as being an inherent explanation or an extrinsic explanation. In the rare instance that a participant wrote something indicating both an inherent and extrinsic explanation, their response was coded as extrinsic (that is, in the direction contrary to my hypothesis).

I predicted that—consistent with influence of an inherence heuristic on scientific explanation—making extrinsic facts more salient would lead to fewer inherent explanations. That is, I predicted that participants assigned to the Forced-Choice condition, the condition in which the extrinsic fact is explicitly stated as a possible explanation, would provide more extrinsic explanations than those assigned to the Production condition. However, because inherent explanations are still less relationally complex than extrinsic explanations (Cimpian & Salmon, 2014), I also predicted that there would be an inherence bias in participants’ explanations regardless of condition. This is because, as I noted, even if an extrinsic factor is salient to participants, they still need to determine *how* it is that this factor interacts with the entities in the pattern to explain the observation—this increases relational complexity. Inherent explanations, in contrast, require no such working out because no additional explanatory variables are added to the equation.

Results and Discussion

In Study 3 I fit a Bayesian logistic mixed-effects model predicting inherent explanation responses (1 = Inherent, 0 = Extrinsic), on the basis of a participant’s condition assignment (Production Condition = 1, Forced-Choice Condition = 0). I fit a model estimating the effect of condition on inherent explanations, and included random intercepts for subject and vignette structure.

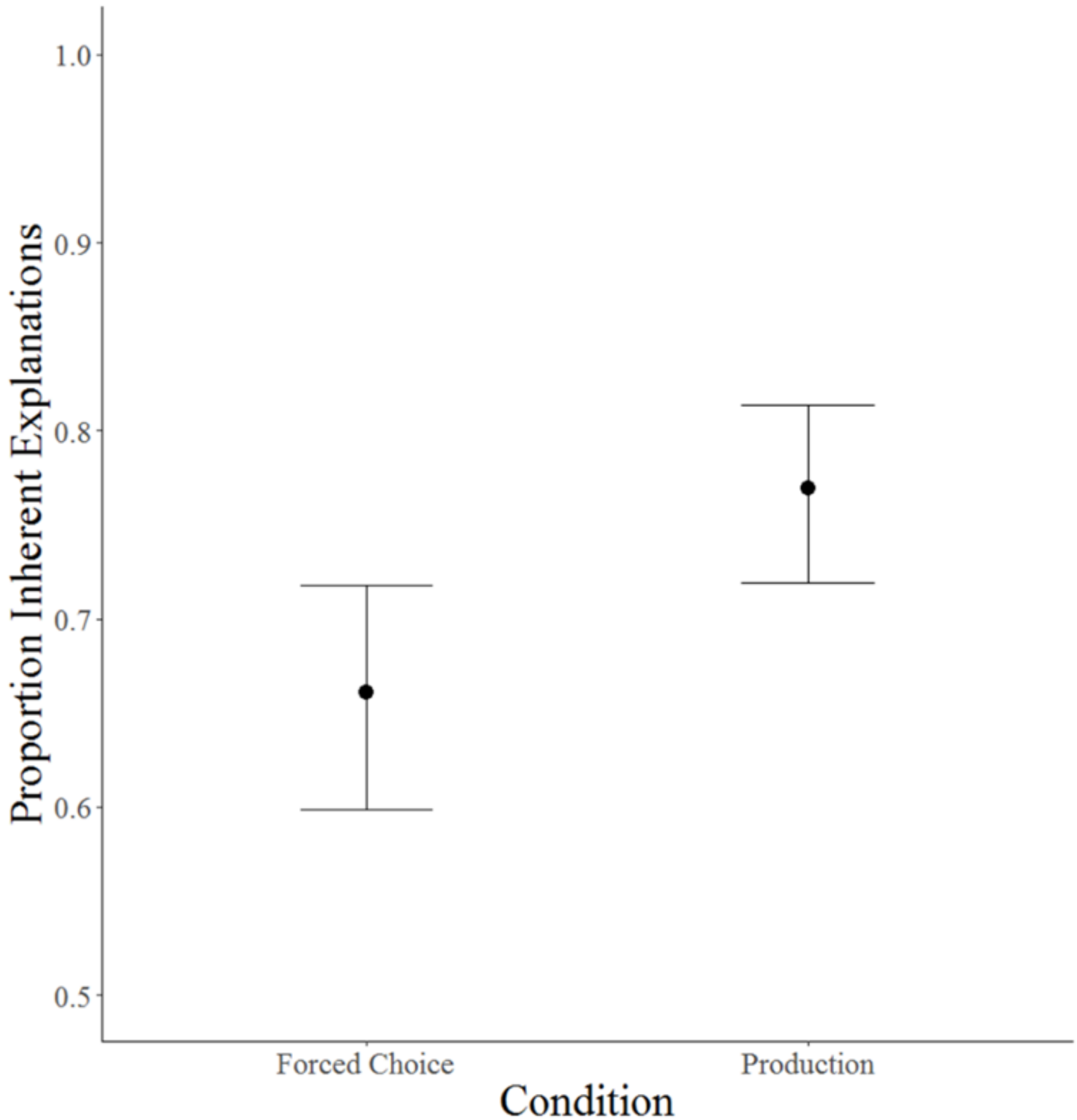

Figure 5. The marginal likelihood of participants in the Forced Choice and Production conditions providing inherent explanations in Study 3. Error bars represent 50% Credible Intervals.

Table 2

Bayesian logistic mixed-effects regression predicting proportion of inherent explanations on the basis of condition in Study 3.

Population-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
Intercept	.68	.69	4,107	−.70	2.18	1
Condition	**.54**	**.11**	**11,563**	**.31**	**.76**	**1**
Group-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
SD(Subject)	.88	.07	6,115	.76	1.01	1
SD(Vignette)	.66	.13	5,188	.46	.95	1

Note. Reference group is the Forced-Choice Condition. Population-Level effects are also known as “Fixed-Effects” and Group-Level Effects are also known as “Random Effects” (see Gelman et al., 2014).

I predicted that the Production condition would lead to more inherent explanations than the Forced-Choice condition because in the Production condition, extrinsic explanations are not as salient. Consistent with this hypothesis, I observed a credible difference between participants’ performance in the Forced-Choice and Production conditions. Participants in the Production condition were more likely than participants in the Forced-Choice condition to provide an inherent (vs. extrinsic) explanation, $\beta = 0.54$, 95% CI [.32 to .76]. Further, as I predicted, in both conditions participants still exhibited an inherence bias, suggesting that salience alone cannot account for the inherence bias in explanatory reasoning—even when an extrinsic fact is accessible, a participant still must work out how it is that the environment and the entity in the pattern interact to account for the observation, BF > 3. Of the 24 vignettes, 23 vignettes exhibited a bias in the predicted direction (i.e., towards inherence), BF > 3. Of 485 participants, 408 participants’ explanation averages exhibited a bias in the predicted direction, BF > 3.

In sum, Study 3 provides two important pieces of evidence for the influence of an inherence heuristic on scientific explanation. First, using two different tasks than those used in Studies 1 and 2, I found that participants still explained scientific phenomena more inherently than extrinsically. This confirms that our prior results were not due to idiosyncrasies of the task participants performed in Studies 1 and 2. Second, and consistent with the influence of a domain-general inherence heuristic on scientific explanation, I found that manipulating the salience of extrinsic facts lowered participants' tendency to endorse inherent explanations.

Study 4: The influence of deliberation time on inherent explanations

It is well known that heuristic thinking tends to be relatively automatic and effortless rather than deliberative and effortful (Stanovich & West, 2000). In Study 4, I tested whether the inherence bias in scientific explanation would be weakened among those participants who reflected for a longer amount of time on the explanandum. According to the inherence heuristic account, what comes to mind first when participants explain an observation are inherent facts about the entities in that observation rather than extrinsic facts about them. Thus, I predicted that participants who deliberate longer about the phenomenon to be explained should explain this phenomenon less inherently because they are more likely to retrieve (and thus use) extrinsic facts. This prediction is also consistent with prior findings that relationally complex reasoning—of the sort often involved in extrinsic explanations—requires longer deliberation times (Waltz et al., 2000).

Methods

Participants. Participants in Study 4 were 125 Amazon Mechanical Turk workers who were paid \$0.60 for participating in the study. An additional six participants were tested but excluded because they indicated that they didn't pay attention during the study.

Materials and Procedures. Participants read 8 vignettes from chemistry, biology, and physics, for a total of 24 vignettes in a within-subjects design. The vignettes used in Study 4 were identical to the vignettes used in Study 3 and the design of Study 4 was otherwise identical to the Forced-Choice condition of Study 3. Participants simply chose between an inherent and extrinsic explanation of a scientific observation. I recorded participants' response times using Qualtrics survey software to test the predicted relationship between the inherence bias in explanation and deliberation time.

Results and Discussion

As in Study 3, I fit a Bayesian logistic mixed effects model predicting inherent explanation choices (1 = Inherent, 0 = Extrinsic) based on the amount of time (log-transformed, Kliegl et al., 2010) participants took to make their choice. Trials that were extremely fast (less than 5 seconds to read and choose an explanation) or extremely slow (greater than 100 seconds) were removed before analyzing the data leading to the exclusion of 6.2% of trials. The model included random intercepts and random slopes (for reaction time) over subject and vignette, allowing the relationship between time and explanation choices to vary at the level of subject and vignette.

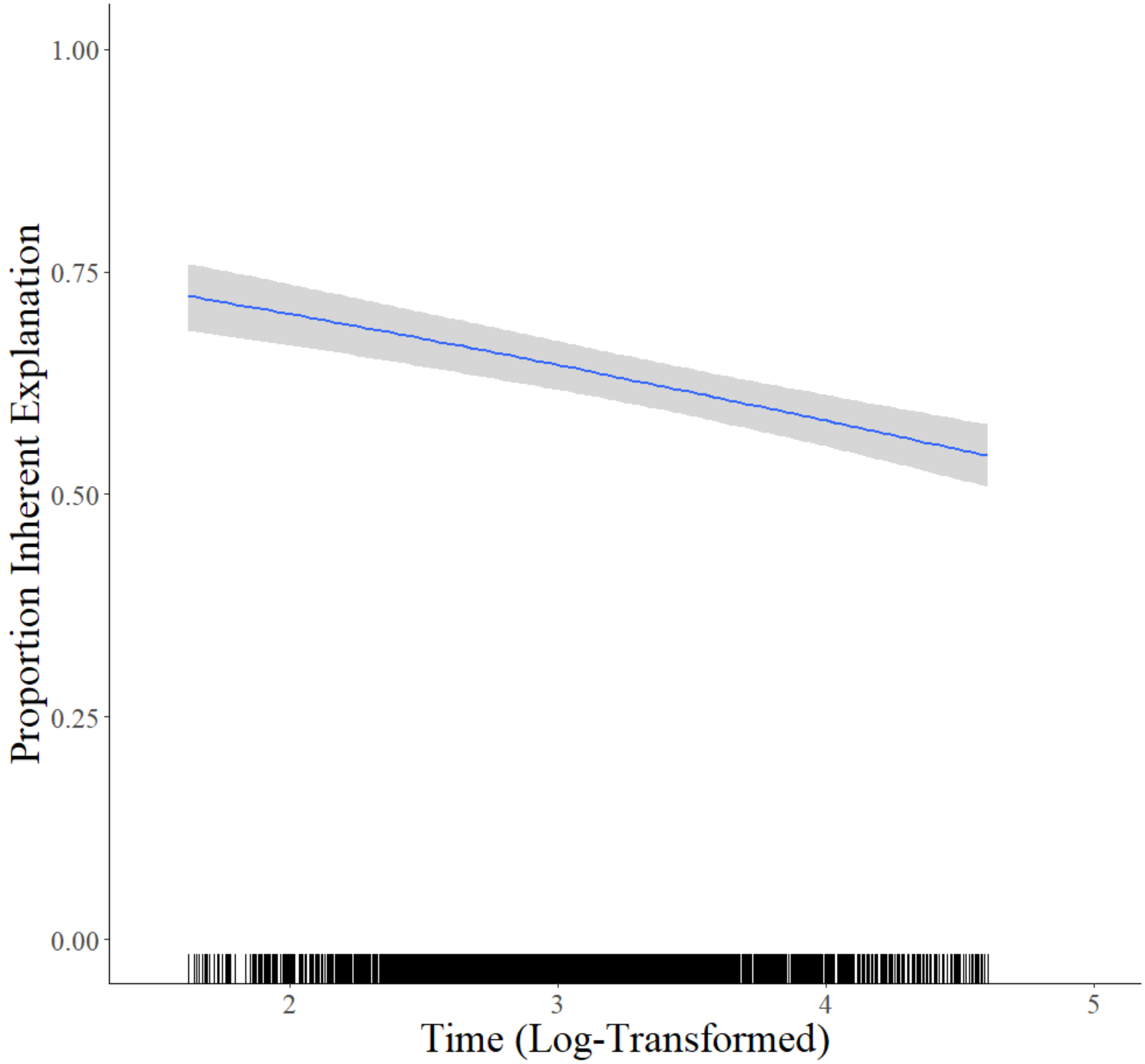

Figure 6. The marginal likelihood of participants making inherent explanations across response time in Study 4. Shaded region represents 50% Credible Intervals.

Table 3

Bayesian logistic mixed-effects regression predicting proportion of inherent explanations on the basis of decision time in Study 4.

Population-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
Intercept	1.39	.46	7,737	.5	2.30	1
Time	**−.26**	**.12**	**10,581**	**−.50**	**−.03**	**1**
Group-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
SD(Subject)	2.64	.51	4,331	1.65	3.69	1
SD(Time)	.57	.16	3,151	.27	.89	1
r(Subject, Time)	−.97	.03	2,293	−1.00	−.90	1
SD(Vignette)	1.04	.38	4,018	.49	1.93	1
SD(Time)	.16	.11	2,484	.01	.42	1
r(Vignette, Time)	−.61	.45	5,171	−.99	.72	1

As predicted, I found that participants' explanation choices were predicted by how long they spent deliberating about their choice – participants that reflected on the explanandum tended to choose the inherent explanation less often than participants who made their decisions quickly, β = -0.26, 95% CI [-.50 to -.03]. Here again I observed an inherence bias in explanation: Of the 24 vignettes participants were tasked with explaining, 19 vignettes exhibited a bias in the predicted direction (i.e., towards inherence). Of 125 participants, 86 participants' explanation averages exhibited a bias in the predicted direction, all BFs > 3. This finding provides further support for the claim that people's tendency to explain scientific phenomena inherently is rooted in heuristic processing; longer decision-times were negatively correlated with inherent explanation choices.

Studies 3 and 4 thus provide additional evidence that the inherence bias in scientific explanation that I observed in Studies 1 and 2 is rooted in a domain-general inherence heuristic.

Study 5: The influence of prompting deeper search on inherent explanations

Studies 1 – 4 have provided evidence for an inherence bias in scientific explanation that exhibits the signatures of heuristic thought. In Study 5, I sought to quasi-experimentally test the claim that requiring participants to reflect longer by having them produce additional explanations would reduce the subsequent proportion of inherent explanations.

Methods

Participants. Participants in Study 5 were 98 Amazon Mechanical Turk workers who were paid $0.60 for participating in the study. An additional two participants were tested but excluded because they indicated that they were not paying attention during the study.

Materials and Procedures. Participants were presented with six vignettes (developed on the basis of the vignettes used in Study 1) and asked to perform a fill in the blank task modeled on the Production condition from Study 3. In Study 5, however, participants were asked to fill in the blank in the sentence "Something about _____ explains the observation", which was presented three times on the same page after each of 6 vignettes for a total of 18 explanations. Thus, participants were required to reflect further on the observation because they were required to produce multiple distinct explanations for the same observation. The design of Study 5 was otherwise identical to the Production condition of Study 3.

Results

I predicted that participants would be more likely to provide inherent explanations initially but this tendency would decrease as they provided additional explanations of the same observation. To test this prediction, I fit a Bayesian logistic mixed-effects model predicting the production of an inherent explanation (1 = Inherent, 0 = Extrinsic) based on the order in which the explanation was produced. This model included random intercepts for subject and vignette. The results are summarized below in Figure 12 and Table 13.

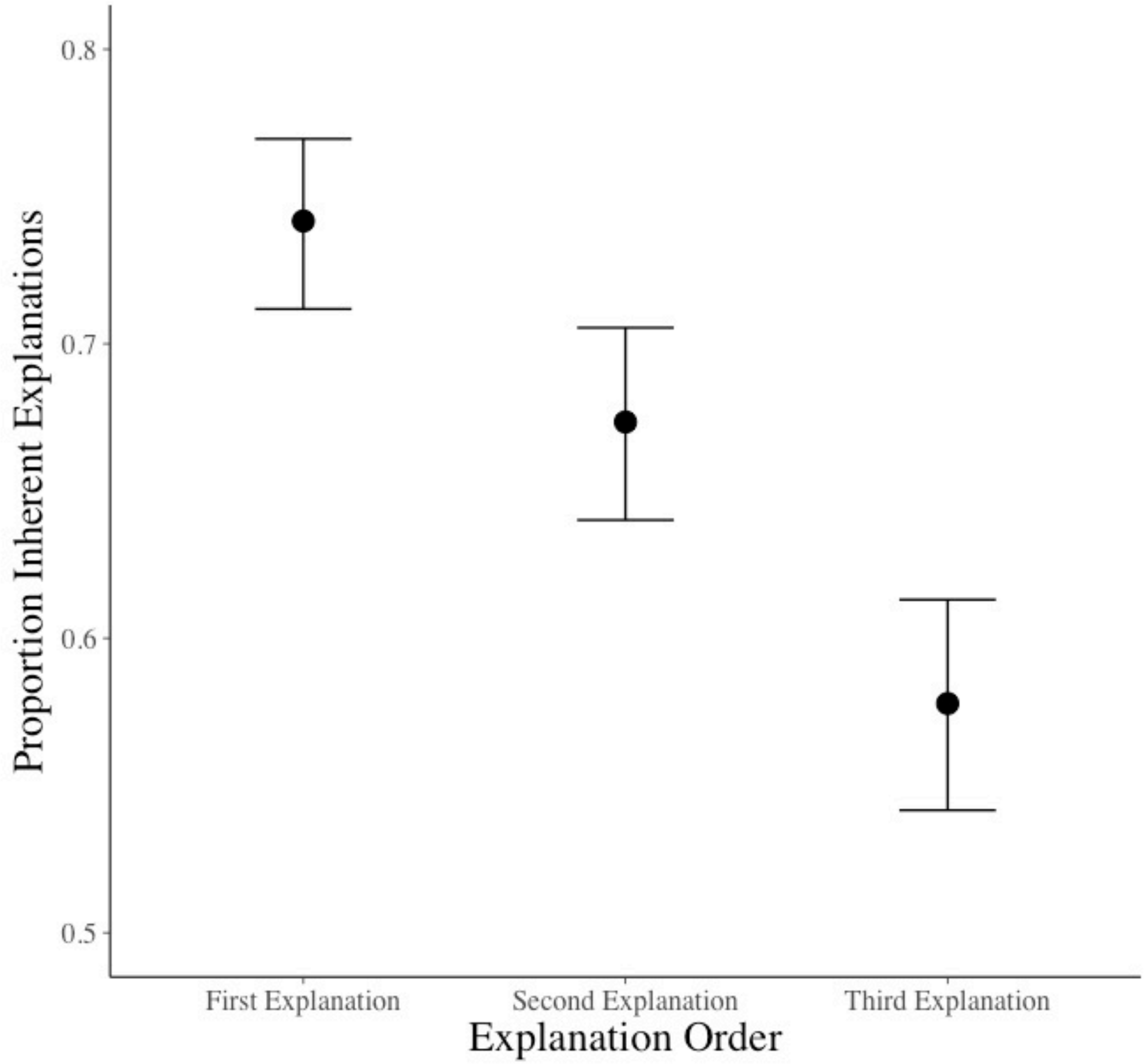

Figure 7. The mean proportion of inherent explanations for novel scientific observations by the order in which the explanations were produced in Study 5. Error bars represent 50% Credible Intervals.

Table 4

Bayesian logistic mixed-effects regression predicting proportion of inherent explanations on the basis of the order in which they were produced in Study 5.

Population-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
Intercept	1.06	.25	4,787	.56	1.55	1
Second Explanation	**−.33**	**.15**	**16,000**	**−.61**	**−.05**	**1**
Third Explanation	**−.74**	**.15**	**16,000**	**−1.03**	**−.45**	**1**
Group-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
SD(Subject)	.89	.11	5,884	.7	1.11	1
SD(Vignette)	.44	.24	4,837	.17	1.05	1

Note: The reference group is the First Explanation group.

Consistent with the results of Study 4, I found that participants' first explanations of an observation tended to be more inherent than extrinsic, with each subsequent explanation of the same observation tending to be less inherent and more extrinsic; as participants were forced to reflect more deeply on the observation to be explained, they provided more extrinsic explanations.

The emergence and the development of the inherence bias in scientific explanation

Intuitive biases affect children's understanding of scientific phenomena inside and outside the classroom. Children make a several errors when they reason about the scientific world. Moreover, we see reduced inherent fact use as children develop (Cimpian & Steinberg, 2014), consistent with the relational shift in thinking across development (e.g., Doumas et al., 2008) and decreased intuitive thinking across childhood. In line with this prior work, I sought to further test the heuristic roots of the inherence bias in scientific explanations by examining how young children between ages 6 and 9 explain scientific observations. These results would provide further evidence for the proposed heuristic source of the bias toward inherence in scientific explanation.

Study 6: Inherent explanations across development

Methods

Participants. Twenty-nine children (ages 6 – 9; 58.3% Female; M_{Age} = 7.52 years old) were recruited in a Midwestern college town to participate on a study in which they explained why a series of scientific events occurred.

Materials and Procedures. In Study 6, children were read four stories about scientific phenomena (two from biology, one from chemistry, and one from physics). I adapted the scientific vignettes from Study 1 to be age-appropriate and further altered these vignettes with the aim of very strongly emphasizing the relevance of the extrinsic feature to the observed pattern. My hope was that this would increase the variance in children's responses. As such, I was not interested in overall inherent (vs. extrinsic) responding given that prior work has already shown that children explain more inherently than extrinsically (Cimpian & Steinberg, 2014). For example, children were read with the following vignette (see Table B of the Supplementary Materials for additional examples):

> There is a plant called a Flurp Plant that lives high on a volcano in Africa. The volcano that this plant lives on is *really* tall, way taller than most volcanos! Now, you know how most plants need lots of air for them to grow? Well, this kind of plant barely needs any air at all even though it's like other plants in every other way. The Flurp Plant can live for a long time without air.
>
> Don't you think that's cool!? What's really amazing is that the Flurp Plant is like other plants in every other way, so it's really strange that it does not need air.

After each vignette, children were asked to explain why the observation occurred.

> If you had to make a guess, why do you think that is? Why doesn't the Flurp plant need air?

After children provided an explanation for the observation, I then asked them to describe to the experimenter as much as they could remember about the story. This was to confirm children were paying attention and could recall central details about each vignette. Children correctly recalled the central details of the vignettes over 95% of the time.

My primary aim was to test the hypothesized influence of a heuristic process on scientific explanation by examining whether the proportion of inherent explanations decreases with age and the proportion of extrinsic explanations increases with age.

Results and Discussion

Children's responses were coded for inherence and extrinsicness just as in Studies 1 and 2. I then performed Bayesian logistic mixed-effects modeling, predicting the production of an inherent (or extrinsic) explanation based on a child's age. This model also included random intercepts for subject and vignette.

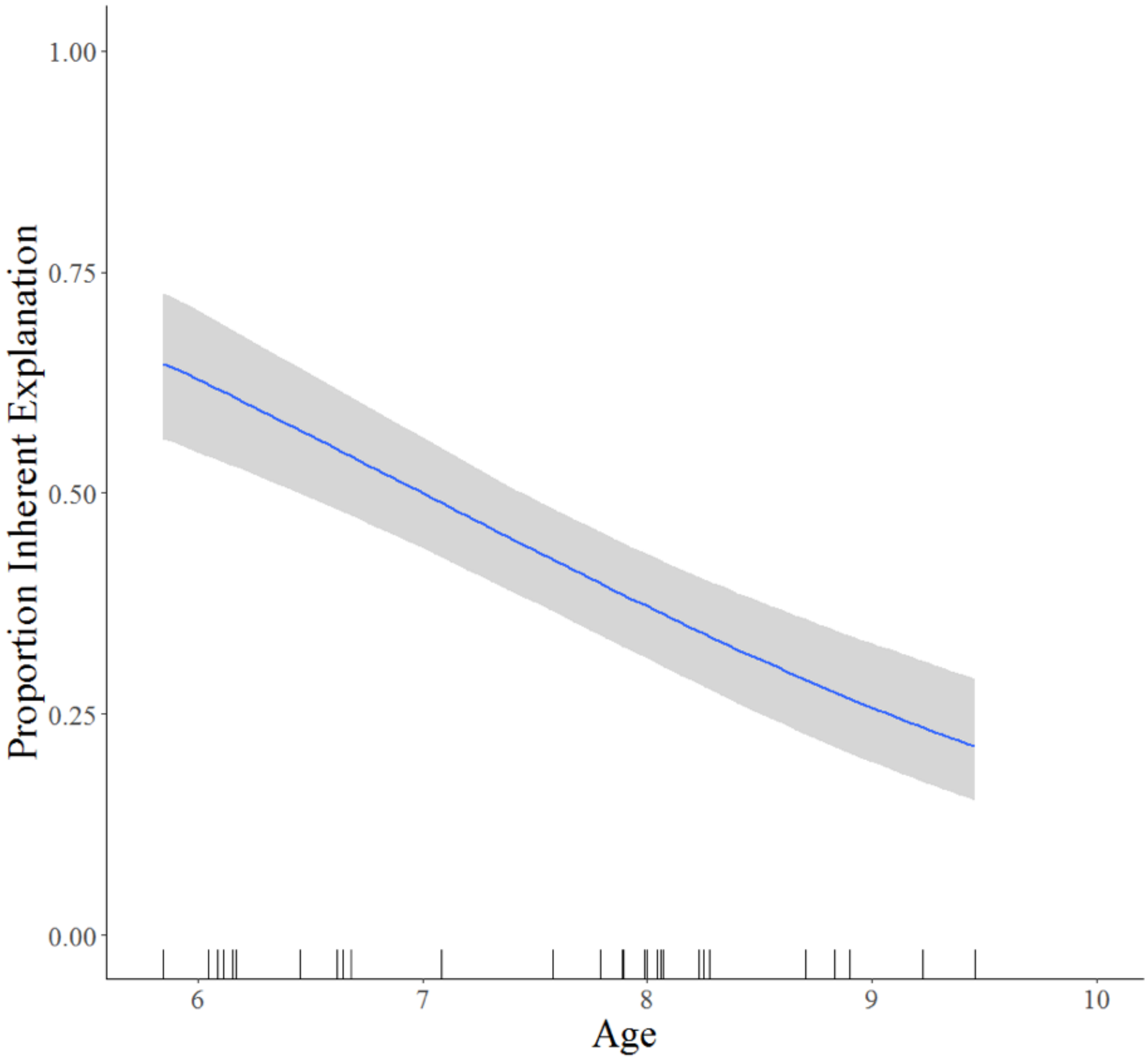

Figure 8. The marginal likelihood of children of different ages producing inherent explanations in Study 6. Shaded region represents 50% Credible Intervals.

Table 5

Bayesian logistic mixed-effects regression predicting proportion of inherent explanations on the basis of age in Study 6.

Population-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
Intercept	3.68	1.88	5,434	.02	7.46	1
Age	**−.53**	**.24**	**16,000**	**−1.02**	**−.07**	**1**
Group-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
SD(Subject)	.66	.40	4,058	.04	1.52	1
SD(Vignette)	.64	.78	2,552	.02	2.82	1

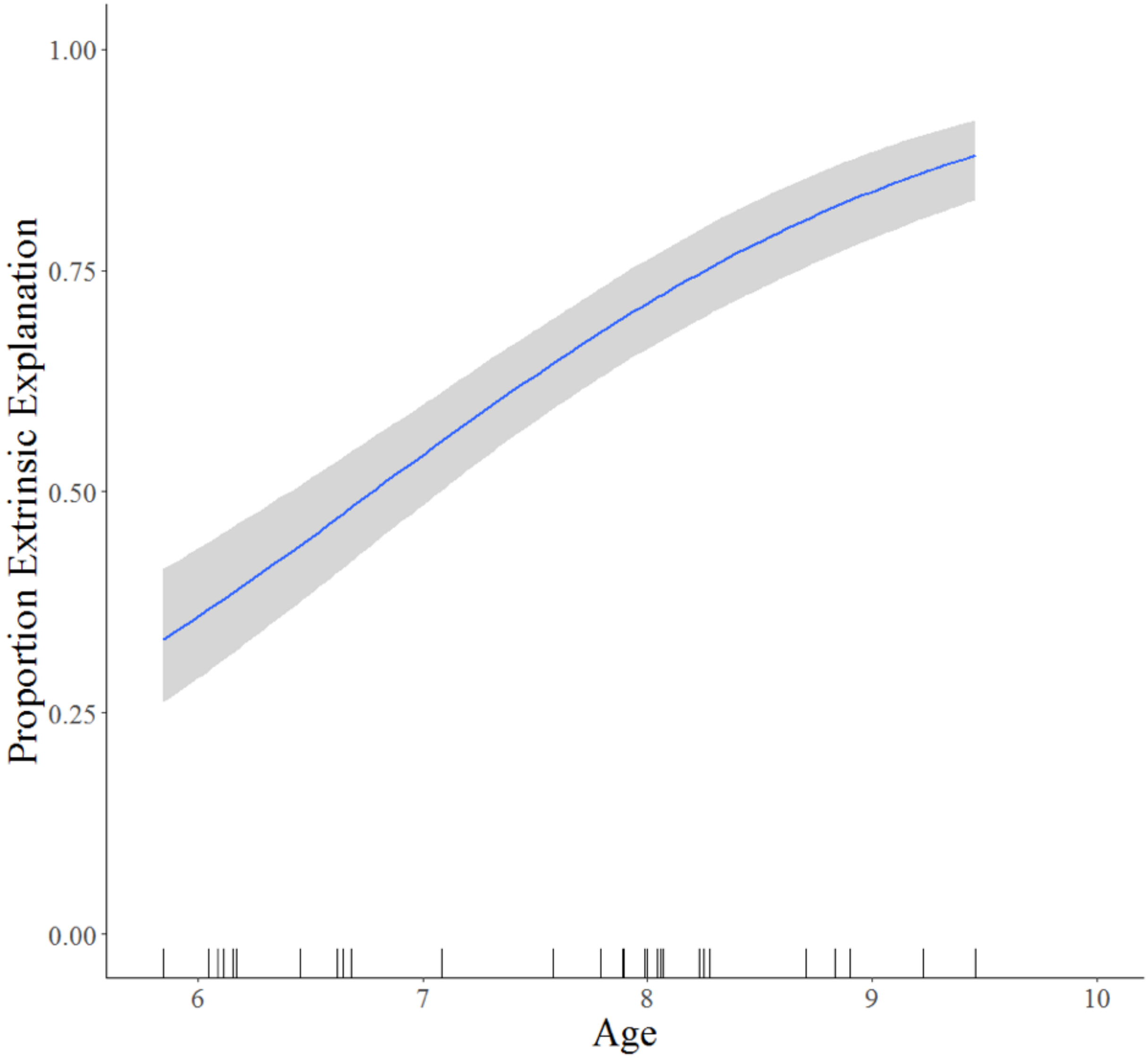

Figure 9. The marginal likelihood of the predicted effect of age on extrinsic explanations in Study 6. Shaded region represents 50% Credible Intervals.

Table 6

Bayesian logistic mixed-effects regression predicting proportion of extrinsic explanations on the basis of age in Study 6.

Population-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
Intercept	-5.07	1.90	16,000	−9.00	−1.48	1
Age	**.75**	**.25**	**16,000**	**.28**	**1.27**	**1**
Group-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
SD(Subject)	.61	.38	4,925	.03	1.47	1
SD(Vignette)	.46	.57	4,123	.01	2.00	1

As predicted, I found that older children (as compared to younger children) were less likely to explain a pattern inherently (β = -0.53, 95% CI [-1.0 to -.08] and more likely to explain a pattern extrinsically, β = 0.75, 95% CI [0.29 to 1.26] (see Figures 8 and 9). These findings are consistent with the hypothesis that children with more limited cognitive resources – that is, younger children – tend to explain scientific phenomena more inherently than extrinsically. In Study 7, I replicated the results of Study 6.

Study 7: Replicating the results of Study 6

In Study 7, I aimed to replicate the results of Study 6 with a larger sample size.

Methods

Participants. Eighty-two children (ages 6 – 9; 51.2% Female; M_{Age} = 7.95 years old) were recruited in a Midwestern college town to participate in a study in which they learned about and explained a series of scientific phenomena.

Materials and Procedures. The materials used and the procedures followed were identical to those of Study 6.[7] I predicted that I would replicate the results of Study 6 – older children would explain less inherently and more extrinsically than younger children. If extrinsic facts are harder to use to produce an explanation because they are more complex, then children with larger working-memory capacities should have less trouble using these facts in their explanations than children with more limited working-memory capacities.

Results

I observed comparable results to those of Study 6: Older children (as compared to younger children) again explained more extrinsically (β = 0.54, 95% CI [0.20 to 0.91] and less inherently, β = -0.37, 95% CI [-0.71 to -0.06] (see Figures 9 and 10; Tables 7 and 8 below).

[7] Study 7 included an additional measure aimed at measuring a child's working-memory capacity – the Backwards Digit Span Task. However, there were problems in the administration of the task which led us fail to detect the well-validated finding that older children have larger spans than younger children (see Wechsler, 2003, p. 267). There were no differences in the spans 6 (*M* = 3.4, *SD* = .75), 7 (*M* = 3.36, *SD* = .49), 8 (*M* = 3.42, *SD* = .81), or 9-year-old children (*M* = 3.79, *SD* = .85). In light of these issues, I have omitted discussion of this measure.

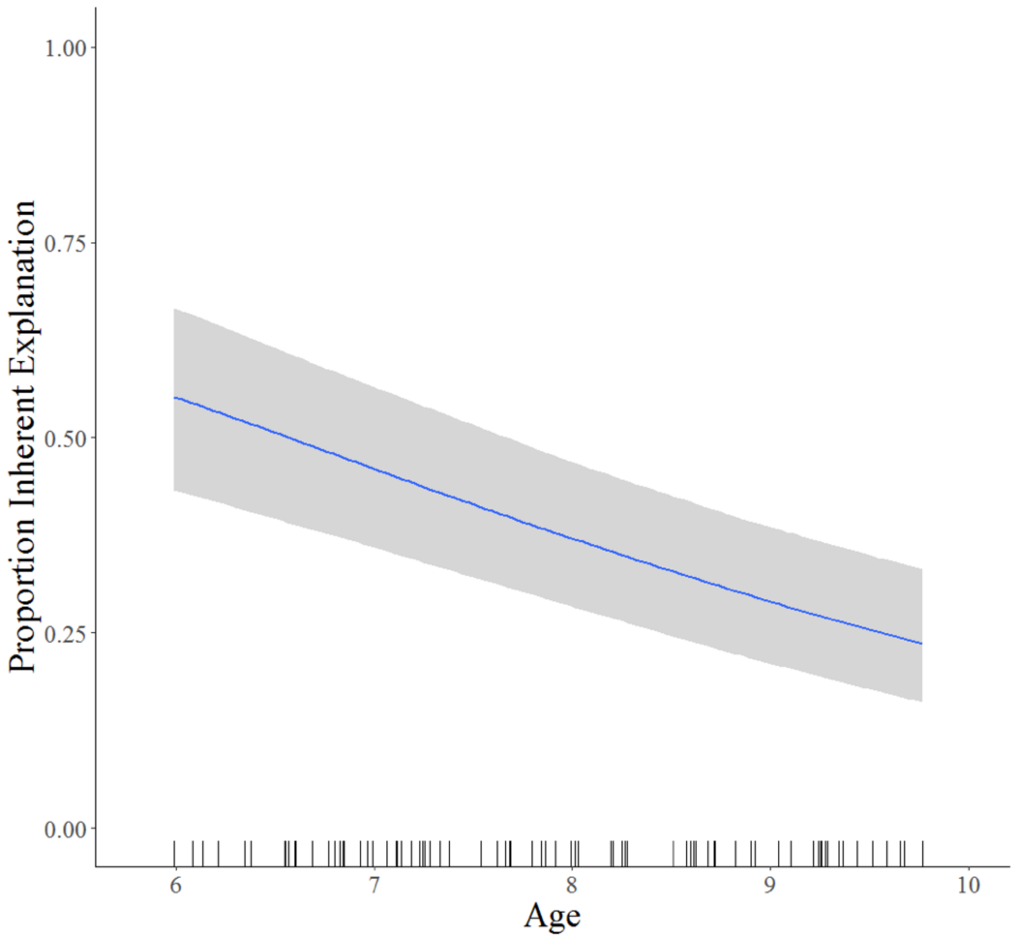

Figure 10. The marginal likelihood of children of different ages producing inherent explanations in Study 7. Shaded region represents 50% Credible Intervals.

Table 7

Bayesian logistic mixed-effects regression predicting proportion of inherent explanations on the basis of age in Study 7.

Population-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
Intercept	2.41	1.58	6,648	**−.69**	5.60	1
Age	**−.37**	**.16**	**11,913**	**−.69**	**−.06**	**1**
Group-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
SD(Subject)	1.00	.26	4,519	.49	1.53	1
SD(Vignette)	1.47	1.2	4,065	.43	4.86	1

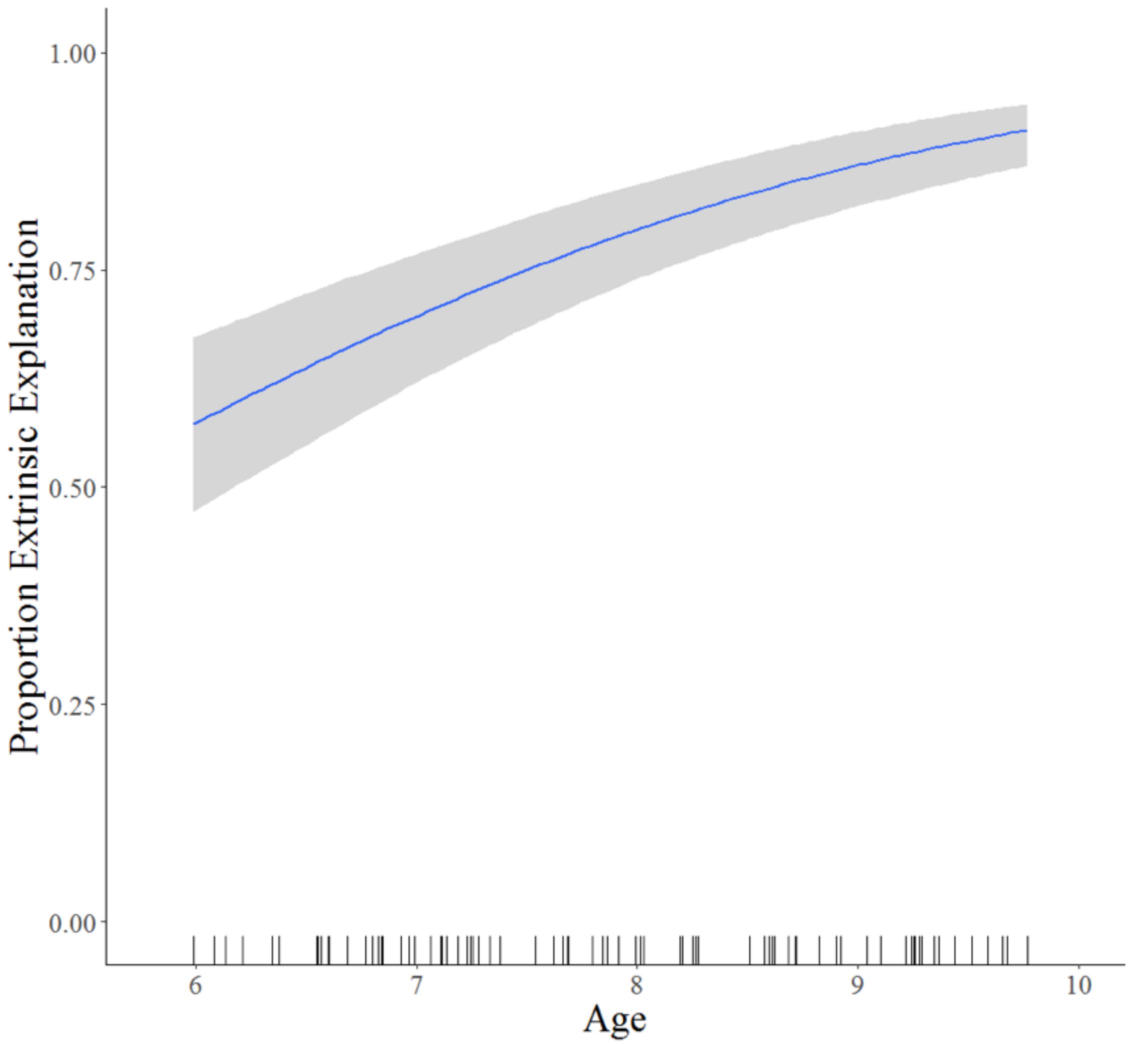

Figure 11. The marginal likelihood of children of different ages producing extrinsic explanations in Study 7. Shaded region represents 50% Credible Intervals.

Table 8

Bayesian logistic mixed-effects regression predicting proportion of extrinsic explanations on the basis of age in Study 7.

Population-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
Intercept	-2.96	1.61	16,000	**−6.19**	**.12**	1
Age	**.54**	**.18**	**16,000**	**.20**	**.91**	**1**
Group-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
SD(Subject)	.99	.30	3,452	.38	1.59	1
SD(Vignette)	1.2	1.14	2,444	.24	4.34	1

Chapter 4: Implications for the History of Science and Study 8

A systematic investigation suggest that heuristic processes influence the scientific explanations of naïve participants. However, one might wonder the extent to which naïve scientific explanation diverges from the explanations of practicing scientists. That is, how much do naïve participants' explanatory biases generalize to practicing scientists?

As discussed earlier, it appears that psychological constraints can even affect how practicing scientists conduct science and explain phenomena. For example, Dunbar (1997) found that in top research labs, scientists often relied on featural similarity in drawing analogies between a source and target rather than abstract, structural similarity (Chen & Khlar, 1999). Even the most characteristically deliberative exercises will be constrained by cognitive limitations of working-memory and, more generally, basic facts about human psychology. Beyond these data, is there evidence that these cognitive constraints can also lead to an inherence bias in scientific explanation? A careful inspection of the history of science suggests that inherent explanations may be particularly common among the first researchers who have attempted to explain a scientific phenomenon—that is, those scientists that most clearly mirror the position of our naïve participants. For example, consider the progression of theories of motion: According to Aristotelian mechanics, an object falls because of the object's inherent downward tendency; the object's natural place is on the ground, and that is a property of the object itself. Impetus theorists (e.g., Buridan) continued to attribute movement to a certain force within the moving object (i.e., impetus) but also began to acknowledge the importance of other forces extrinsic to the object itself. Later, Newton explained motion by primarily invoking an external gravitational force, an extrinsic explanation.

In the science education literature, researchers have also noted the similarity between the biases of naïve participants and working scientists. Driver and colleagues (2005) note:

> Throughout history the same everyday experiences that led earlier scientists like Aristotle and Buridan to develop theories of motion have a lot in common with gut and lay dynamics.

Finally, some evidence for the link between naïve participants' explanations and working scientists' explanations comes from the results of Study 2 – participants' explanations in Study 2 often mirrored early scientists' explanations of these same phenomena. To consider just one example, several participants provided quasi-Aristotelian explanation of why sediment emerged in distilled boiling water. One participant wrote,

> "Something in the water has broken down and the soil is no longer suspended in the water so it fell to the bottom."

In contrast, another participant explained the same phenomenon extrinsically in a manner extremely similar to that of Lavoisier, who famously refuted Aristotle's original hypothesis by carefully measuring the weight of the water and vessel before and after he conducted the study:

> "I believe that the vessel decomposed with in the 100 days and caused a solid sediment to form."

Aristotle hypothesized that water can be transmuted into soil. In contrast, Lavoisier argued, based on a systematic investigation in which he weighed both the beaker and the water prior to performing the study, that the sediment came from the beaker rather than the water. Although this is just one example, several participants' explanations for a variety of historical scientific observations often mirrored explanations that appear throughout the history of science.

Based on empirical findings, as well as a careful inspection of the historical record, I hypothesized that the pioneers of a scientific field—those who cannot benefit from the accumulated insights of past generations—may be particularly likely to use the seeds supplied by their explanatory intuitions to formulate systematic hypotheses about an observation. In turn, if these explanatory intuitions are often inherent, as I have found in naïve participants, then early theoretical forays into a domain may also tend to be skewed toward inherence. This is not to say that early scientific theories are somehow less thoughtful or well-reasoned than later ones. Rather, I suggest that one's basic explanatory intuitions may end up exerting a more noticeable influence on one's scientific thinking if these intuitions are not reined in by exposure to the evidence accumulated by previous generations of practitioners. My claim should not be construed as deterministic: Obviously, science is not a march toward ever-more-extrinsic explanations. However, if my hypothesis is correct, I should nevertheless see a greater preponderance of early inherence-based explanations than one might expect otherwise, along with a reliable tendency for later explanations of those same phenomena to correct for some of the early inherence bias.

Rather than testing this hypothesis by examining only a few examples from the history of science, I recruited historians from history and philosophy of science departments in the US and UK.

Method

Participants. Participants were 25 experts in the history and philosophy of science recruited from major research universities with history and philosophy of science departments in the US and UK.

Materials and Procedure. In Study 8, participants were asked to list with the most important explanatory transitions in the history of science. I explained to participants that explanatory transitions are a situation in which a scientist explained an observation in one way (e.g., Aristotelian dynamics explained why an object falls) and a later scientist explained this same observation in a different way (e.g., Newtonian mechanics). Of the 25 historians of science that participated in the study, 23 provided examples of major scientific transitions (initial and subsequent explanations of the same observation), yielding 81 transitions. (Two of the scientists failed to complete the task because they just listed names without reference to any specific observation (e.g., Aristotle, Einstein)). Two hypothesis blind coders with graduate-level education in the history of science coded historians' responses for inherence and extrinsicness (Overall Explanation Agreement: 82%).[8] An additional 24 transitions were provided by participants but the coders judged that these responses were either not explanatory transitions (e.g., shifts in taxonomy), or they were orthogonal to the inherence-extrinsicness distinction.

Recall that I predicted that earlier explanations of a phenomenon would tend to be more inherent because early scientists' intuitive explanatory beliefs would be less likely to be reined in by the evidence accumulated by previous generations of practitioners. Table 9 below shows the pattern of major explanatory transitions from the history of science.

[8] I recruited coders with a background in the history and philosophy of science because some background in history was necessary to interpret the often terse responses of professional historians and philosophers of science.

Table 9

Initial and Subsequent Major Explanatory Transitions from the History of Science in Study 8.

		Subsequent	
		Inherent	Extrinsic
Initial	Inherent	12	**46**
	Extrinsic	9	14

Table 10

Bayesian categorical regression by explanatory transition type (Intercept only model) in Study 8.

Population-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
Inherent to Inherent (Intercept)	-1.53	1.36	2,008	-5.03	.05	1
Extrinsic to Inherent (Intercept)	-1.28	.84	3,860	-3.25	-.10	1
Inherent to Extrinsic (Intercept)	**1.11**	**.33**	**15,000**	**.50**	**1.79**	**1**
Group-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
SD(Subject (Inherent to Inherent Intercept))	1.83	1.41	1,927	.12	5.31	1
SD(Subject (Extrinsic to Inherent Intercept))	.93	.86	3,245	.03	3.09	1
SD(Subject (Inherent to Extrinsic Intercept))	.44	.35	6,392	.02	1.29	1

Note. Reference group is the explanatory transition “Extrinsic to Inherent.”

Consistent with my proposal, historians of scientists generated many more examples of explanatory transitions in which the early explanation of an observation was inherent and the subsequent explanation of this same phenomenon was more extrinsic, β = 1.11 95% CI [0.50 to 1.79] (Table 9 and 10 above).

This work reveals an important facet of early scientific explanations that had not yet been articulated in prior work on the psychology of the history and philosophy of science (e.g., Holyoak & Thagard, 1996; Nersessian, 1992; Wiser & Carey, 1983): Early explanations of major scientific explanatory transitions tend to be more inherent and less extrinsic than later explanations of the same observation.

Are practicing scientists really exhibiting an inherence bias?

Although I have provided evidence that naïve participants exhibit an inherence bias in their explanations of scientific phenomena, one might wonder to what extent this claim accurately characterizes the behavior and theorizing of practicing scientists. Turning to the case of Brown, one an alternative explanation of how Brown began to make sense of the motion of the pollen is not that he exhibited an inherence bias rooted in heuristic processes but rather that he was working with the only information that he had available. After all, technological limitations did not allow Brown to observe the submicroscopic particles that were pushing the pollen around. Generalizing from this example, perhaps it also the case that many instances of early "inherent" explanations from the history of science are cases in which the practicing scientist had no other information at their disposal—there was no extrinsic fact on which they could base their explanations.

First, the results from the studies with naïve participants speaks against this hypothesis. I have found that inherent explanations decrease over development (Studies 6 and 7), decrease as people deliberate longer (Studies 4 and 5). These findings are consistent with the presence of a heuristic process that is biased towards inherence and inconsistent with the claim that an inherence bias follows from exclusive reliance on the data at hand.

Second, if Brown and other vitalists, for instance, were solely relying on the data that was available to them then we would not expect them to posit unseen vitalist forces to explain the motion of the pollen. While they might be able to conclude that "something about the pollen" explains its motion because no other information was available, positing vitalistic forces goes well beyond the data. Given the similarity in the vitalist explanations of Brown, naïve participants, and even young children (Driver et al., 2005), it seems just as likely that their explanations were constrained by basic cognitive mechanisms.

Third, even if this explanation can account for the case of Brownian motion, many other historical examples suggest an inherence *bias* because plausible extrinsic factors were available. For instance, Aristotle and many of his followers thought that water could be transmuted into soil based on studies showing that when water was boiled in a sealed vessel, solid sediment would begin to emerge at the bottom of the vessel. In this case, the decomposition of the vessel rather than properties of the water explained the presence of the solid sediment. Yet, Aristotle and many other scientists prior to Lavoisier thought that something about the nature of water—the focal entity in the pattern to be explained—explained the emergence of the sediment rather than the vessel in which the water was contained. The vessel was entirely accessible as an alternative explanation, but it simply was ignored in theorizing. There are many other similar cases in the history of science, including early explanations for organisms being well-adapted to their environments (the presence of the environment was observable), retrograde motion of the planets (the presence of the Sun was observable), spontaneous generation (the presence of flies was observable) and so forth. Thus, looking at the historical record provides evidence for an inherence bias which cannot be easily explained by the hypothesis that practicing scientists were relying solely on the information at their disposal.

In sum, the present empirical work, practicing scientists' willingness to posit unobservables, and the historical record speak against the hypothesis that there was no other information available for early scientists to rely on as they constructed their explanations. Instead, these results suggest that naïve participants and scientists alike, at least under certain circumstances, rely on their intuitions as they begin to make sense of the world.

Chapter 5: General Discussion

People have a drive to explain and understand the world (e.g., Gopnik, 1998; Keil, 2006; Khemlani & Johnson-Laird, 2011; Lombrozo, 2006; Murphy & Medin, 1985; Premack & Premack, 1996; Ross, 1977; Weiner, 1985). Yet, explanatory reasoning is an enormously complex task. This may be particularly true in the domain of science where minor changes in environmental factors could greatly impact the stability of a chemical reaction or the behavior of animals. However, rather than being dumbfounded at every turn, people produce explanations and do so relatively effortlessly—indeed, our participants knew little about the focal entities in the pattern and their interactions with the environment but they nonetheless managed to come up with explanations for observations involving these entities. Prior research has suggested that one source of people's explanations are domain-specific intuitive beliefs. The present work suggests that in addition to many of these domain-specific biases there is a domain-general inherence bias rooted in heuristic processing that systematically influences people's explanations across chemistry, biology, and physics.

In Studies 1 and 2, I found that naïve participants explained scientific phenomena in chemistry, physics, and biology in terms of the inherent facts of the entities in the observation, even when a plausible extrinsic explanation was accessible.

Next, I provided evidence that the inherence bias in scientific explanation is rooted in heuristic processes. Together, the increased salience of inherent facts and the increased relational complexity of extrinsic explanations lead people to oversample inherent facts when they produce explanations. In Study 3, I found that the salience of inherent facts leads people to over-rely on them when they construct explanations. In Studies 3 and 4, I found that even when participants were aware of the relationship between an environmental factor and a pattern under consideration—that is, when participants did not have to retrieve this variable—people still exhibited an inherence bias. However, I found that the inherence bias was reduced when people reflected on the explanandum longer (Studies 4 and 5) and among those with additional cognitive resources (older children compared to younger children; Studies 6 and 7). Together, these findings provide evidence for the hypothesis that domain-general cognitive constraints give rise to an inherence bias in scientific explanation that is rooted in heuristic processes.

Finally, I provided evidence for an inherence bias even among practicing scientists. I found that early explanations of major observations in the history of science—that is, the explanations produced by scientists who most closely resemble the position my naïve participants found themselves in—tended to be more inherent than subsequent explanations of the same observations. This work suggests that the inherence bias in scientific explanation I observed among naïve participants can even influence the explanations and theories of practicing scientists and sheds light on an important but as of yet unnoticed facet of early scientific explanations.

To conclude, Brown peered through his microscope and observed pollen rapidly zigzagging about. He explained its motion as owing to "the particle itself." This event embodies both naïve participants' and early scientists' explanatory strategy. Naïve participants, and even early scientists, appear to explain novel scientific observations more inherently than extrinsically and this tendency appears to be rooted in a heuristic cognitive process.

References

Brewer, W. F., Chinn, C. A., & Samarapungavan, A. (1998). Explanation in scientists and children. *Mind and Machines*, *8*(1), 119-136.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, *70*(5), 1098-1120.

Cimpian, A., & Salomon, E. (2014). The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences*, *37*(5), 461-480.

Cimpian, A., & Steinberg, O. D. (2014). The inherence heuristic across development: Systematic differences between children's and adults' explanations for everyday facts. *Cognitive Psychology*, *75*, 130-154.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*, 781.

Doumas, L. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review*, *115*(1), 1-43.

Driver, R., & Erickson, G. (1983). Theories-in-action: Some theoretical and empirical issues in the study of students' conceptual frameworks in science, *Studies in Science Education*, *10*(1), 37-60.

Driver, R., Rushworth, P., Squires, A., & Wood-Robinson, V. (Eds.) (2005). *Making Sense of Secondary Science: Research into children's ideas*. Routledge. New York, New York.

Dunbar, K. (1997). How scientists think: On-line creativity and conceptual change in science. In T. B. Ward, S. M. Smith & J. Viad (eds.), *Creative Thought: An Investigation of Conceptual Structures and Processes*. American Psychological Association, 461-493.

Evans, J. S. B. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review*, *13*(3), 378-395.

Gelman, A. (2016, November 5). *Why I prefer 50% rather than 95% intervals*. Retrieved from http://andrewgelman.com/2016/11/05/why-i-prefer-50-to-95-intervals/

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2014). *Bayesian data analysis* (Vol. 2). Boca Raton, FL, USA: Chapman & Hall/CRC.

Gelman, S. A. (2003). *The Essential Child: Origins of essentialism in everyday thought*. New York: Oxford University Press.

Gilbert, D. T., Krull, D. S., & Pelham, B. W. (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology, 54*(5), 733–740.

Gopnik, A. (1998). Explanation as orgasm. *Minds and Machines, 8*, 101–118.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, *21*(06), 803-831.

Halloun, I. A., & Hestenes, D. (1985). The initial knowledge state of college physics students. *American Journal of Physics*, *53*(11), 1043-1055.

Hatano, G., & Inagaki, K. (1994). Young children's naive theory of biology. *Cognition*, *50*(1-3), 171-188.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In: *Social psychology: Handbook of basic principles*, ed. E. T. Higgins & A. W. Kruglanski, pp. 133–168. Guilford Press.

Higgins, E. T., King, G. A., & Mavin, G. H. (1982) Individual construct accessibility and subjective impressions and recall. *Journal of Personality and Social Psychology, 43*, 35–47.

Higgins, E. T., Rholes, W. S. & Jones, C. R. (1977) Category accessibility and impression formation. *Journal of Experimental Social Psychology, 13,* 141–154.

Hoffman, M. D., & Gelman, A. (2014). The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, *15*(1), 1593-1623.

Holyoak, K. J., & Thagard, P. (1996). *Mental Leaps: Analogy in Creative Thought*. MIT press.

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110*(2), 220–264.

Hussak, L. J., & Cimpian, A. (2015). An early-emerging explanatory heuristic promotes support for the status quo. *Journal of Personality and Social Psychology*, *109*(5), 739-752.

Inagaki, K., & Hatano, G. (2004). Vitalistic causality in young children's naive biology. *Trends in Cognitive Science*, *8*(8), 356-362.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kahneman, D. & Frederick, S. (2002). Representativeness Revisited: Attribute Substitution in Intuitive Judgment, in T. Gilovich, D. Griffin, D. Kahneman (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment*, 49–81, Cambridge: Cambridge University Press

Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology*, *57*, 227-254.

Khemlani, S. S., & Johnson-Laird, P. N. (2011). The need to explain. *The Quarterly Journal of Experimental Psychology*, *64*(11), 2276-2288.

Khemlani, S., & Johnson-Laird, P. N. (2013). Cognitive changes from explanations. *Journal of Cognitive Psychology*, *25*(2), 139-146.

Leslie, S. J. (2013). Essence and natural kinds: When science meets preschooler intuition. *Oxford Studies in Epistemology*, *4*, 108-166.

Lombrozo, T. (2006). The structure and function of explanations. *Trends in cognitive sciences*, *10*(10), 464-470.

Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology*, 55(3), 232-257.

McRae, K., & Jones, M. N. (2013). Semantic memory. In: *The Oxford Handbook of Cognitive Psychology*, ed. D. Reisberg. Oxford University Press.

Maybery, M. T., Bain, J. D., & Halford, G. S. (1986). Information-processing demands of transitive inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(4), 600-613.

McCloskey, M. (1983). Intuitive physics. *Scientific American*, *248*(4), 122-130.

McCloskey, M., & Kohl, D. (1983). Naive physics: the curvilogistic impetus principle and its role in interactions with moving objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(1), 146-156.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92*, 289–316.

Neal, R. M. (2011). MCMC using Hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo*, *2*, 113-162.

Nersessian, N. J. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. *Minnesota Studies in the Philosophy of Science*, *15*, 3-45.

Premack, D., & Premack, A. J. (1995). Origins of human social competence. In M. S. Gazzzaniga (Ed.), *Neural and Psychological Development* (pp. 205–218). Cambridge, Massachusetts: The MIT Press.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605.

Ross, L. D. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology, 10*, 173–220.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t*-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225-237.

Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, *26*(5), 521-562.

Russell, T., & Watt, D. (1990). *Primary space project research reports: Sound*. Liverpool University Press.

Simon, H. A. (1982). *Models of bounded rationality: Empirically grounded economic reason* (Vol. 3). Cambridge, MA: MIT Press.

Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and Brain Sciences*, *23*(05), 701-717.

Sutherland, S. L., & Cimpian, A. (2015). An explanatory heuristic gives rise to the belief that words are well suited for their referents. *Cognition*, *143*, 228-240.

Tworek, C. M., & Cimpian, A. (2016). Why Do People Tend to Infer Ought from Is? The Role of Biases in Explanation. *Psychological Science*, *27*(8), 1109-1122.

Vosniadou, S., & Brewer, W. F. (1987). Theories of knowledge restructuring in development. *Review of Educational Research*, *57*(1), 51-67.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, *24*(4), 535-585.

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p-values. *Psychonomic Bulletin & Review*, *14*(5), 779-804.

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, M., ... & Miller, B. L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science*, *10*(2), 119-125.

Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory & Cognition*, *28*(7), 1205-1212.

Waskan, J., Harmon, I., Horne, Z., Spino, J., & Clevenger, J. (2014). Explanatory anti-psychologism overturned by lay and scientific case classifications. *Synthese*, *191*(5), 1013-1035.

Weiner, B. (1985) "Spontaneous" causal thinking. *Psychological Bulletin, 97*(1), 74–84.

Wilkenfeld, D. A., Plunkett, D., & Lombrozo, T. (2016). Depth and deference: When and why I attribute understanding. *Philosophical Studies*, *173*(2), 373-393.

Wilson, R. A., & Keil, F. (1998). The shadows and shallows of explanation. *Minds and Machines*, *8*(1), 137-159.

Wiser, M., & Carey, S. (1983). When heat and temperature were one. *Mental Models*, D. Gentner & A. Stevens (Eds.), 267-297. New York, NY: Psychology Press.

Yu, E., Sprenger, A., Thomas, R., & Dougherty, M. (2014). When decision heuristics and science collide. *Psychonomic Bulletin & Review*, *21*(2), 268-282.

Appendix: Supplementary Materials and Analysis

Validating the anomaly assumption

Studies 1 – 5 provide evidence for an inherence bias in scientific explanation that is rooted in heuristic processing. However, in all of these studies I have assumed that anomalies should lead to more extrinsic and less inherent responding. To validate this assumption, I directly tested whether claiming that an observation is an anomalies does indeed reduce inherent explanations.

Methods

Participants. Participants were 99 Amazon Mechanical Turk workers who were paid $0.60 for participating in the study. An additional participant was recruited but indicated they were not paying attention and so were excluded from analyses.

Materials and Procedures. I presented participants with fictitious scientific observations (all vignettes taken from Study 5) that only differed in whether the vignette stated that the experienced scientists were surprised by the observation in question (Anomaly Condition) or merely noted it (No Anomaly Condition) (Condition: within subjects). Participants read each vignette and then chose which of four explanations (two inherent and two extrinsic) best explained the observation. Participants were provided with additional explanation options in order to address the concern that the particular response options I provided in Studies 3 and 4 increased the inherence bias. An example vignette is shown below:

> **Anomaly Condition:** A group of experienced biologists wanted to observe the mating behavior of the Siberian Nyctereutes dogs' first offspring. Typically, dogs have multiple mates over the course of their lives. Surprisingly, these biologists observed that the first offspring of the Nyctereutes dog maintained one mate for its entire life.

> **No Anomaly Condition:** A group of experienced biologists wanted to observe the mating behavior of the Siberian Nyctereutes dogs' first offspring. Dogs often have at least one mate over the course of their lives. These biologists observed that the first offspring of the Nyctereutes dog maintained one mate for most of its life.

Results

To confirm the assumption in Studies 1 – 5, I fit a Bayesian logistic mixed-effects model predicting inherent explanation choices (1 = Inherent, 0 = Extrinsic) based on condition assignment (1 = No Anomaly condition, 0 = Anomaly condition) with random intercepts for both subject and vignette. Consistent with my prior assumptions, the Anomaly condition (86% inherent) led to less inherent responding than did the No Anomaly condition (93% inherent), suggesting that participants recognize—at least upon reflection—that anomalous behavior is inconsistent with an inherent explanation of that behavior.

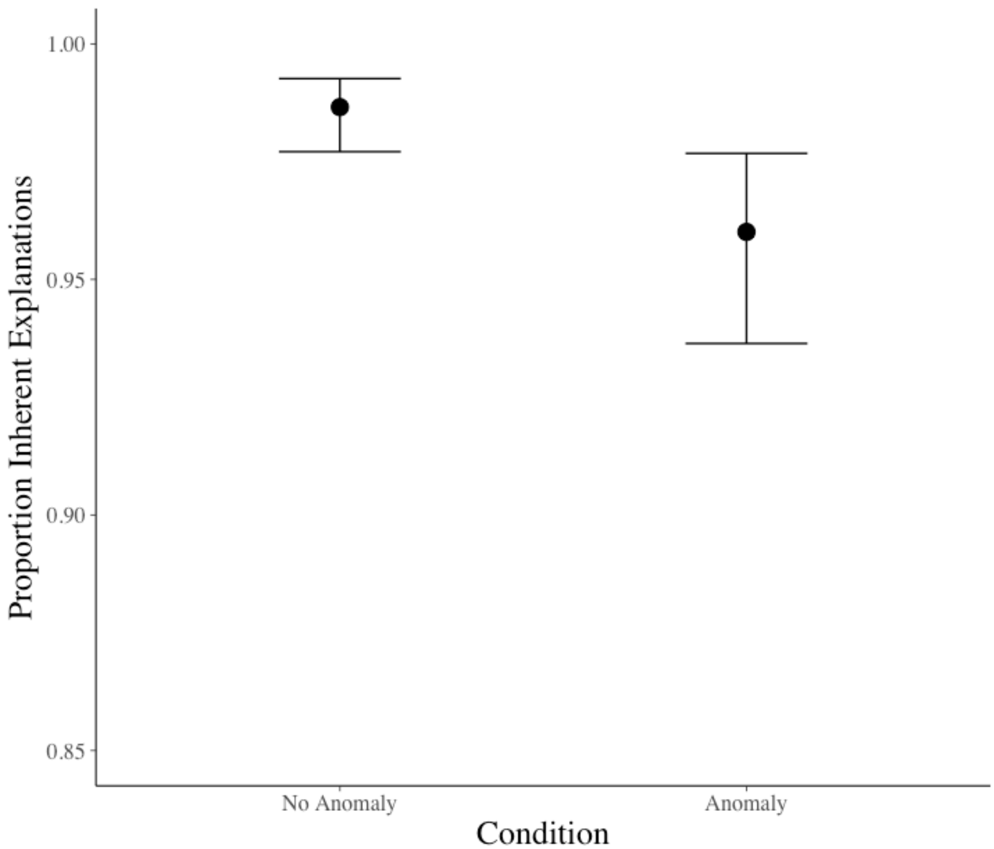

Figure 12. The mean proportion of inherent explanations for novel scientific observations by condition. Error bars represent 50% Credible Intervals.

Table 11

Bayesian logistic mixed-effects regression predicting proportion of inherent explanations on the basis of condition.

Population-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
Intercept	3.21	.93	4,349	1.45	5.11	1
Condition	**1.12**	**.36**	**14,000**	**.44**	**1.85**	**1**
Group-Level Effects	Estimate	*Error*	Effective Steps	95% *CI*		$\hat{r}$
SD(Subject)	2.10	.42	4,239	1.38	3.03	1
SD(Vignette)	1.69	.94	4,720	.62	4.14	1

Table 12

Examples of additional fictitious vignettes with sample explanations and coded responses

Fictitious Observation	Domain	Vignette Text	Sample Inherent Explanation	Sample Extrinsic Explanation
	Chemistry	Chemists in a lab high in the Colorado Rockies were investigating the possibility of storing hydrogen atoms in lithium nitride crystals. In a strange turn of events, these chemists found that when they attempted to store the hydrogen atoms in the lithium nitrides crystalline structures, it led to a result that completely defied their expectations. Instead of the lithium nitride structure storing 15% of hydrogen atoms, the crystaline structures stored 0% of hydrogen atoms. This is despite the fact that these scientists had more than 40 years of experience between them working with lithium nitride.	"The lithium nitrides crystallines struture are not homogenous enough to store the hydrogen atoms." "The atoms in the lithium nitride might have repelled the Hydrogen atoms. Perhaps all of the bonds that could be made have already been made, and so any outside protons and electrons would be repelled no matter. Another possibility is that the lithium nitride is storing another atom without the scientists realizing it."	"The altitude may have caused the hydrogen atoms not to be able to be stored in the lithium structures." "The environment did not facilitate storage."
	Physics	A group of experienced physicists were looking to create thermal solar panels for use in urban environments. These physicists were conducting field measurements of experimental solar panels in a densely populated square of Beijing, China. Typically, their panels are capable of converting 75% of the energy they absorb into usable power. Surprisingly, when they measured the conversion of energy into usable power, they noticed that the experimental solar panels were converting nearly 90% of the power they received into usable energy.	"The panels were better constructed then they believed and this accounted for the better results." "The solar panels must simply be more efficient at solar energy conversion."	"Perhaps the buildings in the dense area were built of a material that helped the solar panels capture more energy." "In china there are a lot of skyscrapers that will refract the better, making more energy for the solar panels"
	Biology	A group of experienced biologists wanted to observe the rare *Nyctereutes* dogs' mating behavior. These biologists were working in the forests of Siberia, the *Nyctereutes* dogs' habitat. Typically, dogs have multiple mates over the course of their lives. Surprisingly, these biologists observed that the *Nyctereutes* dogs maintained one life-long mate for most of their lives.	"the dogs probably feel that they'd rather be with someone they are loyal to." "Maybe they have a different gene in them that allows them to be more loyal to their mating partner. Something different than what other dogs have, like crabs."	"Maybe this group of dogs were a isolated from the standard pack of dogs and it changed." "Because they live in a remote area, this particular dog has greater bonds with their mates because there are not a lot of other choices and therefore stays with the same mate."

Table 13

Examples of additional children's vignettes with sample explanations and coded responses

Domain	Vignette Text	Sample Inherent Explanation	Sample Extrinsic Explanation
Chemistry	Daxy water is really cold and found deep underground in a salt-cave [show pictures]. The place that daxy water is found in is a *really* salty cave, way saltier than most caves! Now, you know how water turns to ice when it's cold outside? Well, daxy water is really cold but it still doesn't turn to ice. Daxy water doesn't turn to ice, even though it is really, really cold. Don't you think that's cool!? What's really amazing is that daxy water is like other water in every other way, so it's really strange that it doesn't turn to ice when it's really cold.	"Maybe it's because it's a special kind of water that only lives in caves and it has a special chemical to it that makes it not freeze." "Because the daxy water is already cold that's why it never gets turned into ice."	"Because it's in a cave, instead of outside where it should be, the other water turns to ice because it's not protected." "because of the salt in the cave."
Physics	An astronaut standing on the Moon wanted to do a really cool experiment [show picture]. They dropped a blarky hammer and a wuggy feather at the same time to see which would hit the ground first [show picture]. Well, the astronaut found that the blarky hammers and the wuggy feather hit the ground at the same time, even though these were like regular hammers and feathers in every other way. The blarky hammer and the wuggy feather hit the ground at the same time! Don't you think that's cool!? What's really amazing is that the blarky hammer and the wuggy feather are like other hammers and feathers in every other way, so it's really strange that they hit the ground at the same time.	"Maybe because they way the same pounds and um they um have like the same um like weight or something. They weigh maybe 50 pounds the same." "Because they're the same weight"	"Because there is no gravity on the moon." "So it's gravity, on the moon in space, so they dropped at the same time, even though on earth there's gravity, but there isn't gravity so it feels the same weight."
Biology	There is a wild dog called a tyco dog that lives in the coldest part of Alaska [show pictures]. The forest that this dog lives in is *really* cold, way colder than most places! Now, you know how most dogs have pink tongues? Well, this kind of dog has a striped tongue, even though it's like other dogs in every other way. The tyco dog has a striped tongue! Don't you think that's cool!? What's really amazing is that the tyco dog is like other dogs in every other way, so it's really strange that it has a striped tongue.	"Because it is the wildest creature in world." "I think it's because it needs to have a difference from other dogs. So I don't think it's a normal dog."	"I think that the Tyco dog has a striped tongue because it is half really cold because the forest is really cold sometimes isn't cold too because it has shelter." "It has a striped tongue because um Alaska on the rocks the snow kind of makes kind of striped lines."

Table 14

Examples of historical vignettes with sample explanations and coded responses

Historical Observation	Domain	Vignette Text	Sample Inherent Explanation	Sample Extrinsic Explanation
	Chemistry	A chemist conducted the following experiment: They applied heat to a piece of manganese and then weighed it, noting that the manganese had gained mass. This result was quite striking to the chemist.	"The heat caused the manganese to expand due to the material the manganese is made of." "The manganese expanded because heat was applied to it. Because the molecules were excited and started spreading apart. "	"A chemical reaction occurred that changed air particles into physical additions to the original object." "Heating the manganese created a chemical reaction between the metal and the oxygen combining the two. Therefore, the product had more mass at the end."
	Physics	A physicist performed the following experiment: They fired a cannonball at a target 1000 meters to the north. The physicist noticed that the cannonball curved right of the target repeatedly even though there was no wind that day and the cannon was perfectly straight. This result was quite striking to the physicist.	"The cannon ball might have not be perfectly round." "The cannonball is not smooth. There is an indentation on one side that is making it pull to one side in flight."	"Rotation of the Earth" "I would guess that the rotation of the earth is affecting the trajectory of the cannonball."
	Biology	A biologist conducted the following study: The biologist placed pollen from the *Clarkia pulchella* flower in water under a microscope and observed it. The biologist noticed that the bits of pollen were vigorously moving around despite the fact that the pollen had been place in storage for months. This result was quite striking to the biologist.	"The bits of pollen had not died during the storage period." "The pollen moved around because it was from a living organism, the Clarkia pulchella flower. Despite being put in storage, the pollen was not impacted by that."	"The water was moving." "The microbes in the water reacted to the pollen and resulted in the motion. Or something in the pollen was reacting to the water."