A GENERAL EXPLANATORY TENDENCY UNDERLIES COMMON ASSUMPTIONS ABOUT HOW THE WORLD WORKS: AN INVESTIGATION OF NOMINAL FIT AND ESSENTIALIST BELIEFS

BY

SHELBIE LLOY SUTHERLAND

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, 2016

Urbana, Illinois

Doctoral Committee:

Associate Professor Andrei Cimpian, Chair
Professor Renee Baillargeon
Professor Cynthia Fisher
Professor John Hummel
Professor Aaron Benjamin
Abstract

People are constantly making sense of their environment, explaining a wide variety of events and facts that they encounter. Often, these explanations are the product of a heuristic process that relies on highly accessible information to make sense of a phenomenon (Cimpian & Salomon, 2014a, 2014b; Thomas, Dougherty, Sprenger, & Harbison, 2008). Importantly, highly accessible information is often skewed to include features inherent to that which is being reasoned about, without reference to external impact or happenstance. Because of this bias in highly accessible information, generating explanations heuristically has downstream consequences for the beliefs people hold about the way the world works. Here, I focus on two beliefs that I propose are underlain by inherent reasoning. In Part 1, six studies provide evidence that children’s and adults’ inherent reasoning promotes a belief that words and their referents are not arbitrarily paired but rather “fit” particularly well together (nominal fit). In Part 2, I provide evidence across three studies that inherent reasoning may also play a role in the development of essentialist beliefs. Overall, this work suggests that general inherent reasoning buttresses disparate beliefs about the way the world works. These findings highlight the value in investigating the ways in which abstract, higher-order judgments and beliefs are shaped by basic cognitive processes.
Acknowledgements

The support that has brought me to this point has been so impactful, humbling and continuously astounding. This note does not do everyone justice, and cannot adequately express the extent of my gratefulness, but here is what I can say:

Andrei has gone way beyond advising my doctoral research. I can confidently say (and do confidently say) that he has helped me to reach my potential as a researcher at this point, laying a solid foundation to develop my writing and critical thinking. Other faculty have also been incredible mentors and have provided valuable support. Renee has been an active mentor for every milestone throughout the program and has consistently encouraged me, most notably during my setting up next steps. Such encouragement has also been explicit in working with Susan—from my first project, to talking about ideas, and discussing the future. Ori sparked my interest in research and remains an inspiring mentor. Cindy, Aaron, and John each helped to develop my scientific inquiry through their feedback on my projects as well as in their classes.

Beyond faculty support, I am so grateful to have had fellow graduate students who are both great people as well as scientists (Lin, Christina, Larisa, Michelle, Zach, Daniel, Erika, Xin, Ethan, Mike(s), Miatta, Noam, and Liz). They have all made such a difference to me in my time in grad school. My work was made possible in large part because of many research assistants, lab managers, schools, parents, children, and departmental staff; it was fantastic working with them. And the financial support from the NSERC and APF helped me to focus on my projects.

So many others have provided priceless emotional and personal support. My parents (Jeff and Kim), brother (MacKenzie), and grandparents (Skip and Nancy; as well as Muriel, Lloy, and Cyrus) have always stood close, encouraging, cheering, and guiding me. They mean the
world to me. And I was always comforted in knowing that the rest of my family also had my back in any circumstance.

Anyone who knows me well right now will know the names Christina, Larisa, and Michelle. These unbelievable ladies, soul-mates, have held my hand and walked with me day in and day out throughout the past four years. They are incredibly generous, supportive, fun, sassy, thoughtful, scientific, true-friends. I could not have gotten through grad school without them.

I was also lucky enough to have more invaluable support and friendship from fantastic humans that I look up to and will work hard to remain close with. First, Marie has not stopped being there for me, near and far, since we first became friends in high school. I am so excited that my next step is going to bring us nearer than ever. Second, Caroline’s and Kye’s genuineness and consideration makes for a deeply rooted friendship. It allows for a thought provoking, loving, and incredibly comfortable space that I am so grateful to have had during this time. There are two other people who have made the most palpable (positive) effect on my perception and reasoning that I know will maintain its impact through the next many chapters of my life: Ellen takes caring and thoughtfulness to a new level. I love thinking with her, and the perspective that she continues to show me. Luke has helped me get to a distinctly healthier way of thinking and seeing things. The difference is so obvious and unbelievable. I don’t have the words to describe how grateful I am for that. Lastly, in this non-exhaustive list, I want to thank the many other friends in psych, at NCF, back home, and (now) scattered around the world, for letting me lean on you in moments throughout my time at Illinois.

I have learned so much academically and personally in these past five years and wouldn’t be who I am and where I am now without all of these people.
## Table of Contents

Chapter 1: Introduction ......................................................................................................................... 1  
   The complexity underlying the process of explanation ................................................................. 2  
   Explanations tend to rely on highly accessible information ....................................................... 3  
   Accessible information is often inherent to the explanandum .................................................. 3  
   Evidence for the inherence heuristic ............................................................................................ 5  
   Consequences of the inherence heuristic ...................................................................................... 6  
Chapter 2: Part 1. Inherent Reasoning Gives Rise to the Belief that Words Are Well Suited for  
   Their Referents .............................................................................................................................. 8  
   Prior Evidence on Beliefs about Nominal Fit .............................................................................. 9  
   Nominal Fit Intuitions Are a Product of the Heuristic Processes Underlying Explanation ........ 11  
   The Current Studies ..................................................................................................................... 13  
Chapter 3: Study 1 .............................................................................................................................. 15  
   Methods ........................................................................................................................................ 15  
   Participants ................................................................................................................................. 15  
   Procedure .................................................................................................................................... 15  
   Measures ....................................................................................................................................... 15  
   Data analyses ............................................................................................................................... 19  
   Results and Discussion .................................................................................................................. 20  
Chapter 4: Study 2 .............................................................................................................................. 22  
   Methods ........................................................................................................................................ 23  
   Participants ................................................................................................................................. 23  
   Measures ....................................................................................................................................... 23  
   Results and Discussion .................................................................................................................. 25  
Chapter 5: Study 3 .............................................................................................................................. 26  
   Methods ........................................................................................................................................ 26  
   Participants ................................................................................................................................. 26  
   Measures ....................................................................................................................................... 26  
   Results and Discussion .................................................................................................................. 27  
Chapter 6: Study 4 .............................................................................................................................. 29  
   Methods ........................................................................................................................................ 29  
   Participants ................................................................................................................................. 29  
   Measures ....................................................................................................................................... 29  
   Results and Discussion .................................................................................................................. 30  
Chapter 7: Study 5 .............................................................................................................................. 31  
   Methods ........................................................................................................................................ 32  
   Participants ................................................................................................................................. 32  
   Measures ....................................................................................................................................... 32  
   Results and Discussion .................................................................................................................. 34
Table 3 .................................................................................................................................................96
Table 4 ..................................................................................................................................................97
Table 5 ..................................................................................................................................................98
Table 6 ..................................................................................................................................................99
Table 7 ................................................................................................................................................100
Table 8 ................................................................................................................................................101
Table 9 ................................................................................................................................................102
Appendix A .........................................................................................................................................103
Appendix B .........................................................................................................................................104
Appendix C .........................................................................................................................................105
Appendix D .........................................................................................................................................106
Chapter 1: Introduction

Beginning in infancy (e.g., Aguiar & Baillargeon, 2002; Saxe, Tenenbaum, & Carey, 2005), and across development, people constantly need to make sense of their environment, explaining events that they encounter. Not only are people explaining on a routine basis, but they also explain a wide variety of phenomena (e.g., why was there a bang in the hall? why is there always a line at that store? why does it feel so relaxing to be by water?). When making sense of an experience or phenomenon, there are often countless factors, at least in principle, to consider in order to put together an accurate and complete judgment. At the same time, people have limited time and cognitive resources to devote to generating explanations. Because of these seemingly opposing forces, people might often engage in a simpler process, generating explanations that are “good enough” using relevant information that is most easily accessible in memory (Cimpian & Salomon, 2014a). Here, I outline prior work on a heuristic explanatory process, and propose that a skew in the content of the information that tends to be used in this process gives rise to many common beliefs about the way the world works. Specifically, in Part 1 I will focus on people’s intuitive beliefs about how language works. In Part 2, I investigate children’s intuitive beliefs about categories—in particular, the assumption that many categories in the world have deep underlying “essences.” Before moving on to these particular proposals, I will first discuss the explanatory process that may underlie the development of these beliefs.

Accumulating evidence suggests that people typically generate explanations heuristically (Cimpian & Salomon, 2014a, 2014b; Salomon & Cimpian, 2014; see also Thomas, Dougherty, & Buttaccio, 2014). That is, the process of generating everyday explanations (1) does not always require extensive cognitive resources, (2) often relies on highly accessible information, and (3)
does not use an exceptionally high standard for evaluating the quality of the explanation generated.\(^1\) Explaining in this way is likely quite frequent: People are able to make sense of new, even unexpected, situations constantly throughout the day, without focused deliberation at every turn. Furthermore, even preschoolers generate explanations frequently and with relative ease despite the, at least in-principle, complexity in the cognitive processes involved in generating explanations (e.g., Hickling & Wellman, 2001). Thus, it seems plausible to claim that our cognitive systems take shortcuts when generating explanations—that is, that people generate explanations heuristically. The idea of explanation as a heuristic process is key for the current purposes of investigating the beliefs that may follow, as the properties of this heuristic process might give rise to many common assumptions about how the world works. Next, I discuss the argument for explanations being heuristic.

**The complexity underlying the process of explanation**

In theory, generating an explanation for a particular event or phenomenon is an incredibly complex task. Even a simple explanandum such as eggs being a common breakfast food has countless factors that could be considered when coming up with an explanation. An adequate explanation for this phenomenon would probably require a great deal of in depth information. At the same time, people make sense of their surroundings in the moment. Even very young children quickly and easily generate explanations for their observations (e.g., Hickling & Wellman, 2001). The contrast between the number of factors that could be considered in generating an explanation, and the ease with which people explain, can be reconciled if the act

---

\(^1\) The term “heuristic” is used in two fairly distinct ways in the literature on reasoning: One refers to a more conscious implementation of a rule of thumb, to deliberately simplify a task, and another to an implicit process that uses intuitive reasoning, little cognitive resources, and is highly dependent on easily available information (e.g., Chaiken & Ledgerwood, 2012; Evans, 2008). Here, I will be using the latter sense of the term.
of explaining was at least at times underlain by a heuristic process.²

Explanations tend to rely on highly accessible information

If explanations are often a product of a heuristic process they will also often over-rely on highly accessible information. This tendency is found with other heuristic judgments (e.g., Chaiken & Ledgerwood, 2012; Kahneman & Tversky, 1973; Thomas et al., 2014; Tversky & Kahneman, 1974). For instance, when asked to judge the frequency of words that begin with the letter “k” compared to words that have the letter in the third position, people will often incorrectly choose the former as more frequent due to the higher accessibility of words that start with “k” in memory (e.g., Tversky & Kahneman, 1973). Such an over-reliance on highly accessible information is considerably less prevalent with more analytic, effortful thinking (Evans & Stanovich, 2013). In the context of generating explanations, what impact might such an over-reliance on highly accessible information have on the content of the explanations people form?

Accessible information is often inherent to the explanandum

Both direct and indirect evidence suggests that highly accessible information is often about inherent features of the entities whose behaviors or properties are being explained (Hussak & Cimpian, under review; see also Ashcraft, 1978; McRae, de Sa, & Seidenberg, 1997; Rosch & Mervis, 1975; Wong & Weiner, 1981).³ That is, it is often easier to retrieve facts about

---

² To be clear, I am not claiming that heuristic reasoning is the only way through which people explain. First, there are likely specific theories that people have about particular phenomena that they can use to direct their explanatory processes, without relying on finding the relevant factors among a wide array of possibilities (e.g., objects fall when unsupported; e.g., Hespos & Baillargeon, 2008; Needham & Baillargeon, 1993). And second, people sometimes recall previously learned explanations rather than generating an explanation anew. For instance, when explaining to a child why the light went out, people can use their knowledge about how lights works to arrive at two possibilities: either the bulb burnt out or the power is off. However, I am arguing that in many cases people use heuristic reasoning processes in generating their explanations (Cimpian & Salomon, 2014a; Thomas, Dougherty, Sprenger, & Harbison, 2008; Salomon & Cimpian, 2014).

³ The use of the term “inherent” refers to a psychological notion—the way most people would judge whether a feature is inherent—not to whether features are truly inherent in a metaphysical sense.
the explanandum itself, which do not reference external entities, than it is other facts. For instance, an inherent explanation for why we eat eggs more in the morning (than other times of the day) might be that eggs have a lot of protein (an inherent feature). In the same way, when explaining why girls paint their nails and why cats have been domesticated, it is likely that features inherent to girls and cats will come to mind quickly and easily. Such inherent facts tend to come to mind faster and more often than non-inherent facts (Hussak & Cimpian, 2015): people are more likely to generate inherent explanations when they are under time pressure than when they are able to take more time (Hussak & Cimpian, under review).

In line with these findings, Cimpian and Salomon (2014a, 2014b) proposed that people typically generate explanations through a heuristic explanatory process that typically outputs inherence-based explanations—a process that they termed the *inherence heuristic*. According to the inherence heuristic proposal, the process of explaining begins when the need for an explanation is detected. Accessible (thus often inherent) information might then come to mind right away, and people will tend to use this information to put together an explanation. This explanation is then checked against a fairly low standard of plausibility (see Cimpian and Salomon, 2014a, 2014b, and Cimpian, 2015, for a comprehensive description of the process). The process is not engaged for every explanation people generate. It is instead one process that is frequently used, and that often results in a skew towards inherent explanations.

---

4 In work testing for an inherence bias, it is important for the researcher to keep clear the explanandum of the particular explanation—just knowing the explanation alone is not enough to determine whether it is inherent. The importance of taking into account the explanandum can be seen in the following example. An inherent explanation for why there are so many female nurses would be “because women really enjoy doing that.” However, the same explanation (“because women really enjoy doing that”) would not be considered inherent if it was given as an answer to why there are so many nail salons—women’s preference is a factor that’s extrinsic to the explanandum in this case (i.e., nail salons and their abundance).
Evidence for the inheritance heuristic

If inherent explanations are often the product of an intuitive, low-effort heuristic process, then such explanations should be more common when fewer cognitive resources are available to explain (Gilbert & Hixon, 1991; Gilbert, Pelham, & Krull, 1988; Trope & Gaunt, 2000). Consistent with this prediction, when adults were experimentally induced to use few cognitive resources, they showed a stronger preference for inherence-based explanations (Hussak & Cimpian, 2015; Salomon & Cimpian, in preparation). Similarly, adults with a general cognitive tendency to avoid deep contemplative reasoning are also more likely to reason inherently (Hussak & Cimpian, 2015; Salomon & Cimpian, 2014; Sutherland & Cimpian, 2015; Tworek & Cimpian, in press). If an inherence bias in explanation is due to a heuristic process, the bias should be more prevalent in children. Children have fewer cognitive resources in general than adults, and thus they should be more likely to rely on an available heuristic process (Kokis et al., 2002; Toplak, West, & Stanovich, 2014). Initial tests of children’s explanations suggest this prediction likely holds: when children were provided with both inherent and extrinsic explanations, they showed a consistent preference for the inherent ones (Cimpian & Steinberg, 2014). Ongoing work testing the explanations that children generate themselves is also consistent with this prediction (Sutherland & Cimpian, in progress). Moreover, children showed a stronger preference for inherent explanations than adults did. Evidence of a heightened inherence bias in children, who also tend to have fewer cognitive resources, supports the heuristic nature of the explanatory process.

Before moving on to discuss how the process of heuristic explanation might influence people’s beliefs about the way the world works, there are three important points to keep in
mind about the inherence bias. First, the inherence bias in explanation is a product of the information that comes to mind quickly and easily. Therefore, while on average inherent information about the entity at hand will be over-represented in the most accessible information, in cases where non-inherent information is most accessible to the reasoner, the bias would not be expected to be present. For instance, someone’s cherished object (e.g., a lucky pen) will likely have accessible its history and specific external associations, especially compared to reasoning about an object that is merely a representative of its category (e.g., just any pen; Gelman, Manczak, & Noles, 2012; Gelman, Manczak, Was, & Noles, 2016; Nancekivell & Friedman, 2013). Additionally, an expert on a particular topic will likely have a lot of other information readily available, much of which might be non-inherent (Tworek & Cimpian, in prep).

Second, inherent explanations are not hypothesized to be a result of the reasoner searching for inherent facts intentionally, but simply a byproduct of the fact that inherent information comes to mind quickly and easily, along with the ease of putting together an explanation with this accessible information. Third, because the expected inherence bias is a product of the information that tends to be accessible, explainers do not need to be able to identify inherent facts on their own.

**Consequences of the inherence heuristic**

While a heuristic explanatory process is largely beneficial for people in making sense of their surroundings without having to devoting a large amount of cognitive resources to the task, as is the case with most heuristics, there are also unintended consequences for such reasoning. If people tend to generate explanations heuristically, and these explanations tend to rely on inherent facts, this skew will be mirrored in people’s beliefs about their surroundings. What
might these beliefs look like? Inherent features tend to be stable features of an entity (e.g., Vitamin C is an inherent feature of oranges, and it is also a stable feature of oranges). Inherent features are also often believed to be non-arbitrary (or one that couldn’t have easily been otherwise). Thus, beliefs about different aspects of the world being stable and non-arbitrary might have their roots in the tendency to reason inherently. In this dissertation proposal, I focus on two beliefs from different domains that rely strongly the ideas of stability and non-arbitrariness: the belief that words and the objects they refer to are not contingent (Part 1), and the belief that natural and social category members all hold category-specific internal essences (Part 2). I provide initial evidence that both may be partly rooted in inherent reasoning.
Chapter 2: Part 1. Inherent Reasoning Gives Rise to the Belief that

Words Are Well Suited for Their Referents

The words of a language are arbitrary social conventions. In English, for example, we call trees *trees* and cows *cows*, but we could have just as well called trees *cows* and vice-versa. There is no inherent reason why particular words refer to particular objects, in English or any other language. However, that may not be exactly how people see this matter. In the present research, I build on prior evidence by Piaget (1967) to suggest that children—and even adults—systematically endorse a belief that words “fit” the objects they denote rather than being arbitrarily paired with them. Further, I test a potential source of these wide-ranging intuitions about word–object fit (or, as I will refer to it, *nominal fit*), suggesting that they are a product of the heuristic processes through which people explain the world. More generally, this work seeks to provide new insights into how people conceive of language and to reveal the deep links between these conceptions and the cognitive processes involved in explanation and understanding. In the first section below, I briefly review previous evidence regarding beliefs about nominal fit. I then detail my proposal concerning how these beliefs arise and describe six studies that investigated this proposal.

5 Part 1 has been published in Sutherland, S. L. & Cimpian, A. (2015). An explanatory heuristic gives rise to the beliefs that words are well suited for their referents. *Cognition*, 143, 228-240. doi:10.1016/j.cognition.2015.07.002 Copyright permission to include this work in the dissertation has been granted.

6 Phonetic symbolism, the phenomenon whereby certain sounds have non-arbitrary links to certain meanings (e.g., Maurer, Pathman, & Mondloch, 2006), is more of an exception than a general rule for how the vocabulary of a language maps onto its referents.

7 This term is a play on Piaget’s (1967) nominal realism. The latter is broadly used to refer to a number of beliefs about a non-arbitrary link between words and objects (including, for example, a belief that names are features of the objects or that names have causal effects on the world [see Rozin, Millman, & Nemeroff, 1986; Rozin, Markwith, & Ross, 1990]). For this reason, I introduced the term nominal fit to pick out specifically the phenomenon studied here—namely, a belief that words are particularly suitable for their referents.
Prior Evidence on Beliefs about Nominal Fit

According to Piaget (1967), beliefs about nominal fit represent an intermediate stage in children’s reasoning about words (see also Brook, 1970; Ianco-Worrall, 1972; Vygotsky, 1962). Early in development, children assume a word to do more than just pick out an object; they also believe that the word matches its referent at some level. This belief that words have a special match with the features of their referents was suggested by children’s reasoning about why objects have the names they do (Piaget, 1967; see also Brook, 1970; Ianco-Worrall, 1972; Vygotsky, 1962). For instance, the children interviewed by Piaget thought that the sun is called *sun* not because of arbitrary conventions established in the past but “because it shines” or “because it is all red”; or, to take another example, the mountains are called *mountains* “because they are all white” (Piaget, 1967, p. 84). The underlying intuition here is that an object has the name it does because this name appropriately captures how this object is constituted—names function almost like descriptions (see Russell, 1905). Because they assume nominal fit, children also fail to realize that the names of objects could have been different than they currently are. For example, children claimed that the sun couldn’t have been called *moon* and vice-versa because “the sun makes it warm and the moon gives light” or because “the sun shines brighter than the moon” (Piaget, 1967, p. 81).

Some of this evidence, however, was subsequently criticized on methodological grounds. For example, Markman (1976) pointed out that Piaget’s questions often presumed a sophisticated meta-level understanding of words as objective units of language. Reasoning about words as words may be difficult for young children (e.g., Osherson & Markman, 1975; 8 Piaget labeled the idea that names fit their referents *logical realism*. Since the term *logical realism* is relatively opaque, I prefer to use *nominal fit*.)
Papandropoulou & Sinclair, 1974), and thus children may have answered Piaget’s questions as if they were about the referents instead (see also Bialystok, 1987). When children claimed, for instance, that the sun couldn’t have been called moon, they might have been doing so because they were reasoning about the referent objects rather than about the names of these objects (as if they had been asked whether the sun is the same thing as the moon). Another prominent criticism concerns the fact that some of Piaget’s questions required counterfactual reasoning (e.g., “could we have called the sun moon?”), which might have been too taxing for children’s limited cognitive resources (e.g., Rosenblum & Pinker, 1983). Without appropriate scaffolding, young children might have failed to understand the hypothetical scenario, believing instead that they were being asked about the current names of the relevant objects (e.g., “do we call the sun moon?”).

While reasonable, these criticisms are not sufficient to undermine the claim that children see words as being suited for their referents. Note, for instance, that children give answers that suggest they believe in nominal fit well into the elementary school years (e.g., Ball & Simpson, 1977; Brook, 1970; Piaget, 1967), by which point they have fairly mature metalinguistic awareness of words as elements of language (e.g., Doherty, 2000; Karmiloff-Smith, Grant, Sims, Jones, & Cuckle, 1996) and well-developed counterfactual reasoning skills (e.g., Beck, Robinson, Carroll, & Apperly, 2006; German & Nichols, 2003; Harris, German, & Mills, 1996).

Although existing evidence for intuitions about nominal fit cannot be fully explained by low-level alternatives, research on this topic has been largely dormant in the past few decades, preventing progress on understanding these intriguing beliefs. Most notably, we know very little about their source: Why would children believe so firmly, and for so long, in a fit between names
and objects, especially given that this fit is illusory? Here, I test the proposal that these beliefs are a product of the broader processes by which people make sense of the world across development. Aside from yielding new insights into a decades-old phenomenon that has not yet received a satisfactory account, this research contributes more broadly to our understanding of the ways people conceive of language. By demonstrating that people’s intuitive understanding of the relation between words and objects is continuous with their understanding of non-linguistic phenomena, these studies forge new links between the research on explanation and reasoning and the work on metalinguistic conceptions, perhaps spurring new interest in this often-overlooked topic.

**Nominal Fit Intuitions Are a Product of the Heuristic Processes Underlying Explanation**

The most common and effective route to understanding a phenomenon is to explain it (e.g., de Regt, 2013; Hempel, 1965; Keil, 2006; Lombrozo, 2012), and therefore people’s understanding of the relationship between words and their referents may likewise be rooted in the explanations they generate for these mappings. In other words, attending to the mechanisms that underlie everyday explanations might provide insight into people’s nominal fit intuitions. Indeed, there is already considerable evidence that people tend to generate explanations using heuristic processes that tend to over-rely on information inherent to that being explained. Furthermore, the presence of such an inherence heuristic in explanation is common across development (see previous section; e.g., Cimpian & Markman, 2009, 2011; Cimpian & Steinberg, 2014; Hussak & Cimpian, 2015; Salomon & Cimpian, 2014). But how might this inherence heuristic promote intuitions about nominal fit?

To start, it is worth noting that a phenomenon explained inherently is thereby
understood as arising simply as a product of the entities that make it up, which is also likely to make this phenomenon appear natural and sensible rather than one that could have easily been otherwise (Cimpian, 2015; Cimpian & Salomon, 2014a, 2014b; Cimpian & Steinberg, 2014; Hussak & Cimpian, 2015; Tworek & Cimpian, 2015). For example, if one explains why stop signs are red using the fact that red is a bright color, then—in light of this explanation—the pairing of stop signs and red might also seem natural, even appropriate. This explanation suggests that red is a *suitable* color for stop signs, perhaps even the color that stop signs *should* be. In other words, explaining an observation in inherent terms often licenses further, value-laden judgments about it (for a similar argument concerning the link between explanation and normativity, see Prasada & Dillingham, 2006). This is the rationale for the current proposal of a link between the inherence heuristic and beliefs about nominal fit. I hypothesize that, when children attempt to understand why certain objects have the names they do, they often rely on the inherence heuristic and, as a result, they arrive at explanations couched in terms of the highly accessible inherent features of the referents (e.g., their appearance) or the names (e.g., their sound). In turn, these inherent explanations further suggest to children, and perhaps even to adults, that words are appropriately paired with their referents. If children make sense of why, say, trees are called *trees* by invoking some feature of these objects (e.g., because they have branches)—as if the name describes them—then this name might also seem fitting and sensible to children rather than an arbitrary convention that could have easily been otherwise.

Initial support for this proposal can be found in children’s unambiguously inherent explanations to Piaget’s (1967) questions about why objects have the names they do (see also Brook, 1970; Rosenblum & Pinker, 1983). For example, children said the sun is called *sun*
because of its brightness, its heat, or its color, and they also thought that it couldn’t have any other name. Also consistent with my argument, some evidence suggests that children’s failure to understand words as arbitrary conventions is strongly correlated with their failure to understand other regularities as conventions (e.g., rules of etiquette such as eating with utensils rather than with one’s hands; Lockhart, Abrahams, & Osherson, 1977). These tight relationships are exactly what would be expected if children’s explanations for a broad range of linguistic and non-linguistic regularities relied on the same fundamental cognitive process (the inherence heuristic; Cimpian & Steinberg, 2014).

The Current Studies

Building on this preliminary evidence, the six studies reported here in Part 1 provide the first direct test of the proposal that nominal fit beliefs are rooted in an inherence heuristic in explanation. These studies investigated three predictions of the proposal. First, the extent to which children rely on the inherence heuristic should predict the extent to which they endorse beliefs in nominal fit (Study 1). Second, because the inherence heuristic is influential beyond childhood (Cimpian & Salomon, 2014a, 2014b; Hussak & Cimpian, 2015; Salomon & Cimpian, 2014), I expected to find evidence for the same relationship in adults’ reasoning as well, even though their intuitions about nominal fit should obviously be weaker than children’s and thus show less variability (Studies 2–5). Moreover, the relationship between endorsement of nominal fit beliefs and reliance on the inherence heuristic should hold even when adjusting for a number of control variables.

---

9 The present argument does not conflict with the developmental literature on children’s awareness of the conventionality of words (e.g., Buresh & Woodward, 2007; Clark, 1988; Diesendruck, 2005; Diesendruck & Markson, 2011; Graham, Stock, & Henderson, 2006; see also Cimpian & Scott, 2012). In this literature, conventionality is operationalized as an understanding that a word is widely known—a different dimension of conventionality than the one I am investigating here. Conventionality is a multifaceted construct, and some facets of it (i.e., words are widely known) will likely be much easier to grasp than others (i.e., words are arbitrary).
of other individual differences that could provide alternative explanations for it (e.g., fluid intelligence, motivation for effortful thought, familiarity with foreign languages). Finally, an experimental manipulation that temporarily alters the extent to which adults rely on inherence-based intuitions should have downstream effects on their nominal fit beliefs (Study 6). Across these six studies, I found consistent support for the proposal that an inherence heuristic in explanation gives rise to the intuition that words are well suited for their referents.
Chapter 3: Study 1

In the first study, I investigated whether children’s nominal fit beliefs are predicted by their reliance on the inherence heuristic, as proposed. I focused on a broad age span in this study (children between the ages of 4 and 7) to ensure that children’s beliefs about nominal fit would show sufficient variability.

Methods

Participants. Sixty-four children (half girls and half boys) were recruited from a small city in the Midwestern US (age range = 4.18 to 7.96 years, \( M = 5.89, SD = 1.15 \)). Although demographic information was not formally collected, the children were socioeconomically diverse, and most were European American.

Procedure. Children received two sets of questions: one set measured their nominal fit beliefs and the other measured their use of the inherence heuristic. The order of the two sets of questions was counterbalanced across participants. Each set was accompanied by pictures to maintain children’s attention.

Measures.

Nominal Fit. Nominal fit is the idea that the words people use are, at some level, right or fitting, and thus that no other word–object pairings would have worked as well. To measure endorsement of this idea in a child-friendly way, I presented children with brief vignettes that asked whether familiar objects (e.g., zebras) could have had another name or whether they had to have the name they do. Believing that the name couldn’t have been otherwise might indicate that children believe there is a special fit between objects and their assigned names. Below is the script for one of the vignettes:
Ok, so a long time ago, people didn’t have a name for this kind of animal. [The experimenter points to a picture of a zebra.] They didn’t have a name for it, and they wanted to come up with one. How did they do that? When people were first coming up with a name for this animal, did they have to call it a zebra, or could they have called it something else, like a diby or a pearsa?

The task included three other vignettes identical to that above, except about different words: giraffe, lion, and bear. Each trial used a different pair of novel words as alternatives to the conventional word. The order of the two response alternatives (i.e., that the object had to have that name vs. that the name could have been different) was constant across trials for any one child but was counterbalanced across subjects. The order of the four vignettes was also counterbalanced.

This measure was constructed so as to avoid some of the concerns that have been raised about previous measures in the literature. First, I sought to minimize the possibility that nominal fit responses would stem from a misunderstanding of the question as being about the object, not the word (e.g., Osherson & Markman, 1975). The task blocked this sort of misunderstanding by making it very explicit that they were being asked about the link between the name and the object (“They didn’t have a name for it, and they wanted to come up with one. How did they do that?”). Second, the task scaffolded children’s understanding of the critical question about the hypothetical alternative names by first setting the stage for this question (“Ok, so a long time ago, people didn’t have a name for this kind of animal…”), thereby minimizing the difficulties associated with counterfactual reasoning (Rosenblum & Pinker, 1983).

In addition, because the two response alternatives were long, children could respond
non-verbally by pointing to one of two body parts. For example, children could touch their chin to select one alternative and their ear to select the other. The mapping of the two gestures to the two response alternatives was counterbalanced across children. To minimize the memory load, the experimenter always reminded children which gesture goes with which answer just before asking them to respond. This method allowed children to answer without having to verbally articulate the option with which they agreed. (For a previous use of this procedure, see Cimpian & Park, 2014).

Children’s responses were coded as 1 if they selected the nominal fit option (e.g., they had to call it a zebra) and as 0 if they selected the other option. Children’s scores across the four trials were averaged to derive a composite score that could range from 0 to 1, with higher scores indicating stronger intuitions about nominal fit.

**Inherence Heuristic.** The measure of children’s reliance on the inherence heuristic was borrowed from a prior series of studies investigating the developmental trajectory of this explanatory process (Cimpian & Steinberg, 2014). None of the items in this measure concerned words; rather, they pertained to non-linguistic aspects of the world (e.g., the fact that coins are round). For each item, children were asked three types of questions. The first type measured children’s explanatory preferences. Children were presented with inherent explanations (e.g., coins are round “just because they are coins”) and extrinsic explanations (e.g., coins are round “just because people thought it might be a nice idea”) and were asked to evaluate them on a 4 point scale (“really not right” = 1, “a little not right” = 2, “a little right” = 3, “really right” = 4). These explanations were intentionally nonspecific in the inherent feature or societal fact used. Such generality enables the respective explanations to cover a wide range of possible
explanations that children might have generated on their own. Therefore, children could agree with an explanation despite the fact that it didn’t mention the particular feature or fact that they might have spontaneously thought best explains the phenomena.

The second and third types of questions tapped intuitions that follow from inherence-based explanations. If a particular pattern is explained in inherent terms (e.g., some feature of coins explains why they’re round), one might also reasonably assume that this pattern cannot be changed and is temporally stable. Thus, the second type of question assessed whether children thought it would be okay to change the relevant pattern (e.g., “Imagine if people wanted coins to be a different shape, and everyone agreed that they wanted coins to be a different shape. Would it be okay to make a change so that coins are not round, or would it not be okay?”). Children made their answers on a scale from 1 (“okay”) to 4 (“really not okay”). Finally, the third type of question assessed whether children thought the relevant pattern had always been and would always be as it currently is (e.g., “Do you think coins have always been round, even way back when the first ever coin was made?”). Children’s “yes” answers were scored as 1, and their “no” answers were scored as 0.

Each of the questions was asked about three items (coins are round; birthday cakes have candles; and school buses are yellow). The order of these items was counterbalanced, as was the order of the three types of questions. Children’s answers were standardized within each question and then combined into an inherence heuristic composite score ($\alpha = .63$), with higher scores indicating greater reliance on the output of this explanatory process.

**Control variables.** To address potential alternative explanations for the hypothesized relationship between nominal fit beliefs and inherence-based explanations, the analyses
included two control variables. First, both nominal fit beliefs and use of the inherence heuristic might covary with *children’s age*. The nominal realism literature has consistently found a developmental trend away from realist reasoning about words (e.g., Brook, 1970; Osherson & Markman, 1975; Piaget, 1967). Similarly, use of the inherence heuristic has been shown to decline somewhat with age (Cimpian & Steinberg, 2014). Therefore, it is possible that the two variables of interest would correlate simply because they both happen to decrease with age (and not because beliefs about the suitability of words are rooted in an inherence heuristic in explanation). I therefore adjusted for children’s exact chronological age in the analyses. Note that this adjustment can also account, at least to some extent, for the potential influence of other variables that are strongly age-linked (e.g., cognitive ability, language skill).

Second, experience with multiple languages is usually accompanied by more sophisticated reasoning about the arbitrariness of words (e.g., Bialystok, 1987; Cummins, 1978; Ianco-Worrall, 1972; Rosenblum & Pinker, 1983). This factor might also affect the output of the inherence heuristic, insofar as children with a wider variety of linguistic and cultural experiences might have a wider variety of information (including extrinsic information) to draw on when generating an explanation (e.g., Kinzler & Sullivan, 2014). To address this potential confound, I measured children’s *exposure to multiple languages* by asking parents whether they spoke any languages other than English to their child at home. Children’s score on this dimension was the number of languages other than English that they had been exposed to (*M* = .19, *SD* = .43, range = 0 to 2).

**Data analyses.** The dependent variable (i.e., nominal fit beliefs) in this and all subsequent studies in Part 1 was non-normally distributed, as indicated by Shapiro-Wilk tests. I took two
steps to accommodate the nature of these data: First, I used bootstrapping techniques (1,000 replications) to derive standard errors, $p$ values, and bias-corrected and accelerated (BCa) 95% confidence intervals for the multiple regression estimates reported in Part 1 (Studies 1-6). Second, I used nonparametric, rank-based tests (namely, Mann-Whitney $U$) for any between-group comparisons (see Study 6).

**Results and Discussion**

If nominal fit beliefs are rooted in inherent reasoning, children’s reliance on inherence-based intuitions should predict unique variance in their nominal fit beliefs, above and beyond their chronological age and their exposure to multiple languages. Consistent with this hypothesis, children’s inherence heuristic composite scores significantly predicted their nominal fit scores in a regression analysis controlling for age and multilingualism, $b = .26$ [.15, .38], $SE = .05$, $p < .001$ (see Appendix A for the full correlation matrix and Appendix B for the means).

The other outcomes of this regression were consistent with prior work: First, nominal fit beliefs decreased with age, $b = −.13$ [−.19, −.06], $SE = .03$, $p < .001$. Second, exposure to multiple languages was associated with lower endorsement of nominal fit beliefs, $b = −.19$ [−.31, −.07], $SE = .07$, $p = .006$.

It is worth noting that the results above were replicated when just the questions that assessed explanations per se were used as a measure of children’s reliance on the inherence heuristic (that is, leaving out the questions about change and stability). This narrower measure of reliance on the inherence heuristic still predicted children’s nominal fit beliefs, $b = .17$ [.07, .28], $SE = .05$, $p < .001$. Thus, the most direct measure of children’s explanatory preferences also predicted unique variance in their beliefs about the fit between names and objects.
To conclude, these findings support the first prediction: namely, that children’s use of the inherence heuristic in making sense of the world predicts their belief that words and their referents are suitably connected.
Chapter 4: Study 2

Study 2 tested the second prediction of the current account—that reliance on the inherence heuristic should be correlated with endorsement of nominal fit beliefs even in adults. This prediction follows directly from the argument that nominal fit beliefs are the consequence of an inherence heuristic in explanation—a heuristic that is influential in adults’ reasoning as well, not just children’s (Cimpian & Salomon, 2014a, 2014b; Hussak & Cimpian, 2015; Salomon & Cimpian, 2014). Note, however, that this is a conservative test of the proposal because intuitions about nominal fit are likely to be (at best) weak among literate, educated adults. In fact, prior theories have explicitly maintained that intuitions about nominal fit are just an intermediate stage in the development of children’s metalinguistic knowledge (e.g., Brook, 1970; Piaget, 1967) and should thus be absent from adults’ reasoning altogether.

Additionally, I explored another possible interpretation of the predicted relationship: Perhaps nominal fit beliefs and inherence-based explanations are related simply because they co-occur in people with less complex cognitive styles (e.g., Stanovich, 1999; Stanovich & West, 2000). That is, individuals who—regardless of their cognitive abilities—prefer simple, black-and-white judgments or dislike effortful cognitive activity might be more likely to adopt both nominal fit beliefs and heuristic explanations, which would, in and of itself, give rise to a correlation between these two types of judgments. I explored this possibility by statistically adjusting for two widely used measures of cognitive style: the Need for Closure Scale (Kruglanski, Webster, & Klem, 1993) and the Need for Cognition Scale (Cacioppo, Petty, & Kao, 1984). I expected that participants’ explanatory preferences would predict unique variance in the extent to which they endorse beliefs about word–referent fit, above and beyond the two measures of cognitive style.
Methods

Participants. Undergraduate students from a large US university \((N = 126; M_{\text{age}} = 20.6\) years; 72% female) participated in this study.\(^1\) Approximately half of the students completed the study online, whereas the other half completed it in a psychology computer lab. An additional participant was tested but excluded from the final analyses for failing the catch items in the Inherence Heuristic Scale (see below).

Measures.

**Nominal Fit.** I developed an eight-item nominal fit measure \((\alpha = .65)\) to capture adults’ intuitions concerning whether words are particularly suitable or appropriate for their referents (see Table 1 for sample questions). Participants’ average rating across the eight items served as a dependent variable in the analyses. Unsurprisingly, average endorsement of nominal fit was fairly low overall \((M = 2.14 \text{ on a } 1–7 \text{ scale, } SD = 0.85, \text{ range } = 1.00 \text{ to } 4.38)\).

Participants were also asked to justify their scale ratings. These justifications provided a means of assessing the construct validity of the measure. If participants’ scale ratings truly tap into their nominal fit beliefs, then these beliefs should also be detectable in their open-ended justifications. The justifications were coded independently by two researchers (both of whom were blind to participants’ scale ratings) for the extent to which they expressed beliefs about the fit between words and referents \((1 = \text{no nominal fit}; 2 = \text{unclear}; 3 = \text{nominal fit})\). Inter-rater agreement was excellent \((\text{kappa} = .82)\).\(^11\) Coders’ ratings of these open-ended justifications correlated highly with participants’ scale ratings, \(r(124) = .72, p < .001\), which provides evidence

---

\(^1\) Due to a programming error, demographic information was collected from only 78 participants.

\(^{11}\) Twenty subjects were used for training the second coder, and thus reliability was calculated over 106 participants. I used a prevalence- and bias-adjusted kappa formula that was designed for ordinal coding scales (see Byrt, Bishop, & Carlin, 1993).
for the construct validity of the scale measure. That is, participants’ scale ratings seem to genuinely capture the extent to which they view word–referent pairings to be non-arbitrary. A sample of the justifications that revealed a belief in nominal fit can be found in Table 2. (Overall, 66.7% of participants provided at least one such justification.) To take an example, when asked whether *yaroo* could have been a suitable name for what is now called a *zebra*, one participant reasoned that, “Maybe it would work, but *zebra* sounds more suitable to me. Maybe it’s because of the stripes.”

Finally, to test whether the eight items indeed measure a *single* latent construct (namely, a belief in nominal fit), I conducted a confirmatory factor analysis using a correlated uniqueness model that allowed the error terms of similarly-structured items to correlate (Kenny & Kashy, 1992; Marsh, 1989). The hypothesized one-factor model provided a good fit to the data ($\chi^2[8, N = 126] = 12.10, p = .147; \text{RMSEA} = .064; \text{SRMR} = .039; \text{CFI} = .968$), consistent with the claim of a single underlying dimension (see Hu & Bentler, 1999, for additional information about interpreting fit indexes).

**Inherence Heuristic.** The Inherence Heuristic Scale consisted of 15 randomly ordered items that assessed participants’ tendency to rely on inherent explanations (e.g., “It seems natural that engagement rings typically have diamonds”). Extensive evidence concerning the construct validity, internal consistency, and factor structure of this scale can be found in Salomon and Cimpian (2014). The scale included four catch items (e.g., “It seems right to kill people for fun”) designed to screen out participants who were inattentive. Participants who failed two or more of these items (e.g., agreeing that it is right to kill people for fun) were excluded from the final analyses ($n = 1$). For this scale and the two described below, participants
marked their answers on a scale from 1 (disagree strongly) to 9 (agree strongly).

**Need for Closure.** The Need for Closure Scale consisted of 42 randomly ordered statements that assessed participants’ preference for order, simplicity, and quick, unambiguous judgments (e.g., “I usually make important decisions quickly and confidently”; Kruglanski et al., 1993).

**Need for Cognition.** The Need for Cognition Scale consisted of 18 randomly ordered statements that assessed participants’ motivation to engage in effortful cognitive activity (e.g., “I find satisfaction in deliberating hard and for long hours”; Cacioppo et al., 1984).

**Results and Discussion**

In line with the present proposal, participants’ scores on the Inherence Heuristic Scale were a significant predictor of their nominal fit beliefs even when adjusting for the two measures of cognitive style, $b = .24 [.09, .39], SE = .07, p = .002$ (see Table 3 for full regression results, Appendix A for the correlation matrix, and Appendix B for the means). Neither measure of cognitive style uniquely predicted participants’ nominal fit scores. Together, these findings suggest it is unlikely that the relationship between nominal fit beliefs and inherent explanations is an artifact of their co-occurrence in individuals with less complex cognitive styles. Rather, these results are more compatible with the current proposal, according to which beliefs about an inherent link between names and objects are simply an instantiation of a broader reliance on inherent explanations in making sense of the world.
Chapter 5: Study 3

Study 3 was modeled on Study 2, with two main changes. First, because animal names are occasionally non-arbitrary (e.g., onomatopoeia such as *cuckoo* or *chickadee*), I revised the nominal fit measure so that it asked about artifact names instead (e.g., *fork*). This change should provide a more conservative estimate of participants' nominal fit beliefs, as well as increase the generalizability of our conclusions.

Second, I included a new pair of control variables, which can address two more alternative explanations for the findings so far: Perhaps nominal fit beliefs and inherent explanations are related just because they happen to co-occur in individuals who (1) have lower fluid intelligence, or (2) are less creative. These individuals may be less able to imagine the myriad ways in which our words could have been different, and, independently, they may also be more likely to explain heuristically. In contrast to this alternative, I predict that the relationship between nominal fit and inherence will be present even when accounting for these cognitive variables.

Methods

Participants. Participants were recruited from Amazon’s Mechanical Turk ($N = 122; M_{age} = 36.2$ years; 51% female). In this and all subsequent studies with adults, Mechanical Turk participants were paid $0.75. An additional eight participants were tested but excluded from the sample because they failed the catch items in the Inherence Heuristic Scale.

Measures. The measures of inherent reasoning and nominal fit were administered as in Study 2, except that the nominal fit measure asked about artifacts instead of animals. As in Study 2, two researchers independently coded participants’ justifications for their scale ratings on the
nominal fit measure (kappa = .81). The codes assigned to these justifications were again strongly correlated with participants’ scale ratings, \( r(120) = .57, p < .001 \), suggesting that the latter truly tap into participants’ beliefs about the word–object link. Overall, 38.5% of participants provided at least one open-ended justification that indicated a belief in nominal fit.

**Fluid Intelligence.** I used a subset of Raven’s Standard Progressive Matrices (Raven, 1960), a standard test of fluid intelligence. Participants were required to solve 12 pattern completion problems. The problems were presented for 1 minute each, in increasing order of difficulty. Scores were calculated by summing the number of correct responses.

**Creativity.** Participants completed Gough’s (1979) Creative Personality Checklist, a widely used measure of creativity. Participants indicated which of 30 personality traits described them. Scores were calculated by adding the number of selected traits that had previously been found to correlate with creative achievements (e.g., “original”) and subtracting the number of selected traits negatively associated with such achievements (e.g., “cautious”). Trait order was randomized.

**Results and Discussion**

The findings were again consistent with the second prediction that adults’ inherent explanations would be related to their nominal fit beliefs. Participants’ scores on the Inherence Heuristic Scale significantly predicted their nominal fit beliefs in a regression analysis that adjusted for fluid intelligence and creativity, \( b = .23 [0.06, 0.41], SE = .08, p = .004 \) (see Table 3 for full regression results, Appendix A for the correlation matrix, and Appendix B for the means). These findings suggest that intelligence and creativity cannot account for the relationship between people’s explanatory tendencies and their endorsement of beliefs about word–object
fit. Instead, the results of this study provide more evidence supporting the claim that the inheritance heuristic is a source of nominal fit intuitions.
Chapter 6: Study 4

Study 4 provided a further test of the second prediction (namely, that adults’ reliance on heuristic explanations would be related to their belief in the non-arbitrariness of words). Additionally, Study 4 explored the possibility that this predicted relationship may be due to individual differences in either (1) counterfactual thinking (which might suppress both nominal fit beliefs and inherent explanations; e.g., Salomon & Cimpian, 2014; Stanovich & West, 1997) or (2) exposure to multiple languages, and by extension, multiple cultures (which, as argued in Study 1, could likewise lower both variables of interest; e.g., Bialystok, 1987; Kinzler & Sullivan, 2014).

Methods

Participants. Participants were recruited from Amazon’s Mechanical Turk (N = 122; M_age = 36.0 years; 57% female). An additional seven participants were tested but excluded from the sample because they failed the catch items in the Inherence Heuristic Scale (n = 6) or because they indicated during debriefing that they had not paid attention to the survey (n = 1).

Measures. The measures of nominal fit beliefs and inherent explanations were administered exactly as in the preceding study (Study 3). Inter-coder agreement for participants’ open-ended justifications on the nominal fit measure was again high (kappa = .72), as was the correlation between this coding and participants’ scale responses, r(120) = .68, p < .001. Overall, 51.6% of participants provided at least one open-ended justification that indicated a belief in nominal fit.

Counterfactual Thinking. To measure counterfactual thinking, I used a two-item scale from Stanovich and West (1997; e.g., “My beliefs would not have been very different if I had
been raised by a different set of parents” [reverse-scored]). Responses to each item were recorded on a scale from 1 (disagree strongly) to 9 (agree strongly).

**Multilingualism.** Participants were first asked how many languages other than English they were familiar with. Next, participants rated how fluent they were in each of these languages on a scale from 1 (limited familiarity [e.g., a year of instruction in school]) to 5 (native speaker [learned from birth]). Participants’ multilingualism scores were then calculated by adding up their fluency ratings across however many languages they were familiar with, other than English \((M = 2.07, SD = 2.70, \text{range} = 0 \text{ to } 17)\).

**Results and Discussion**

The results mirrored those from Studies 1–3. Once again, participants’ tendency to rely on inherence-based intuitions was a significant, unique predictor of their beliefs in the non-arbitrariness of words, even when controlling for their counterfactual thinking and their exposure to multiple languages, \(b = .27 \,[.09, .45], \, SE = .09, \, p = .003\) (see Table 3 for full regression results, Appendix A for the correlation matrix, and Appendix B for the means). These results reinforce the argument that the heuristic processes people rely on to explain the world in general also underlie how they make sense of the mapping between words and objects.
Chapter 7: Study 5

Unlike Study 1, the three adult studies so far (Studies 2–4) did not measure participants’ endorsement of actual inherent explanations. Rather, these studies tapped intuitions that follow from inherent explanations (e.g., “It seems natural that engagement rings typically have diamonds”; Salomon & Cimpian, 2014). For a more direct assessment, in Study 5 participants were presented with two sets of explanations that explicitly appealed to inherent features (see Table 4): a set that consisted of inherent explanations about a wide range of non-linguistic facts (which I termed Inherence–Global) and a set that focused more narrowly on the linguistic relation between names and objects (which I termed Inherence–Language). Using these new measures, I tested whether the global tendency to explain the world inherently is accompanied by a more specific tendency to explain the mapping between words and objects inherently, which then promotes intuitions about nominal fit. In the context of a mediation model, there should be a significant indirect path linking participants’ scores on the Inherence–Global scale with their endorsement of nominal fit via their scores on the Inherence–Language scale (see the diagram in Figure 1).

For a stronger test of this prediction, I added two covariates to the mediation model. The first was a version of the Cognitive Reflection Task (CRT; Frederick, 2005)—a widely used measure of heuristic thinking. I predicted that participants’ heuristic explanations will account for unique variance in nominal fit beliefs, above and beyond participants’ general propensity for heuristic thinking (as measured by the CRT). The second covariate assessed the extent to which participants view randomness and chance as influencing their lives (Levenson, 1981). Perhaps a tendency to underestimate the influence of chance leads people to explain the world by
appealing to inherent (rather than extrinsic) facts, as well as to endorse beliefs about nominal fit. Contrary to this alternative, I predicted that the hypothesized path linking inherent explanations with nominal fit beliefs will remain significant even when adjusting for participants’ understanding of chance.

Methods

Participants. Participants were recruited from Amazon’s Mechanical Turk (N = 128; M<sub>age</sub> = 33.9 years; 54% female). An additional 13 participants were tested but excluded from the sample because they failed more than one catch item across the scales (see below).

Measures. Participants completed five scales in randomized order. The measure of nominal fit beliefs was administered exactly as in Studies 3 and 4 (and thus focused on artifact names). Inter-coder agreement for participants’ open-ended justifications on the nominal fit measure was again high (kappa = .84), as was the correlation between this coding and participants’ scale responses, r(126) = .65, p < .001. Overall, 46.1% of participants gave at least one nominal fit explanation in their open-ended justifications.

The Inherence–Global Scale. The Inherence–Global scale consisted of eight explanations for diverse phenomena (see top of Table 4 for full list; α = .82). All explanations included explicit mention of inherent features. Participants indicated their agreement with each explanation on a scale from 0 (strongly disagree) to 10 (strongly agree). This measure also included two catch items with obvious answers (one true, one false).

The Inherence–Language Scale. The Inherence–Language scale consisted of eight explanations for specific word–object pairings (see bottom of Table 4 for full list; α = .93). Similar

---

12 Two participants did not provide demographic information.
to the Inherence–Global scale, all explanations included explicit mention of inherent features. Participants indicated their agreement with each explanation by using a sliding marker on a scale that ranged from 0 (completely disagree) to 100 (completely agree). This scale also included two attention checks that asked participants to move the sliding marker to one end of the scale. Any participant who missed more than one of the four attention/catch items across the Inherence–Global and Inherence–Language scales was excluded from the final sample (n = 13).

**Cognitive Reflection Test (CRT).** I used a version of the CRT (Frederick, 2005) that was superficially different from the original so as to avoid previous exposure among Mechanical Turk workers (Finucane & Gullion, 2010; see also Salomon & Cimpian, 2014). The task consisted of three word problems with salient and intuitive, but incorrect, answers. For example, one problem was as follows: “If it takes 2 nurses 2 minutes to measure the blood pressure of 2 patients, how long would it take 200 nurses to measure the blood pressure of 200 patients?” The easy, heuristic response to this problem is 200 minutes, but the correct response is actually 2 minutes. Responses were coded as either correct (= 1) or incorrect (= 0), and then summed across the three items. Higher scores on the CRT indicate more analytic (and less heuristic) reasoning.

**The Chance Scale.** To measure participants’ appreciation for the role of chance events, I used five items from Levenson’s (1981) Chance scale (e.g., “To a great extent, my life is controlled by accidental happenings”). Response options ranged from −3 (strongly disagree) to +3 (strongly agree), with no midpoint. The scores were summed across the five items, and 15 was then added to the grand total so as to arrive at a possible range of 0 to 30 (with greater scores indicating greater appreciation for the importance of chance).
Results and Discussion

According to the present proposal, participants’ heuristic tendency to make sense of the world in inherent terms should be apparent in their reasoning about word-object relations as well, which should in turn lead participants to see a special fit between words and their referents. This predicted path (global inherence → inherence about words → nominal fit) was indeed significant in a bootstrapped product-of-coefficients mediation analysis, $ab = .05 \,[.02, .10], SE = .02$, Sobel test $p = .020$ (Hayes, 2013; see Appendix A for the full correlation matrix and Appendix B for the means). The direct effect from global inherence to nominal fit was nonsignificant, $c' = .01 \,[-.09, .11], SE = .05$, $p = .828$.

Next, to test whether the proposed model provides a better fit to the data than other possible models of the relationships between these three variables, I switched the dependent variable (nominal fit) and the mediator (inherence about words). According to my proposal, people’s inherent explanations of word–object mappings precede (and give rise to) their beliefs about nominal fit, not vice-versa. Thus, I predicted this alternative model would be less compatible with the data. Indeed, the path tested in this model (global inherence → nominal fit → inherence about words) was not statistically significant, $ab = .05 \,[-.03, .14], SE = .04$, Sobel test $p = .257$. This result provides additional confidence in the hypothesized model.

In a separate set of analyses, I added the CRT and the Chance Scale as covariates to the

---

13 Surprisingly, the zero-order correlation between the Inherence–Global scale and the nominal fit scale did not reach significance, $r(126) = .11$, $p = .227$. Further inspection of the data revealed that the relationship between these variables was not linear. Rather, this relationship had a significant (and negative) quadratic component, $b = -.03 \,[-.06, -.001], SE = .02$, $p = .040$, in addition to a significant (and positive) linear component, $b = .46 \, [.10, .83], SE = .19$, $p = .014$. Specifically, the two variables showed a robust positive linear relationship over most of the Inherence–Global scale, a relationship that flattened off at the high end of this scale (hence the negative quadratic component). In sum, the non-significant correlation coefficient masks a much stronger positive relationship between Inherence–Global and nominal fit scores over most values of the Inherence–Global scale.
original mediation model. The predicted path linking global reliance on inherent explanations with endorsement of nominal fit beliefs through reliance on language-specific inherent explanations remained significant, $ab = .04 [.01, .08]$, $SE = .02$, Sobel test $p = .045$ (see Figure 1 for full results). Moreover, switching the mediator and the dependent variable again led to a non-significant indirect path, $ab = .03 [-.03, .10]$, $SE = .03$, Sobel test $p = .436$.

In sum, these results provide further evidence for the prediction that adults’ reliance on inherent explanations to make sense of the world, and of word–object mappings in particular, may lead to intuitions that words are a good fit for their referents.
Chapter 8: Study 6

The evidence thus far suggests that the belief in a fit between words and their referents in both children and adults is related to their inherence-based reasoning, even when controlling for a variety of alternative factors. In this final study of Part 1, I tested the causal direction of this relationship (the third prediction): Would manipulating participants’ tendency to explain inherently lead to subsequent changes in their endorsement of beliefs about the inherent suitability of words?

Methods

Participants. Participants were recruited from Amazon’s Mechanical Turk and an undergraduate subject pool (N = 500; M_age = 35.6 years; 61% female). All participants completed the study online. An additional 61 participants were tested but excluded from the sample for the following reasons: (1) because they indicated they had not paid attention during the study (n = 5), (2) because they guessed the purpose of the study when asked at the end of the session (n = 17), or (3) because they failed a comprehension check (n = 39) (see below for details).

Procedure and Materials. Participants were randomly assigned to one of two conditions that differed in whether they undermined or reinforced participants’ tendency to explain via the inherence heuristic (the Anti-Inherence and Pro-Inherence conditions, respectively). After the manipulation, participants completed a distractor task (which helped disguise the manipulation), followed by a measure of nominal fit and a brief manipulation check.

Manipulation. Participants read an article as part of a task on “reading comprehension and memory.” In both conditions, the article was titled, “Ever Wonder, ‘Why Do We Do It That Way?’” and focused on the example of why recycling bins are blue. In the Anti-Inherence
condition, the article explained that the color blue was a “historical accident,” having been chosen for no particular reason by the founders of the first recycling program in southern Ontario. The article went on to provide more examples of extrinsic forces that shape the way we do things (e.g., marketing campaigns, influential people). This information was designed to undermine participants’ typical explanatory intuitions, which focus on the inherent natures of the things explained rather than extrinsic factors. The article in the Pro-Inherence condition was similar in many respects to that in the Anti-Inherence condition (e.g., length, layout, and examples) except that it claimed blue was chosen because of its visibility and its ability to endure damage from the sun (inherent properties). The article then went on to provide more examples illustrating how the way we do things is in accord with the inherent natures of those things, and is thus optimal (e.g., scientists conduct experiments to determine how to do something optimally). This information was designed to reinforce participants’ typical, inherence-based explanatory intuitions.

After reading one of these articles, participants in both conditions completed a series of reading comprehension and memory questions, in keeping with what they had been led to believe was the purpose of the task. In addition to camouflaging the manipulation, these questions allowed us to test whether participants read and understood the information in the manipulation articles. In particular, one of the comprehension questions asked participants to choose which of four statements best summarized the main argument of the article. The correct answer should have been obvious to anyone who read the passage. Thus, this question served as an attention check as much as a comprehension check. Participants who answered this question incorrectly were excluded from subsequent analyses (n = 39, as mentioned above).
**Distractor.** Following the manipulation, participants completed a three-minute distractor task consisting of four spatial puzzles.

**Nominal Fit.** I used the measure of nominal fit beliefs from Studies 3–5.

**Manipulation Check.** Finally, I asked participants to rate their agreement with four items designed to check whether the manipulation was effective in influencing their explanatory tendencies (e.g., “A lot of things that are true today are the way they are because of historical accidents that could have been otherwise” [reverse-coded]). These items were always presented last, just before the demographic questions. Participants marked their answers on a 1 (disagree strongly) to 9 (agree strongly) scale.

**Results and Discussion**

The manipulation check revealed the expected condition difference, with the Pro-Inherence participants showing a stronger preference for inherence-based reasoning than the Anti-Inherence participants ($M_s = 5.09$ and 4.22, respectively; both $SD_s = 1.19$), Mann-Whitney $Z = 8.15$, $p < .001$. The key question, however, was whether this manipulation would affect participants’ beliefs about the word–object link. As predicted, participants in the Pro-Inherence condition showed stronger intuitions about nominal fit ($M = 2.21$, $SD = 0.86$) than participants in the Anti-Inherence condition ($M = 2.08$, $SD = 0.94$), Mann-Whitney $Z = 2.08$, $p = .038$. In further support of the current claim, the effect of the Pro- vs. Anti-Inherence manipulation on nominal fit beliefs was mediated by participants’ inherence-based reasoning (as measured by the manipulation check), $ab = .12 [.06, .19]$, $SE = .03$, Sobel test $p < .001$ (see Figure 2 for the full mediation model). Once again, the direct effect was nonsignificant, $c' = .01 [-.15, .12]$, $SE = .09$, $p = .882$, suggesting that the influence of the manipulation on participants’ nominal fit beliefs was
fully mediated by the effect of the manipulation on participants’ inherent reasoning.

Thus, experimentally manipulating participants’ reliance on the inherence heuristic had a downstream effect on their tendency to see words as fitting with their referents. These findings support the proposal that this explanatory process gives rise to intuitions about the non-arbitrary nature of words.
Chapter 9: Discussion of Part 1

Theoretical Contributions

The studies from Part 1 make two important contributions. First, they provide evidence for the psychological reality of the belief that words fit, or are somehow suitable for, their referents. Nominal fit beliefs were present not only in children’s reasoning but were also articulated quite explicitly by many literate American adults (e.g., “The name sounds fitting for what it’s meant to do”; see Table 2 for additional examples). The second contribution of this work is that it proposes and tests a potential mechanism underlying these beliefs about nominal fit. I hypothesized that these beliefs are a product of how people make intuitive sense of the world more generally. Recent arguments and evidence suggest that many everyday, in-the-moment explanations are generated using a heuristic process that leads people to routinely appeal to inherent facts, which are highly accessible in memory (Cimpian, 2015; Cimpian & Salomon, 2014a, 2014b; Cimpian & Steinberg, 2014; Salomon & Cimpian, 2014). Given that this process is invoked to explain a wide range of observations, it is reasonable to suppose that it would also shape people’s understanding of why objects have the names they do. Due to this heuristic, people might assume that word–referent pairings are explained by inherent aspects of the words or referents themselves rather than by social conventions established in the distant historical past. For example, as one of the adult participants suggested, perhaps a zebra’s stripes are part of the reason for the name zebra, which suggests that this name is fitting. The six studies reported here support the claim that such nominal fit beliefs stem from the broader tendency to explain heuristically. It is also worth noting that the results of these studies converged on the same conclusion despite considerable variation in the designs they employed.
(correlational vs. experimental), in the characteristics of their participants (4- to 7-year-olds vs. undergraduates vs. Mechanical Turk workers), and in the questions they used to assess participants’ nominal fit beliefs and their reliance on the inherence heuristic. In summary, these results highlight the promise of the proposal that use of the inherence heuristic is at the root of beliefs about the fit between objects and their names.

**Why Do Beliefs about Nominal Fit Persist into Adulthood?**

The results of these studies suggest that even adults show traces of a belief that words are suitable for their referents. Why would this be? Surely the (literate American) adult participants would have had plenty of opportunities to realize that language is arbitrary. In today’s world, for example, it is nearly impossible to avoid exposure to different languages and thus to different ways of referring to the same objects; many of the subjects even *spoke* more than one language. Why weren’t such experiences sufficient to dispel any doubts about the arbitrariness of language?\(^{14}\) I argue that, in and of themselves, linguistic differences of this sort do not *necessarily* undermine the idea of nominal fit. It is not logically inconsistent to believe that names fit their referents while simultaneously acknowledging that different languages use different words for the same thing. The fact that, say, *dog* and *perro* refer to the same animal does not preclude the possibility of explaining both inherently: Perhaps the two words are rooted in, or inspired by, two different aspects of the animal and are thus both a good match. Words can fit an object in many respects. More generally, it is by no means obvious that the only way to make sense of the existence of multiple words for the same thing is to assume that words are largely arbitrary. **Arriving at this conclusion might actually require more systematic effort**

\(^{14}\) Exposure to different languages may be informative up to a point: In Study 1, children with such exposure showed greater understanding of the arbitrariness of words.
than is typically expended in the course of ordinary cognitive activity.

A related reason for the persistence of nominal fit beliefs might be that the evidence that could potentially weaken such beliefs is often not the most accessible when we’re looking for a quick, in-the-moment answer. Although many people of course realize that, say, a bowl is called by very different names in other languages, this fact may not be the first thing that comes to mind when thinking about why a bowl is called a bowl. Rather, retrieval may be dominated by inherent facts about the entities in the explanandum (the object and the word), which then gives rise to an inheritance bias in the heuristic explanations generated. This is, I would argue, analogous to how many people are aware of counterexamples to familiar non-linguistic regularities (e.g., engagement rings without diamonds) and still often fail to factor this less-accessible evidence into their explanations (Cimpian & Salomon, 2014a, 2014b; Salomon & Cimpian, 2015). In both cases, the explananda are typically understood as being rooted in inherent facts and thus as being natural and how things should be (e.g., Hussak & Cimpian, 2015; Tworek & Cimpian, 2015).

**Nominal Fit Intuitions: Shallow or Sophisticated?**

So far, I have assumed that nominal fit beliefs are the product of shallow, heuristic processes. One might argue, however, that the participants could have also endorsed these beliefs for sophisticated, well-thought-out reasons. For example, if a participant assumes that (1) animal names in English are derived from their scientific names in Latin, and (2) scientific names function as descriptions of sorts (rather than being entirely arbitrary), perhaps it is reasonable to also assume that current English names are, in this roundabout way, not entirely arbitrary. Similarly, perhaps participants gave nominal fit responses because they extrapolated from the—
occasionally non-arbitrary—ways that new words are introduced into English usage. For example, the photographs that people take of themselves are now called *selfies* (the word of the year in 2013, according to the Oxford Dictionaries). Participants may have reasoned by analogy that, just as neologisms like *selfie* or *podcast* seem to provide meaningful descriptions of their referents, established words such as *zebra* or *candle* might likewise provide a good fit for their referents.

The view that participants’ nominal fit responses stemmed from sophisticated chains of inferences such as those illustrated above is contradicted by at least two aspects of the data. First, endorsement of nominal fit beliefs in the current studies was positively correlated with endorsement of explanations that are heuristic rather than the result of careful thought. It seems unlikely that the same participants would effortfully puzzle their way through one set of questions (i.e., the nominal fit questions) and rely on effort-saving heuristics for another (i.e., the inherence heuristic questions). Second, the pattern of relationships between endorsement of nominal fit intuitions and other variables measured in these studies is exactly the opposite of what would be expected under this alternative. For instance, endorsement of nominal fit was negatively correlated with children’s age (Study 1), with adults’ fluid intelligence (Study 3), with their counterfactual reasoning ability (Study 4), and with the degree to which they reasoned analytically (Study 5). These results suggest that the source of intuitions about nominal fit is less examined and rational than proposed by this alternative hypothesis and more consistent with the proposed interpretation of participants’ responses.

**Future Directions**

On a final note, I hope that one of the lasting contributions of the present work will be to
spark new research on people’s reasoning about the arbitrariness of language. There is much important work left to be done here. For instance, one potential avenue for future work is to investigate the relationship between the present findings and other lines of research that similarly suggest people underestimate how arbitrary language is. To illustrate, people also tend to assume that (1) idioms such as “spill the beans” or “take the bull by the horns” connect quite transparently, rather than arbitrarily, with their meanings (e.g., Keysar & Bly, 1995) and that (2) the grammatical gender of a noun (in languages that employ this syntactic device) provides a match to the characteristics of its referent (e.g., Boroditsky, Schmidt, & Phillips, 2003). Are the similarities between the data presented here and these other studies coincidental, or are they perhaps due to an underlying similarity in mechanism? And, if the latter alternative is correct, is the common mechanism explanation-based?

To conclude Part 1, the above studies identify a surprisingly robust belief that words fit their referents. Additionally, they provide evidence that this belief is due to an inherence heuristic in how people generate explanations. This work adds an important piece to our understanding of people’s basic conceptions about language, and it illustrates the extent to which heuristic explanations underlie people’s fundamental theories about how the world works.
Chapter 10: Part 2. Inherent Reasoning Provides Important Foundations for Children’s Essentialist Beliefs

People commonly believe (at least implicitly) that members of natural and social categories (e.g., lions, boys, African Americans) share an internal substance or “essence” that causes their category’s typical properties (e.g., having manes, being interested in fixing things; e.g., Gelman, 2003; Dar-Nimrod & Heine, 2011; Kinzler & Dautel, 2012; Medin & Ortony, 1989). This belief is revealed in phrases such as, “it’s in their bones” or “it’s in their blood.” Phrases of this sort emphasize the extent to which a trait is stable and predictive. For instance, in a news article about the caste system in India, someone explained, “it’s in their blood and they carry it with them all of the time” (D’Souza, 2015) Powerful essentialist beliefs like those about castes are common across a wide variety of categories. Furthermore, essentialist beliefs are not particular to specific cultures (e.g., Birnbaum, Deeb, Segall, Ben-Eliyahu, & Diesendruck, 2010; Gelman, Meyer, & Noles, 2013; Rhodes & Gelman, 2009; Waxman, Medin, & Ross, 2007) nor to particular ages (e.g., Gelman, 2003; Gelman, Heyman, & Legare, 2007). Beyond their being ubiquitous, essentialist beliefs encourage overgeneralization of traits across category members and insensitivity to nuances that exist within the category. Given the broad scope of influence that essentialist beliefs have on people’s reasoning, it is important to understand the origins of essentialist beliefs.

Here, I propose that a key part of the foundation on which essentialist beliefs are built is set by inherent reasoning. Before discussing the proposal in detail, I will first describe important past work on essentialism and its origins. I will then outline why general explanatory heuristics might play an important role in setting up essentialist beliefs. Finally, I will describe three studies
that lend support to this proposal and make important headway in the investigation of the origins of essentialist beliefs.

**Various uses of the term “essentialism”**

To be clear, when I use the term “essentialism” or “essentialist beliefs,” I will be referring to psychological essentialism specifically (i.e., the way people mentally represent categories and their members), rather than metaphysical or biological essentialism (i.e., the actual existence of essences in the real world). Beyond this simple clarification, I need to further specify what beliefs I am investigating the origins of, as even within the literature on psychological essentialism, the term “essentialism” has been used to refer to various beliefs. To do so, I will first outline some of the nuances in way the term is used.

There are multiple potential components of an essentialist belief; however, theorists put these components together in different combinations, all of which are labelled “essentialism.” Some use a relatively bare definition, whereby an essence is a central part of a concept, its underlying reality, which is sometimes responsible for certain properties (however, exceptions in displaying these properties are not indicative of a lack of essence; e.g., Medin & Ortony, 1989). Others agree that the essence is a non-specific “placeholder” of sorts, in that people don’t know exactly what it is, but these theorists believe that more can be said about how people conceive of essences (e.g., Bloch, Solomon, & Carey, 2001; Dar-Nimrod & Heine, 2011; Gelman, 2003; Gelman & Wellman, 1991; Haslam, Rothschild, & Ernst, 2000; Keller, 2005; Kinzler & Dautel, 2012; Newman & Keil, 2008; Rosengren, Gelman, Kalish, & McCormick, 1992; Taylor, Rhodes & Gelman, 2009; Uhlmann & Zhu, 2013; Yzerbyt, Corneille, & Estrada, 2001). Qualifications that are commonly included are a belief that the essence is (1) substantial and internal, (2) that it can be
used to identify individuals’ underlying, true and natural category membership (at least in principle), and (3) that this same substance is causally responsible for category members’ typical properties (Gelman, 2003).

In most cases, essentialist beliefs are said to span natural categories (such as animals and chemical elements) and social categories (such as gender and race; e.g., Gelman, 2004; Gelman & Heyman, 1999; Gelman & Wellman, 1991; Haslam, Rothschild, & Ernst, 2002; Rhodes, Leslie, & Tworek, 2012). However, others extend this belief to cover artifact categories (e.g., Bloom, 1996). The argument is that, if one were to look at what an object was intended to be, the intentions of the creator would mirror important aspects of what constitutes an essentialist belief for natural and social kinds (e.g., Bloom, 1996). For instance, all objects that were intentionally and successfully created to be a vase are in fact vases. From this perspective, once someone knows the intention behind an object’s successful creation they will always know its true kind membership—in this way, the intention of the creator is similar to the essence for natural categories. Note, however, how different these two notions of an essence are: One is substantial and located within each member of the category, while the other is neither.15

Overall, people’s essentialist beliefs about artifacts and people’s essentialist beliefs about natural and social categories seem to stem from different notions of an essence. And even within

---

15 Others still, discuss individuals as having their own essences (e.g., Hitler had a “Hitler-essence”; Leslie, 2013; Newman, Diesendruck, & Bloom, 2011). Individual essences are most often discussed in contexts where an individual contaminates objects with their essence. The remnants of the essence on the objects is then capable of contaminating other individuals with which it comes in contact, and in turn the essence transferred might cause the contaminated individual to have similar properties as the individual from which the essence originated. Importantly, there are key differences in a belief in this type of essence and category essences (e.g., the transferred traces of the essence are not thought to cause the identity of the contaminated individual to change, and the traces can imbue graded degrees of the original individual’s properties rather than making the contaminated individual mirror the original individual). Because of the differences in the properties of essences when considering an individual essence and a category essence, in this paper I will focus on category essences only.
essentialist beliefs about natural and social kinds, the specifics of the belief researchers are arguing about are inconsistent. The variability in what one might be referring to when discussing essentialist beliefs necessitates being specific about which beliefs one is in the business of investigating the origins of, in part because the different ideas of an essence might stem from different cognitive mechanisms. Here, I focus on finding the origin of a belief that strongly influences people’s daily reasoning and is widely studied in the field—the belief that members of natural and social kinds share a microstructural, internal, category-indicative, and causally powerful essence (Gelman, 2003; Gelman & Wellman, 1991; Keil, 2003; Keil, 1989; Rhodes et al., 2012). I will use the term “essentialism” to represent this specific belief.

Essentialism strongly influences the predictions and inferences people make daily (e.g., Gelman & Heyman, 1999; Haslam et al., 2002; Kinzler & Dautel, 2012; Leslie, 2013; Meyer, Leslie, Gelman, & Stilwell, 2013; Waxman et al., 2007; Yzerbyt et al., 2001), and also gives people a sense that they understand the mechanism that underlies many category features (see Keil, 2006; Rozenblit & Keil, 2002; Trout, 2002 for how confidence accompanies a sense of mechanistic understanding). For instance, not only might people predict that the dog at the park will bark and have similar propensity for certain diseases as other dogs, but their essentialist beliefs give them a more concrete sense of knowing why—because dogs share an internal substance. While confidence in daily inferences can help people go about their day smoothly, such confidence can also encourage people to overlook nuanced differences among category members and treat members of a group as if they were interchangeable copies of one another. For instance, people’s essentialist beliefs will justify their assumptions that a female will be best

---

16 This belief can include the idea that DNA fills the role of an essence, though it need not.
at taking care of a crying child and a male will be best at fixing the sink (Dar-Nimrod & Heine, 2011; Haslam et al., 2002; Keller, 2005). While on the surface these consequences may seem relatively small, on a larger scale they can give rise to systematic prejudice and harmful stereotyping. It is thus critical to better understand the initial source of essentialist beliefs to know when and why these consequences arise, and how most effectively to avoid them when they are harmful.

**Possible Sources of Essentialist Beliefs**

Multiple proposals have been discussed for the origins of essentialist beliefs; however, as a field we remain far from a comprehensive understanding. In the next section, I will outline some of the primary contenders for the origins of essentialist beliefs and why they fall short of being able to explain the available data. Subsequently, I will outline the current proposal of an important part of the foundation for the emergence of essentialist beliefs—a foundation found in general explanatory processes.

First, it is possible that essentialist beliefs come from an accurate perception of essences in the world (see Gelman, 2003; Gelman & Hirschfeld, 1999 for discussion of these arguments). However, contrary to many people’s essentialist beliefs, there is no scientific evidence of an internal substance that causes all of the category-typical properties that kind members share. Even DNA cannot accurately predict a person’s personality traits, ethnicity, sexual orientation, disabilities, and so on (see Keil, 1995 and Leslie, 2013 for excellent arguments against the existence of essences in the world). While there are some features in the world that help to distinguish category members, these fail to categorize with 100% accuracy, and often do not play a causal role in a member’s properties.
Second, essentialist beliefs may be conveyed socially, through other people—perhaps through language or cultural messages more generally (e.g., Gelman, Taylor, & Nguyen, 2004). Consistent with this possibility, children are more likely to endorse essentialist beliefs about a novel social category after hearing several generic facts about it (i.e., facts that apply to the entire category; Rhodes et al., 2012). Importantly, however, generic statements such as “girls like pink” do not in and of themselves convey the notion of an essence. Saying “girls like pink” does not in and of itself convey a notion of an internal cause of this preference. Therefore, a general concept of an essence would need to have already been entertained before children’s essentialist beliefs would be affected by such language. Furthermore, people rarely, if ever, discuss essentialist beliefs per se (e.g., that girls all share something inside that makes them like pink), through generic language or otherwise (e.g., Gelman et al., 1998). More plausibly, generic facts might influence which categories people essentialize, once essentialist beliefs have already been entertained in some form. Thus, it seems as though learning from others is not a primary source of essentialist beliefs.

Another possibility is that our cognitive systems are structured in a way that fosters the development of essentialist beliefs. There are at least two distinct ways they might do so. Some theorists argue that a domain-specific, innate module facilitates essentialist inferences about biological kinds specifically (e.g., Atran, 1998; see also Barrett, 2001; Gil-White, 2001). This module theory holds that this predisposition to essentialize biological kinds might then generalize to kinds that are not biologically based (e.g., Spanish speakers and water; Gil-White, 2001; Haslam et al., 2000). However, the view that biological essentialism is the product of an innate module does not fit with two aspects of the available data. First, the claim that children
use their essentialism for biological kinds to infer essentialist views about social kinds predicts that essentializing biological kinds would come online prior to that for social kinds. Children would need time to draw an analogy from biological to social categories. However, essentialist beliefs about animal categories and social categories seem to emerge around the same time in children (e.g., Birnbaum et al., 2010; Gelman, 2003; Taylor, Rhodes, & Gelman, 2009). Second, it is difficult for the module view to explain the variance in essentialist beliefs, both in the tendency to essentialize some biological (and social) categories more than others, as well as in the tendency for some individuals to essentialize more than others across the board (see Haslam et al., 2000, 2002, for evidence of this variability).

Instead of a biological reasoning module, it might be the case that more general cognitive processes support the development of essentialist beliefs (Gelman, 2003; see also Keil, 1995). Susan Gelman (2003, pp.313-314) thoughtfully outlined multiple early-emerging, domain-general “root capacities” that support essentialist beliefs. These include the ability to distinguish appearance from reality, the ability to make inductive inferences from property clusters, a belief in causal determinism, the ability to track identity over time, and a tendency to defer judgments to experts. Other likely elements are the ability to make inferences about individuals based on information about others of the same kind (e.g., Gelman & Markman, 1987; Graham, Kilbreath, & Welder, 2004), and an understanding of internal energy (e.g., Inagaki & Hatano, 2004; Leslie, 1995; Luo, Kaufman, & Baillargeon, 2009). The “domain-general processes” argument is that as multiple domain-general cognitive abilities combine together to form essentialist beliefs. The

---

17 It is highly unlikely that children would be able to draw such a complex analogy at the young ages that they express social essentialist beliefs (i.e., preschool) without very explicit aligning and discussion of essentialist beliefs by others (Gentner, Anggoro, & Klibanoff, 2011). And there is no evidence that this occurs (Gelman et al., 1998).
idea that essentialism emerges from more basic cognitive abilities and propensities has the
advantage of allowing for wide variability in the categories and properties that are essentialized,
as well as in the extent to which they are essentialized, depending on how the precursor abilities
develop and combine within individuals.

While it is reasonable to expect that the abilities and processes discussed above support
essentialist beliefs, we don’t yet seem to have a sufficient list of abilities and beliefs that would
bring children all the way to essentialism. For instance, it is difficult to see how a child might
come to believe specifically that the same internal substance (or essence) causes all of an
individual’s category-typical properties and is in principle able to indicate true category
membership from the idea that reality may differ from what appearances suggest, the ability to
make inductive inferences, and the belief that agents have an internal energy. There are details
in a fully developed essentialist belief that seem unable to form out of the current precursors
outlined in the literature. To gain a better understanding of the cognitive underpinnings of
essentialist beliefs, it is important to test additional early emerging processes that might
contribute to the eventual formation of essentialist beliefs. Here, I identify one key component
of this developmental process that has not yet been discussed in previous work—namely, the
process of generating intuitive explanations.

**The current proposal: Inherent reasoning facilitates the development of essentialism.**

Essentialist beliefs are often called upon to explain properties of natural and social kinds. Thus,
these beliefs support a large portion of people’s explanations of their surroundings—they
provide a framework for making sense of the world (Barrett, 2001; Gelman, 2003; Gelman &
Wellman, 1991; Medin & Ortony, 1989). For instance, essentialist beliefs provide explanations
for why a dog barks, and why girls are nurturing—“it’s just in their blood.” But why do our explanations call upon the idea of an “essence” in these cases? Perhaps in looking at the process of explaining in more detail, new light might be shed onto why these essentialist beliefs are used as we make sense of the world. For instance, it might be the case that aspects of the process of explanation itself are structured in a way that promotes use of essentialist beliefs. In the next section, I will propose that the inherence bias in people’s everyday explanations plays a role in the development of essentialist beliefs. I will then outline evidence consistent with this proposal.

The role of inherent explanations in the development of essentialist beliefs. Because explanations can have a powerful influence on how people (both children and adults) make sense of the world, a skew in the content of those explanations will likely translate into a skewed view of the way the world works. In line with this idea, children’s reliance on the inherence heuristic may be one source of their beliefs that words are non-arbitrarily matched with their referents (see Part 1), that social inequalities are fair (Hussak & Cimpian, 2015), and that current patterns of behavior signal the way people should behave (Tworek & Cimpian, in press). As children repeatedly appeal to inherent features when making sense of their world, over time and in combination with other root capacities, coherent beliefs might form based on these explanations—perhaps more specifically essentialist beliefs. To speculate, inherent reasoning might draw children to the more general belief that it is usually something about the thing itself that makes it the way it is—likely something stable and integral to the thing being what it is, as these are the types of features that tend to be used in many explanations. In combination with an early understanding of internal energy, a developing sense of causal reasoning, as well as accumulating knowledge about biological insides, inherent reasoning might precipitate the
essentialist intuition that there is some inherent, internal feature that explains a category’s features. While testing the precise mechanisms by which inherent reasoning facilitates the emergence of essentialist is beyond the scope of this paper, here I describe the first steps in testing whether inherent reasoning lays an important part of the foundation for the development of essentialist beliefs.

To clarify, inherent features and essences are similar, but distinct. They are similar because an essence is an inherent feature—it is completely within and about the member to which it belongs. Additionally, if an individual’s essence was changed, the individual itself would change—a key indication that a feature is inherent to the entity at hand (Cimpian & Salomon, 2014a, 2014b). Importantly, however, an essence has characteristics that other inherent features do not have. The essence is microstructural, internal, causal, and defining of an individual’s category membership. In this way, inherent features and essences are not the same thing; inherent features form a much broader set. Also consistent with the idea that essentialism and inherent explanations are not the same construct, reasoning about essences appears to come online later in development compared to children’s more general inherent reasoning (e.g., Cimpian & Steinberg, 2004). This difference in emergence suggests that essentialism requires additional information or cognitive abilities on top of general inherent reasoning. In all, looking at an essence as a subtype of an inherent feature (an internal, category-specific, causal inherent feature) is broadly consistent with the view that essentialist beliefs emerge in part as a product of inherence-based reasoning.

**Recent Findings in line with the Proposal**

Prior work on essentialist beliefs in adults is consistent with the proposal that general
inherent reasoning facilitates the initial development of essentialist beliefs. For instance, adults’ endorsement of inherence-based explanations about conventions (e.g., wedding dresses being white) predicts the strength of their essentialist beliefs about social groups (e.g., Asians; Salomon & Cimpian, 2014). Similarly, even within individuals, adults’ essentialist beliefs shift depending on their reliance on inherent intuitions in the moment: When adults’ reliance on inherent intuitions was experimentally reduced, participants’ essentialist beliefs were consequently weaker (Salomon & Cimpian, in preparation). Moreover, when adults’ available cognitive resources were taxed with the use of a speeded task (increasing their likelihood of using heuristic processes), they showed heightened essentialist beliefs compared to participants who were forced to delay reporting their responses (Edison & Coley, 2013). This is consistent with our proposal because inherent explanations are often a product of a heuristic process (Cimpian & Salomon, 2014a, 20014b; Hussak & Cimpian, 2015); thus, if inherent explanations give rise to essentialism, we might expect essentialist beliefs would be more strongly endorsed when people are reasoning heuristically.

While these findings are broadly consistent with the proposal that inherent reasoning influences essentialist-based intuitions, they cannot speak to the influence of inherent thinking on the early development of essentialist beliefs. The adult data is consistent with the possibility that the link between the inherence heuristic and essentialism develops later in life. Given the findings reviewed so far, it remains possible that children’s essentialist beliefs develop independently of their explanatory tendencies. For a more direct test of the role of inherent explanations in the development of essentialism, it is important to investigate the early stages of development, when essentialist beliefs are first emerging.
Current Studies

To test whether inherent reasoning lays part of the foundation for the development of essentialist beliefs, I focus my investigation on the developmental time point at which essentialist beliefs emerge (4 years of age; e.g., Gelman, 2003). Here, I provide the first test of this link across three studies with young children. First, I tested whether 4-year-old children’s inherence-based intuitions about conventions would predict the strength of their essentialist beliefs about animals (Study 7). Next, I tested the causal nature of this relationship by either encouraging or discouraging 4- to 7-year-olds’ reliance on inherence-based intuitions, and measuring the subsequent changes in the strength of their essentialist beliefs (Studies 8 and 9).
Chapter 11: Study 7

To test whether inherence-based reasoning lays the foundation for essentialist beliefs, I first investigated whether individual differences in children’s reliance on inherence-based reasoning predict the extent to which they endorse essentialist beliefs around the time at which children begin to express essentialist beliefs (4 years of age; Gelman, 2003).

Methods

Participants. Four-year-old children (N = 64, half boys and half girls; M = 4.4 years, SD = 0.32; range = 4.0 to 5.0) participated in a quiet room at a research lab or in their school. They were socioeconomically diverse, and most were European American.

Procedure. Children were given two sets of questions: one that captured their inherence-based reasoning, and another that captured their essentialist beliefs. The order of the two sets was counterbalanced across children.

Inherence-based reasoning. To assess the extent to which children explain general phenomena using inherent features, I used 4 different types of questions used in previous studies (Cimpian & Steinberg, 2014). The first three questions are exactly those used in Study 1: One asked children to rate inherent and non-inherent explanations, and two tested inferences that would follow from having reasoning inherently (that the phenomenon would be consistent over time, and that society has little influence on it). A fourth question was added here that also tested a belief that would follow from inherent reasoning—that the phenomenon at hand is non-arbitrary. This question is the same as that used in Study 1 as a measure of nominal fit beliefs. Each question was asked about two different social conventions (sampled from the following three conventions, in counterbalanced order across children: school buses being
yellow, birthday cakes having candles, or coins being round). I measured inherent reasoning about conventions involving artifacts, while essentialist beliefs were measured in a wholly different domain – animal categories. This way, any link found between these constructs cannot simply be attributed to shared content. Each question received a score that could range from 0 to 1, where higher scores represented more inherence-based reasoning. I then averaged the scores to create an inherence composite for each child.

**Essentialist beliefs.** To capture the strength of children’s essentialist beliefs, I used a variety of questions used in the literature: measurements of inductive potential (Gelman & Markman, 1987), stability of a property (Gelman & Heyman, 1999; Gelman, Heyman, & Legare, 2007), innate potential (Gelman & Wellman, 1991; Gelman, Heyman, & Legare, 2007), and the importance of insides for determining category membership (Gelman & Wellman, 1991). To avoid overlap in content with the social conventions used to measure inherent reasoning, each of the questions was asked about properties of animal kinds (two randomly selected from among three: cats, dogs, or birds). As was done with the Inheritance-based Reasoning measure, the questions were averaged to create a single essentialism composite score that could range from 0 to 1, with higher scores indicating greater endorsement of essentialist beliefs.

**Inductive potential.** If children believe that a category-specific, microstructural “essence” causes members of animal kinds to have the properties that they do, then, after learning that a particular member has a property, children may infer that other category members share the same property (e.g., Gelman & Wellman, 1991). To measure this inference, I showed children a picture of an animal and told them a non-visible property of the member (e.g., “Look at this cat. This cat has something called pedicles in its tail.”). I then showed children three different
members of the same kind, with varying visual similarity to the target, and asked whether they thought the new member had the same property (e.g., “Look at this cat. Do you think that this cat has pedicles in its tail like this cat [pointed to the original cat picture]?”; Gelman & Markman, 1987). If they thought the member shared the property, they were given a score of 1, and if not a score of 0. The same was done for two other members of the kind. I then averaged children’s scores across the three novel members (and the two categories) to create a single inductive potential score, which was subsequently combined with the scores for other questions into the essentialism composite.

Stability. Another consequence of having essentialist beliefs about the source of an animal’s properties is the belief that these properties are stable and unchanging. That is, if a particular property is due to the causal essence, and if this causal essence is present in the individual by virtue of its being a member of its kind, then that causal essence should always produce the property as long as the individual is a member of its category (e.g., Gelman & Heyman, 1999). To test this belief, I asked children whether a current property of an animal (e.g., “This cat does something called ‘flehming’”) was present at an earlier age (e.g., “do you think this cat did this thing called ‘flehming’ when it was 4 years old?”), and whether it would always have the same property (e.g., “do you think this cat will always do this thing called

---

18 There are important differences between this measure of stability that is argued to measure a particular aspect of essentialism, and the measure of temporal consistency question used in the inherence-based reasoning scale. In the inherence-based reasoning scale, I always asked about artifact categories maintaining a property over the category’s existence (e.g., have birthday cakes, and will birthday cakes, always have candles). Essentialist beliefs would not support a response to these questions.

19 However, there are exceptions to this form of reasoning that are not seen as evidence against essentialist beliefs. A member of a category can go through significant changes in appearance and still be considered as maintaining their essence, and therefore their category membership (e.g., Keil, Smith, Simons, Levin, 1998; Rosengren, Gelman, Kalish, & McCormick, 1992). For instance, a caterpillar is seen as the same individual when it transforms into a butterfly. Similarly, one can groom a dog to look exactly like a cat, but it would still be considered a dog given its essence.
‘flehming’?”; Gelman & Heyman, 1999; Gelman, Heyman, & Legare, 2007). For each question, children were given a score of 1 if they agreed and 0 if they disagreed, and the average of the two scores across the two categories made up their score for this question.

**Innate potential.** A commonly measured outcome of essentialism is the belief that biological (“essence”-based) causes of properties are more influential than environmental causes (Gelman, Heyman, & Legare, 2007; Gelman & Wellman, 1991; Rhodes et al., 2012; Sousa, Atran, & Medin, 2002). Children in our study were shown two different animals (e.g., a dog and a horse) and were told about a property that differed between the two (e.g., the dog eats bones and the horse eats grass). Then, children were told that the dog had a baby, and right after the baby was born it went to live with the horse. The horse took care of the baby (e.g., “She played with the baby, fed the baby, and loved the baby”), and the baby never saw the birth mother or any other member of that kind again. After asking the children questions to ensure they understood the story, I asked whether they thought the baby (now 6 years old) has the property of the birth mother or the adoptive mother (e.g., “Do you think that the baby eats bones like this dog, or does it eat grass like this horse?”). If children chose the birth mother (an essentialist response), they received a score of 1, and if they chose the adoptive mother, a score of 0.

**Importance of insides for determining category membership.** Two important aspects of essentialism are the belief that the essence is internal, and that it is indicative of category membership. To test these aspects of essentialism, I used a modified version of Gelman and Wellman’s (1991) test of whether the nonvisible insides or the visible properties of an animal were more indicative of its category membership. In the version I used, I told children that two groups of scientists were trying to make a “real live” animal (cat, dog, or bird). They were told
that one team had all the right insides (e.g., “It had dog bones and dog blood, and all of the other things a dog has on the inside”) but none of the right outsides (e.g., “dog skin and dog fur”).

After children heard what the team made, and after they were asked memory questions to ensure understanding, children were asked whether or not the team’s creation was a member of the relevant animal category (e.g., “is what this team made a dog, or not a dog?”). If they said that it was a member of the intended animal kind they were given a score of 1, and if not, a score of 0. The other team made the reverse product (i.e., all of the right outsides, but none of the right insides). I asked children follow-up memory and category identity questions for this team as well, but reversed the scoring so that children received a 0 if they judged the team’s creation with proper outsides but wrong insides to be a member of the relevant category. The scores from the two teams were then averaged.

Results and Discussion

The proposal that inherent reasoning is involved in the development of essentialist beliefs predicts that, around the time when essentialist beliefs tend to first emerge, children’s inherent reasoning would be correlated with their essentialist beliefs. As expected, preschoolers’ inherence-based reasoning (M = 0.61, SD = 0.19) predicted the strength of their essentialist beliefs (M = 0.60, SD = 0.17), \( r(62) = .38, p = .002 \) (see Table 5 for a correlation matrix).

Both inherent reasoning and essentialist beliefs have been shown to change with age (Cimpian & Steinberg, 2014; Gelman, Heyman, & Legare, 2007). Therefore, while the age range for this study was rather restricted (4.0 to 5.0 years of age), it is still possible that the bivariate relationship between inherent reasoning and essentialism is simply a product of independent

---

\(^{20}\) I counterbalanced the order of these response options across children.
changes in these variables with age. For a stronger test of a relationship between inherent reasoning and essentialism, I adjusted for children’s age in a multiple regression analysis. Consistent with my proposal, individual differences in inherence-based reasoning predicted children’s essentialism scores even when age was included in the regression, $b = .34, SE = 0.10, p = .001$. (Age did not significantly predict 4-year-olds’ essentialist beliefs, $b = -.10, SE = 0.06, p = .117$.) This analysis also suggests that other potential confounds that vary strongly with age (e.g., language development, executive functioning) cannot fully account for the relationship of interest. Overall, this study is consistent with the proposal that the inherence heuristic facilitates the development of essentialist beliefs. However, experimental work is necessary to test whether inherent reasoning plays a causal role in the emergence of essentialism.
Chapter 12: Study 8

In Study 8, I tested a stronger prediction of the argument that inherence-based reasoning facilitates the development of essentialist beliefs. Would manipulating the extent to which children rely on inherent explanations in turn affect the extent to which they essentialize? If essentialist beliefs emerge as an elaboration of inherent intuitions, we might expect the answer to be “yes.” To test this prediction, I read children a book intended to either promote or discourage inherence-based reasoning, measuring subsequent changes in the strength of children’s essentialist beliefs. In this study I used a wider age range (4 to 7 years of age, divided into younger and older children), to allow for more variability in children’s essentialist beliefs.

Methods

Participants. Four- to seven-year-old children (N = 112, half boys and half girls; M = 5.94 years, SD = 1.13; range = 4.08 to 7.98) participated in a quiet room at a research lab or in their school. They were socioeconomically diverse, and most were European American.

Procedure. I first assessed the strength of children’s essentialist beliefs using the questions in Study 7 across two animal items. Next, I read them a story book intended to promote or discourage inherence-based reasoning. After reading the book, I assessed children’s essentialist beliefs in the context of two other animals. The animal categories used before and after the book reading, as well as their order within the pre- and post-test, was counterbalanced.

Essentialist beliefs. The questions used to measure children’s essentialist beliefs were similar to those used in Study 7 with a few exceptions. First, I added a new animal trial (about frogs) so as to have a total of four different items (two pre- and two post-book). Second, because reading the book with the children took a relatively long time, I removed the question about the
importance of insides vs. outsides in determining category membership (which was the most involved) from the measure to keep the study at a more reasonable length.

**Inherence-based reasoning manipulation.** Children listened to a story book described as a book that talked about “why some things are the way they are.” The book was titled “Why?” and consisted of 4 “chapters” that each discussed a different social convention (drinking orange juice for breakfast, coins being round, fire trucks being red, and clapping when someone does a good job). For half of the children (Pro-inherence condition), the book provided inherent explanations for the social conventions (e.g., that there are features of the entities involved that best explain why we do things that way). For the other children (Anti-inherence condition), the book provided various facts to promote extrinsic explanations (e.g., “people decided that they wanted it to be that way”; see Appendix A for a sample script).

After each chapter, the children were asked to describe what they heard in the book to a stuffed animal. I used these answers as a comprehension/attention check. Children were considered to have followed along with the book if, for each of the four “chapters” they were able to repeat any of the facts mentioned in that chapter (which consisted of only a few sentences). My prediction of a causal link between inherent reasoning and essentialism is strongest for children who were able to process and retain at least some elements of each chapter’s message. It was for these children that I expected to see the clearest effect of the book manipulation on essentialist beliefs.

Each book (pro-inherence or anti-inherence) had two possible chapter orders. At the end of the book, the children were asked to brainstorm with the experimenter about what a new chapter in the book could be about. This part was intended to encourage further engagement
with the ideas children had just read about. The experimenter made a suggestion and appropriate prompts to keep the child in line with the general idea of the book (see Appendix A for the script).

**Data Analysis.** The data were submitted to a multilevel model with the following predictors: Condition (Pro-inherence vs. Anti-inherence), Age Group (4- and 5-year-olds vs. 6- and 7-year-olds), Phase (Pre-manipulation vs. Post-manipulation), Manipulation Comprehension (those who followed along with the book vs. those who did not), and all of the corresponding interactions. Additionally, I allowed the intercepts for each subject, item, and question type to vary randomly. This analysis helped account for the variability that exists among children’s essentialist beliefs, as well as any variability that might be attributed to particular items and questions used in the procedure. Because children’s essentialism scores were not normally distributed, I used bootstrapping (1,000 replications) to calculate standard errors, p values, and 95% confidence intervals.

**Results and Discussion**

I predicted that the change in children’s essentialism scores from pre-manipulation to post-manipulation would depend on whether they were read the pro-inherence book or the anti-inherence book. However, I also expected the effect of condition on the change in children’s essentialism scores would be strongest for those children who could repeat elements of the chapters they had been read. That is, I expected a stronger two-way Condition (Pro- vs. Anti-Inherence) × Phase (Pre- vs. Post-manipulation) interaction for children who understood the books than for those who did not, which would be reflected in a three-way Condition × Phase × Comprehension interaction.
The two-way Condition × Phase interaction (including all participants) was not significant, $b = .04 [-.03, .11], SE = .04, p = .315$. However, as expected, the three-way interaction with Comprehension was marginally significant, $b = .15 [-.01, .30], SE = .08, p = .056$ (see Tables 6 for the complete regression table and Table 7 for the means). To examine the source of this three-way interaction, I split the data into two groups: those children who passed the comprehension questions ($n = 41$) and those who did not ($n = 71$). These follow-up analyses revealed that children who followed along with the book showed pre- to post-manipulation changes in their essentialist beliefs that were highly dependent on condition. In other words, they showed a significant Condition × Phase interaction, $b = .13 [.02, .25], SE = .06, p = .014$, with a significant increase in essentialism in the Pro-inherence condition ($M_{pre} = .48$ vs. $M_{post} = .57, b = .10 [.03, .17], SE = .04, p = .002$) and a slight decrease in the Anti-inherence condition ($M_{pre} = .59$ vs. $M_{post} = .57, b = -.04 [-.12, .05], SE = .05, p = .47$). However, children who did not follow along with the book did not show this Condition × Phase interaction, $b = -.02 [-.10, .06], SE = .04, p = .616$ (see Table 7 for the means). There was no effect of Age, nor did any interactions with Age reach significance (see Table 6 for the complete regression table).

The findings in this study are consistent with the proposal that inherent reasoning lays an important foundation for children's essentialist beliefs. The children who understood the manipulation books displayed changes in the strength of their essentialist beliefs, and these changes were dependent on their having been read pro-inherence or anti-inherence books about social conventions. This finding is striking given that the content of the books and the essentialism measure were quite distinct (one being about why firetrucks are red, for example, and the other about whether a baby born from a cat and raised by a pig would have a straight or
curly tail).

There were also, however, a few aspects of the data were not predicted a priori. While these findings do not in and of themselves contradict my proposal, they are worth further discussion. First, many children did not follow along with the books used in the manipulation. In hindsight, the scripts in each book chapter may not have connected their arguments well enough to the overall “why?”-question associated with the chapter (e.g., why are coins round?), in turn hindering children’s ability to follow along. In other words, although the facts provided on each page were relatively easy to understand (e.g., coins having faces on them, and faces are round), children may have had some difficulty understanding how this information was relevant for answering the corresponding “why?” question. Along similar lines, the chapters were structured in a way such that the children heard three distinct arguments presented on successive pages. Perhaps it was difficult for the children to follow multiple arguments presented in such quick succession (see Appendix C for a sample script). Second, the effect found was driven by children in the Pro-inherence condition, as children in the Anti-inherence condition did not have lower essentialism scores in the post-test. To speculate, it might be the case that supporting children’s inherent intuitions is feasible through a subtle manipulation while getting children to go against their intuitive explanations, on the other hand, may be more difficult to accomplish, especially with just one book reading.

The above issues with the current manipulation may be addressed with a few changes to the manipulation books used. Perhaps if the chapters presented the arguments in a narrative structure, and the books had an introduction and conclusion to help children tie the chapters to the more general inherent and non-inherent ways of reasoning, more children would follow
along. These changes may also help children in the Anti-inherence condition to better understand and appreciate the non-inherent reasoning used.

One other aspect of the data that was surprising was the fact that the children in the Pro-inherence condition who followed along with the books had particularly low pre-test essentialism scores (see Table 7). To determine the reason for these low scores, I looked for differences on a variety of factors between this subsample (i.e., Pro-inherence children who followed along with the books) and the other groups of children. This sample matched the children who were in the same condition but did not follow along, as well as children in the Anti-inherence condition who followed along, on the following variables: the items they received in the pre-test, the question order and chapter order they were assigned to, as well as the locations they were tested (lab vs. school). The only difference found between the subsamples was in age—not surprisingly, the children who followed along were somewhat older ($M = 6.68$ years for those who followed vs. $M = 5.51$ years for those who did not). Critically, however, there was no age difference between the children who followed along in the Pro- and Anti-inherence conditions ($M_{Pro} = 6.65$ years vs. $M_{Anti} = 6.72$ years), and thus the reason for the lower pre-test scores in the Pro-inherence condition is still unclear.

Without a clear reason for distribution of pre-test essentialism scores, it was particularly important to conduct a replication of this study. Finding an effect of the Pro- vs. Anti-inherence manipulation in an independent sample of children would minimize the possibility that the present study found the predicted results simply due to a failure of random assignment.
Chapter 13: Study 9

Study 9 was a close replication of Study 8, and thus a further test of the predicted causal relationship between inherent-based reasoning and essentialist beliefs.

Methods

Participants. Four- to seven-year-old children (N = 112, half boys and half girls; M = 5.95 years, SD = 1.23; range = 4.0 to 8.1) participated in a quiet room at a research lab or in their school. They were socioeconomically diverse, and most were European American.

Procedure. The procedure was the same as Study 8 with three changes. First, I changed the books to (a) have more of a narrative structure to facilitate children’s attention and interest, (b) include an introduction and conclusion to the book that discussed generally “why some things are the way they are” (see Appendix D), (c) omit the scaffolded brainstorming at the end, to keep the sessions from taking too long and overtaxing children’s attention. Second, again in the interest of keeping the testing session at a reasonable length, I only used one animal item to measure children’s essentialist beliefs before the book reading. I also removed the bird item from the set of counterbalanced items. Third, I changed the innate potential questions so that, for all items, children were asked about behavioral properties (rather than having some about physical properties, as these did not show much variability across children in the previous study).

Results and Discussion

The findings replicated those of Study 8. Most importantly, I again found a significant three-way Condition × Phase × Comprehension interaction, \( b = .20 \ [0.04, .38], SE = .09, p = .021 \). As before, children who followed along with the book (\( n = 44 \)) showed a significant two-way Condition × Phase interaction, \( b = .18 \ [0.04, .30], SE = .07, p = .008 \), that was due to a significant
increase in children’s essentialism in the Pro-inherence condition ($M_{pre} = .43$ vs. $M_{post} = .58, b = .20 [.11, .28], SE = .04, p < .001$) but not in the Anti-inherence condition ($M_{pre} = .55$ vs. $M_{pre} = .57, b = .01 [-.08, 10], SE = .05, p = .828$). In contrast, children who did not follow along with the book ($n = 68$) did not show a significant two-way Condition × Phase interaction, $b = -.02 [-.11, .08], SE = .05, p = .682$. Once again, there was no effect of Age, nor did any of the interactions with Age reach significance (see Table 8 for means and Table 9 for the complete regression table).

Replicating Study 8, and consistent with my proposal, (some) children’s essentialist beliefs were influenced by reading a book that either encouraged or discouraged inherent reasoning. This finding further suggests that inherent explanations lay an important part of the foundation on which essentialist beliefs develop. Once again, however, many children did not follow along with the manipulation books, and the effect found was driven by children in the Pro-inherence group who did follow along. I discuss possible reason for these aspects of the data as well as future work that might test these explanations in the next section.
Chapter 14: Discussion of Part 2

The findings from three studies (Studies 7-9) suggest a link between inherence-based reasoning and the development of children’s essentialist beliefs. At the age when essentialism begins to emerge (4 years), children’s inherent reasoning predicted the strength of their essentialist beliefs (Study 7). This relationship was observed despite the fact that inherent reasoning was only measured in domains that fall outside of the scope of essentialism (i.e., the inherent reasoning measure did not ask about natural kinds or social groups). Furthermore, the predictive relationship between inherent reasoning and essentialist beliefs was not accounted for by children’s age. In Studies 8 and 9, the strength of children’s essentialist beliefs was influenced by reading a book that explained various social conventions (e.g., why fire trucks are red). Half the children were read a version of the book that gave inherent explanations; the others were read a book that explained the conventions in terms of people’s influence on them. After simply being read arguments that support (or contradict) inherent reasoning, some of the children subsequently showed changes in the strength of their essentialist beliefs. These results provide a first test of the role that inherent reasoning plays in the emergence of essentialist beliefs, and they support the broader view that essentialism originates from domain-general cognitive processes.

As future studies build on the current work, it would be valuable to improve the effectiveness of the current inherent reasoning manipulation (i.e., the pro-inherence and anti-inherence books). The manipulation was rather subtle—the books used to influence the children’s reasoning only explained four example phenomena with the intention of influencing children’s general explanatory reasoning. In addition, the children were only read the book once.
Two changes to the manipulation might improve its effectiveness in changing children’s reasoning.

First, it is possible that the exposure to the books in the current study was simply not sufficient to induce a substantial change in children’s reliance on inherent intuitions. Perhaps if children had been read the books multiple times, even over multiple days (as was done in Gelman, Ware, & Kleinberg, 2010, and Rhodes et al., 2012), more children would have been affected by the book. This change might also enable more children to follow along with the books by providing more time to connect the specific explanations on each page to the more general way of reasoning that underlies each example. I predict that such a change would strengthen the effect of the pro- versus anti-inherence manipulation on children’s essentialism.

Second, in addition to increasing the number of readings of the books, it would be valuable for the specific arguments in each book to more directly align with the more general aim of the manipulation (i.e., enhancing or reducing general reliance on inherent intuitions). The current books’ focus on four concrete phenomena may have been too distant to change children’s more general reasoning (even with the introduction and conclusion used in Study 9). Perhaps including more abstract general statements that give voice to an inherent or extrinsic mode of explanation (e.g., “there are a lot of things that make perfect sense because of the way they look or work” or “a lot of the time, people just make decisions about how to do something even though it could be done a lot of other ways too”) might have a stronger influence on children’s tendency to rely on their inherent intuitions more generally. Furthermore, to more directly test the efficacy of the manipulation, two additions to the procedure would be valuable. It would be useful to measure (1) whether children believed the arguments in the books, and (2)
to what extent children’s inherent reasoning was actually lowered by the Pro- versus Anti-inherence manipulation (i.e., a manipulation check). If we assessed the changes in children’s inherent thinking in response to the manipulation, we could also test whether the effect of the manipulation on children’s essentialist beliefs was in fact mediated by changes in children’s inherent reasoning. However, adding such measures may not be practically feasible given that the experimental sessions in Studies 8 and 9 already seemed to have pushed young children’s attention spans to their limit.

Nevertheless, even with the subtle manipulation of inherent reasoning used in Studies 8 and 9, the findings from these studies are consistent with the idea that inherent reasoning forms part of the foundation for the initial development of essentialist beliefs. A question that remains, however, is the following: How exactly does inherent reasoning foster essentialist beliefs? One possibility is that children’s inherent explanations combine with other beliefs and cognitive abilities over time. Together, this aggregate set of general beliefs and cognitive abilities might make more easily available the notion of a specific type of inherent feature for natural and social kinds (i.e., an essence). For instance, perhaps inherent reasoning interacts with children’s noticing that natural kind members are often treated as equivalent or interchangeable. Children are often taught about general categories of animals (e.g., dogs) in the presences of a single member, as if the member is a mere representative for its category, and category members are alike. The realization that kind members are often treated as interchangeable might combine with reasoning that something about the category member explains the way it is (i.e., inherent reasoning) to give rise to the belief that inherent features are shared by category members. This belief might then take children one step (of many) closer to a more fully formed essentialist
belief by supporting the notion that a particularly important inherent feature might be shared by category members. Over time, children’s reasoning may develop into stand-alone essentialist beliefs that are invoked independently of inherent explanations.

It is noteworthy that the tendency to explain inherently is domain general. Thus, the suggestion that it plays a role in setting up children’s initial essentialist beliefs is consistent with the broader view that essentialist beliefs originate from domain general cognitive processes. However, there remain gaps in our understanding of the complete foundation on which, and mechanism by which, children initially develop their essentialist beliefs. It isn’t clear what other general cognitive processes and beliefs facilitate the last step children take to reach fully developed essentialist beliefs. To investigate this, it is likely most informative to use a longitudinal design in which changes in domain-general knowledge and cognitive processes (e.g., knowledge of the insides of natural kinds, the belief in the coherence of natural kinds, beliefs about internal energy, and inherent reasoning) can be used to predict the initial development of essentialist beliefs. A longitudinal design would reveal how various cognitive factors and their interactions support the emergence of essentialist beliefs.

A better understanding of the origins of essentialist beliefs can inform work on beliefs and behaviors that follow from essentialism, and thus also inform the mechanisms that underlie these beliefs and behaviors. Perhaps the downstream effects that are most critical to understand are those that follow from essentialist beliefs about social groups—specifically, stereotyping and prejudice. The current work suggests interventions that target the information that is highly accessible in memory. For instance, increasing the accessibility of extrinsic, environmental factors might decrease the likelihood that essentialist explanations are generated. More
specifically, by highlighting the variety of environmental influences on members of social kinds (e.g., culture-specific historical events), such non-inherent information might become much more accessible in memory and help people to avoid the typical inherent, and more specifically essentialist, reasoning. Along similar lines, making extrinsic information accessible early in life might significantly reduce the likelihood that essentialist beliefs, about at least some social groups, would develop at all.

Overall, the current work advances our understanding the origins of essentialism. It supports the argument that essentialism develops out of domain-general reasoning processes and brings us a step closer to understanding how essentialist beliefs might develop from such domain-general processes.
Chapter 15: Conclusion

The general tendency to explain phenomena with inherent features of the explananda is at the root of many common beliefs about how the world works. Across 9 studies, I provided evidence for the influence of inherent reasoning on people’s beliefs about the non-arbitrariness of words and their essentialist beliefs. Part 1 included 6 studies that suggested inherent explanations may underlie nominal fit beliefs in children and, to some extent, even in adults. Three additional studies in Part 2 provided initial steps towards understanding the role of inherent reasoning in the development of essentialist beliefs.

These findings add to the body of evidence highlighting the important ways in which inherent explanations shape our reasoning (Cimpian & Salomon, 2014a, 2014b; Hussak & Cimpian, 2015; Tworek & Cimpian, in press). For example, in prior work the tendency to generate inherent explanations was also found to give rise to beliefs that legitimize the sociopolitical status quo (e.g., rich people have more money because they’re smarter; Hussak & Cimpian, 2015). The inherence bias in explanations might also contribute to the tendency to assume that the way things are (e.g., people typically give roses for Valentine’s Day) is the way they ought to be (e.g., people should give roses for Valentine’s Day; Tworek & Cimpian, in press). If people reason that inherent features of entities in an observed pattern of behavior explain why it is that way (e.g., roses are beautiful), they may also reason that the pattern is good and right—that is, they may imbue the pattern with sociomoral value. To illustrate, in the past, Americans might have reasoned that racial segregation in public places could be explained using the inherent features of those in each racial group, and if this way of doing things is thought to stem from stable features, it might seem natural and the way things should be.
Thus, the inherence heuristic seems to lay a foundation for not only nominal fit and essentialist beliefs, but also other psychological processes such as system justification (or defense of the status quo), and sociomoral evaluation. The fact that these judgments are influenced by the tendency to rely on highly accessible information when generating explanations suggests the value in investigating how abstract, higher-order judgments and beliefs are shaped by very basic cognitive processes. By adopting this perspective, new light may also be shed on common mechanisms that underlie seemingly disparate beliefs.
References


Hickling, A. K. & Wellman, H. M. (2001). The emergence of children’s causal explanations and


Table 1

Sample Items from the Nominal Fit Measure Used in Study 2

Sample question #1:
Consider long ago when people discovered the kind of animal above [referring to a picture of a lion] and decided to give it a name. In English, we call it a “lion.” Do you think there is something particularly appropriate about this name, or could we have just as easily called this animal something else?

Answer scale: 1 (this name is particularly appropriate) to 7 (we could have easily called this animal something else)

Sample question #2:
Think back to a time long ago when people discovered the kind of animal above [referring to a picture of a pig] and decided to give it a name. In selecting this name, how many suitable options did they have?

Answer scale: 1 (only a few suitable options) to 7 (countless suitable options)

Sample question #3:
The name “yaroo” would have been a suitable name for this kind of animal [referring to a picture of a turtle].

Answer scale: 1 (strongly disagree) to 7 (strongly agree)

Sample question #4:
Imagine the word “cow” was not an English word. That is, imagine we called cows something else in English and the word “cow” didn’t mean anything. Now imagine that at some point in the future English speakers decided to use the word “cow” for a kind of thing that didn’t yet have a name. In selecting a kind of thing to call “cow,” how many suitable options did they have?

Answer scale: 1 (only a few unnamed things would be suitable) to 7 (pretty much any unnamed thing would be suitable)

Note. Studies 3–5 used identical questions, except about artifact names (e.g., “bottle”).
Table 2

Sample Justifications from the Nominal Fit Measure (Studies 2–5)

... the long “a” in “giraffe” appears fitting for the long neck.

Only a few different words could possibly describe a glove.

“Yaroo” is two syllables and doesn’t really fit.

I feel like the name fits the object better than anything else would.

I feel like turtle was given its name for biological reasons and “yaroo” seems just arbitrary.

I think it’s a fitting name [“giraffe”] although other names could be used, but I believe the way the word is pronounced it gives you a feeling of being stretched out like the giraffe’s neck.

The [novel] name doesn’t fit the form and function of a chair, or anything you sit on for that matter.

I think it might be difficult [to call a bowl “a fork”] because the word “bowl” kind of describes the “o” shape of the object. The fork doesn’t.

It is called a “fork” because of the prongs on it, so [using the word “fork” for] a bowl would take away from its real name.

“Yaroo” doesn’t give a sound like “chair,” to imply what you do with it.

It [the novel word] doesn’t really match the object.

... “bowl” does have some connotations due to how the mouth moves to make the sound.

“Yaroo” doesn’t seem to fit the name for a chair. It just sounds wrong to me.

The name sounds fitting for what it’s meant to do.

Note. Quotes were added to participants’ justifications where needed to clarify when they were referring to the word or to the object.
Table 3

Regression Analyses Predicting Participants’ Nominal Fit Beliefs in Studies 2–4

<table>
<thead>
<tr>
<th>Study</th>
<th>Predictor</th>
<th>b</th>
<th>SE</th>
<th>BCa 95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Inherence Heuristic</td>
<td>.24</td>
<td>.07</td>
<td>[.09, .39]</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>Need for Closure</td>
<td>.07</td>
<td>.09</td>
<td>[−.12, .24]</td>
<td>.452</td>
</tr>
<tr>
<td></td>
<td>Need for Cognition</td>
<td>.06</td>
<td>.06</td>
<td>[−.09, .18]</td>
<td>.352</td>
</tr>
<tr>
<td></td>
<td>$R^2$ total</td>
<td>.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>4.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>126</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Inherence Heuristic</td>
<td>.23</td>
<td>.08</td>
<td>[.06, .41]</td>
<td>.004</td>
</tr>
<tr>
<td></td>
<td>Raven’s Progressive Matrices</td>
<td>−.08</td>
<td>.04</td>
<td>[−.16, −.01]</td>
<td>.022</td>
</tr>
<tr>
<td></td>
<td>Creative Personality Checklist</td>
<td>−.01</td>
<td>.02</td>
<td>[−.05, .03]</td>
<td>.805</td>
</tr>
<tr>
<td></td>
<td>$R^2$ total</td>
<td>.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>7.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>122</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Inherence Heuristic</td>
<td>.27</td>
<td>.09</td>
<td>[.09, .45]</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td>Counterfactual Thinking</td>
<td>−.10</td>
<td>.04</td>
<td>[−.18, −.02]</td>
<td>.029</td>
</tr>
<tr>
<td></td>
<td>Multilingualism</td>
<td>.01</td>
<td>.04</td>
<td>[−.06, .13]</td>
<td>.780</td>
</tr>
<tr>
<td></td>
<td>$R^2$ total</td>
<td>.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>7.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>122</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. BCa = bias-corrected and accelerated.
Table 4

The Inherence–Global and Inherence–Language Scales

The Inherence–Global Scale

1) There is some feature of orange juice (maybe something about how it tastes or something else) that explains why we drink orange juice at breakfast.
2) There is some feature of the color red (maybe something about how it looks or something else) that explains why red in a traffic light means “stop.”
3) Girls generally wear pink because of something about the color pink—for example, it looks delicate and flower-like.
4) Engagement rings typically have diamonds because of something about diamonds—for example, it might be something about how rare they are.
5) The reason we don’t keep chipmunks as pets has to do with the way chipmunks act, or maybe something else about them.
6) The reason wedding dresses are typically white has to do with the way white looks, or maybe something else about it.
7) There is some feature of funerals that explains why we associate the color black with them.
8) There is some feature of mint that explains why we use it to flavor toothpaste.

The Inherence–Language Scale

1) We use the word “frog” when talking about the amphibian because of something about the word—perhaps something about how it sounds or looks in print.
2) We call clocks “clocks” because of something about the object—perhaps something about how it works or looks.
3) The fact that we call giraffes “giraffes” can be explained by some feature of the animal—maybe something about how it looks or something else.
4) The fact that we use the word “pizza” to talk about the food can be explained by some feature of the word—maybe something about how it’s pronounced or something else.
5) There is some feature of the word “cat” (maybe something about how it sounds or something else) that explains why it is used to talk about cats.
6) There is some feature of ovens (maybe something about how they work or something else) that explains why the word “oven” is used to talk about them.
7) The reason we call horses “horses” has to do with something about the animal—it could be something about how it looks or something else about it.
8) The reason we use the word “milk” when talking about the white liquid has to do with something about the word—it could be something about how it’s pronounced or something else about it.

Note. The Inherence–Global Scale also included the following two catch items: “There is something about your favorite sports team that explains why coffee keeps us awake” (false) and “The brightness of the day is usually due to the sun’s light, regardless of the season” (true). The Inherence–Language Scale also included two attention checks that directed subjects to slide the response bar to either end of the scale.
Table 5

*Bivariate Correlation Matrix from Study 7*

<table>
<thead>
<tr>
<th></th>
<th>Inherence-based reasoning</th>
<th>Essentialist beliefs</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inherence-based reasoning</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Essentialist beliefs</td>
<td>.382**</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Age</td>
<td>.058</td>
<td>-.162</td>
<td>---</td>
</tr>
</tbody>
</table>

**p < .01**
### Table 6

*Fixed Effects from the Main Mixed Multilevel Regression in Study 8*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>b</th>
<th>Bootstrap SE</th>
<th>Bootstrap Percentile 95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>-.004</td>
<td>.03</td>
<td>-.06, .05</td>
<td>.876</td>
</tr>
<tr>
<td>Age</td>
<td>.02</td>
<td>.03</td>
<td>-.03, .08</td>
<td>.406</td>
</tr>
<tr>
<td>Phase</td>
<td>.01</td>
<td>.02</td>
<td>-.02, .05</td>
<td>.507</td>
</tr>
<tr>
<td>Comprehension</td>
<td>-.06</td>
<td>.03</td>
<td>-.13, -.002</td>
<td>.050</td>
</tr>
<tr>
<td>Condition × Age</td>
<td>-.01</td>
<td>.05</td>
<td>-.12, .09</td>
<td>.816</td>
</tr>
<tr>
<td>Condition × Phase</td>
<td>.04</td>
<td>.04</td>
<td>-.03, .11</td>
<td>.315</td>
</tr>
<tr>
<td>Condition × Comprehension</td>
<td>-.13</td>
<td>.06</td>
<td>-.26, .01</td>
<td>.056</td>
</tr>
<tr>
<td>Age × Phase</td>
<td>.01</td>
<td>.03</td>
<td>-.06, .07</td>
<td>.752</td>
</tr>
<tr>
<td>Age × Comprehension</td>
<td>.01</td>
<td>.06</td>
<td>-.11, .14</td>
<td>.826</td>
</tr>
<tr>
<td>Phase × Comprehension</td>
<td>.03</td>
<td>.04</td>
<td>-.04, .10</td>
<td>.491</td>
</tr>
<tr>
<td>Condition × Age × Phase</td>
<td>-.08</td>
<td>.07</td>
<td>-.21, .05</td>
<td>.232</td>
</tr>
<tr>
<td>Condition × Age × Comprehension</td>
<td>.13</td>
<td>.12</td>
<td>-.11, .38</td>
<td>.226</td>
</tr>
<tr>
<td>Condition × Phase × Comprehension</td>
<td>.15</td>
<td>.08</td>
<td>-.01, .30</td>
<td>.056</td>
</tr>
<tr>
<td>Age × Phase × Comprehension</td>
<td>.03</td>
<td>.07</td>
<td>-.11, .18</td>
<td>.557</td>
</tr>
<tr>
<td>Condition × Age × Phase × Comprehension</td>
<td>-.04</td>
<td>.15</td>
<td>-.33, .27</td>
<td>.801</td>
</tr>
</tbody>
</table>
Table 7

*Mean Composite Essentialism Scores at the Subject Level from Study 8 with SDs in Parentheses*

<table>
<thead>
<tr>
<th></th>
<th>All Participants</th>
<th>Followed the book reading</th>
<th>Did not follow the book reading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><em>N = 41</em></td>
<td><em>N = 71</em></td>
</tr>
<tr>
<td><strong>Pro-Inherence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-manipulation</td>
<td>.56</td>
<td>.48</td>
<td>.63</td>
</tr>
<tr>
<td></td>
<td>(.19)</td>
<td>(.18)</td>
<td>(.17)</td>
</tr>
<tr>
<td>Post-manipulation</td>
<td>.61</td>
<td>.57</td>
<td>.63</td>
</tr>
<tr>
<td></td>
<td>(.18)</td>
<td>(.21)</td>
<td>(.15)</td>
</tr>
<tr>
<td><strong>Anti-Inherence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-manipulation</td>
<td>.58</td>
<td>.59</td>
<td>.57</td>
</tr>
<tr>
<td></td>
<td>(.18)</td>
<td>(.17)</td>
<td>(.19)</td>
</tr>
<tr>
<td>Post-manipulation</td>
<td>.58</td>
<td>.57</td>
<td>.59</td>
</tr>
<tr>
<td></td>
<td>(.18)</td>
<td>(.16)</td>
<td>(.20)</td>
</tr>
</tbody>
</table>
### Table 8

*Fixed Effects from the Main Mixed Multilevel Regression in Study 9*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$b$</th>
<th>Bootstrap SE</th>
<th>Bootstrap Percentile 95% CI</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>.004</td>
<td>.03</td>
<td>-.05, .07</td>
<td>.835</td>
</tr>
<tr>
<td>Age</td>
<td>.01</td>
<td>.03</td>
<td>-.05, .07</td>
<td>.842</td>
</tr>
<tr>
<td>Phase</td>
<td>.07</td>
<td>.02</td>
<td>.03, .11</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Comprehension</td>
<td>-.09</td>
<td>.03</td>
<td>-.15, -.02</td>
<td>.013</td>
</tr>
<tr>
<td>Condition × Age</td>
<td>-.03</td>
<td>.06</td>
<td>-.15, .10</td>
<td>.698</td>
</tr>
<tr>
<td>Condition × Phase</td>
<td>.05</td>
<td>.04</td>
<td>-.03, .14</td>
<td>.205</td>
</tr>
<tr>
<td>Condition × Comprehension</td>
<td>-.10</td>
<td>.06</td>
<td>-.21, .03</td>
<td>.131</td>
</tr>
<tr>
<td>Age × Phase</td>
<td>.04</td>
<td>.04</td>
<td>-.04, .12</td>
<td>.395</td>
</tr>
<tr>
<td>Age × Comprehension</td>
<td>-.02</td>
<td>.06</td>
<td>-.15, .11</td>
<td>.743</td>
</tr>
<tr>
<td>Phase × Comprehension</td>
<td>.05</td>
<td>.04</td>
<td>-.03, .14</td>
<td>.228</td>
</tr>
<tr>
<td>Condition × Age × Phase</td>
<td>-.13</td>
<td>.08</td>
<td>-.28, .03</td>
<td>.118</td>
</tr>
<tr>
<td>Condition × Age × Comprehension</td>
<td>.33</td>
<td>.25</td>
<td>-.16, .87</td>
<td>.167</td>
</tr>
<tr>
<td>Condition × Phase × Comprehension</td>
<td>.20</td>
<td>.09</td>
<td>.04, .38</td>
<td>.021</td>
</tr>
<tr>
<td>Age × Phase × Comprehension</td>
<td>-.11</td>
<td>.09</td>
<td>-.31, .06</td>
<td>.209</td>
</tr>
<tr>
<td>Condition × Age × Phase × Comprehension</td>
<td>-.10</td>
<td>.17</td>
<td>-.43, .23</td>
<td>.565</td>
</tr>
</tbody>
</table>
Table 9

*Mean Composite Essentialism Scores at the Subject Level from Study 9 with SDs in Parentheses*

<table>
<thead>
<tr>
<th></th>
<th>All Participants</th>
<th>Followed the book reading $N = 44$</th>
<th>Did not follow the book reading $N = 68$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pro-Inherence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-manipulation</td>
<td>.58 (.23)</td>
<td>.43 (.23)</td>
<td>.65 (.20)</td>
</tr>
<tr>
<td>Post-manipulation</td>
<td>.64 (.18)</td>
<td>.58 (.14)</td>
<td>.67 (.18)</td>
</tr>
<tr>
<td><strong>Anti-Inherence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-manipulation</td>
<td>.56 (.24)</td>
<td>.55 (.23)</td>
<td>.57 (.26)</td>
</tr>
<tr>
<td>Post-manipulation</td>
<td>.60 (.21)</td>
<td>.57 (.17)</td>
<td>.62 (.23)</td>
</tr>
</tbody>
</table>
### Appendix A

#### Table A1

*Correlation Matrices of the Measures in Studies 1–5*

<table>
<thead>
<tr>
<th>Study 1</th>
<th>Nominal Fit</th>
<th>Inherence Heuristic</th>
<th>Age</th>
<th>Multilang. Exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal Fit</td>
<td>—</td>
<td>—</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>Inherence Heuristic</td>
<td>.53***</td>
<td>—</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>Age</td>
<td>—.43***</td>
<td>—.17</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Multilanguage Exposure</td>
<td>-.09</td>
<td>.12</td>
<td>-.11</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study 2</th>
<th>Nominal Fit</th>
<th>Inherence Heuristic</th>
<th>Need for Closure</th>
<th>Need for Cognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal Fit</td>
<td>—</td>
<td>—</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>Inherence Heuristic</td>
<td>.31***</td>
<td>—</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>Need for Closure</td>
<td>.13</td>
<td>.28**</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>Need for Cognition</td>
<td>-.01</td>
<td>-.28**</td>
<td>-.16</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study 3</th>
<th>Nominal Fit</th>
<th>Inherence Heuristic</th>
<th>Raven’s SPM</th>
<th>Creative Personality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal Fit</td>
<td>—</td>
<td>—</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>Inherence Heuristic</td>
<td>.33***</td>
<td>—</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>Raven’s SPM</td>
<td>-.30**</td>
<td>-.28**</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Creative Personality</td>
<td>-.06</td>
<td>-.03</td>
<td>.15</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study 4</th>
<th>Nominal Fit</th>
<th>Inherence Heuristic</th>
<th>Counterfactual Thinking</th>
<th>Multilingualism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal Fit</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Inherence Heuristic</td>
<td>.34***</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Counterfactual Thinking</td>
<td>-.25**</td>
<td>-.17</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Multilingualism</td>
<td>-.06</td>
<td>-.25**</td>
<td>.04</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study 5</th>
<th>Nominal Fit</th>
<th>Inherence-Language</th>
<th>Inherence–Global</th>
<th>CRT</th>
<th>Chance Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal Fit</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Inherence-Language</td>
<td>.36***</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Inherence–Global</td>
<td>.11</td>
<td>.25**</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>CRT</td>
<td>-.29***</td>
<td>-.28**</td>
<td>-.13</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Chance Scale</td>
<td>.14</td>
<td>-.01</td>
<td>-.01</td>
<td>-.07</td>
<td>—</td>
</tr>
</tbody>
</table>

* * p < .05. ** p < .01. *** p < .001.
Appendix B

Table B1

*Means of the Measures in Studies 1–5*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Study 1 (children)</th>
<th>Study 2 (adults)</th>
<th>Study 3 (adults)</th>
<th>Study 4 (adults)</th>
<th>Study 5 (adults)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nominal Fit</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(possible range [children] = 0 to 1)</td>
<td>0.55 (0.42)</td>
<td>2.14 (0.85)</td>
<td>2.09 (0.95)</td>
<td>2.41 (1.12)</td>
<td>2.31 (1.06)</td>
</tr>
<tr>
<td>(possible range [adults] = 1 to 7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Inherence Heuristic</strong></td>
<td>0.00&lt;sup&gt;a&lt;/sup&gt; (0.79)</td>
<td>5.89 (1.10)</td>
<td>5.85 (1.11)</td>
<td>5.95 (1.28)</td>
<td>—&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>(possible range [adults] = 1 to 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>5.89 (1.15)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>(range = 4.18 to 7.96 years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Multilanguage Exposure</strong></td>
<td>0.19 (0.43)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>(range = 0 to 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Need for Closure</strong></td>
<td>—</td>
<td>5.33 (0.72)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>(possible range = 1 to 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Need for Cognition</strong></td>
<td>—</td>
<td>5.79 (1.25)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>(possible range = 1 to 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Raven’s SPM</strong></td>
<td>—</td>
<td>—</td>
<td>4.47 (2.45)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>(possible range = 0 to 12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Creative Personality</strong></td>
<td>—</td>
<td>—</td>
<td>4.97 (3.99)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>(possible range = −12 to 18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Counterfactual Thinking</strong></td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>5.33 (2.28)</td>
<td>—</td>
</tr>
<tr>
<td>(possible range = 1 to 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Multilingualism</strong></td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>2.07 (2.70)</td>
<td>—</td>
</tr>
<tr>
<td>(range = 0 to 17)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Inherence–Language</strong></td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>45.47 (24.95)</td>
</tr>
<tr>
<td>(possible range = 0 to 100)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Inherence–Global</strong></td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>6.76 (1.83)</td>
</tr>
<tr>
<td>(possible range = 0 to 10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CRT</strong></td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>1.66</td>
</tr>
<tr>
<td>(possible range = 0 to 3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Chance Scale</strong></td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>13.15 (5.97)</td>
</tr>
<tr>
<td>(possible range = 0 to 30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Children’s scores were standardized (hence the zero average).

<sup>b</sup> Please see the Inherence–Language and Inherence–Global measures further down in the table.
Appendix C

Sample scripts from the Coins chapter of the books used in Study 8

This book is called “Why?” It talks about why some things are the way they are! Let’s talk about coins. You know how coins are round, right?

**Pro-Inherence book**

Well you know, it seems right that coins are round. Look, they fit into candy machines just right. If they were a different shape, they wouldn’t fit in so well.

Also, it works really well that coins are round, because then, when you reach into your pocket to get some coins, they won’t poke you. If coins were a different shape, they might be pointy and poke you.

And, you know how there are faces on coins? Well, because faces are round, and we put faces on coins, it seems like a good reason to make coins round. It makes perfect sense that coins are round.

So, can you tell Feppy why coins are round? [*if child does not answer:*] Let’s think about it together. What did we just read about that we could tell Feppy? [*if child still does not answer:*] Well maybe we could tell Feppy that coins are round because then they fit into candy machines.

**Anti-Inherence book**

Well you know, a long time ago in all sorts of places, like Spain, coins were not like coins are today. They were not round, they were all different shapes. Look here [points to page]. They used these as their money.

Today we have round coins, but they don’t have to be that way. They are round today just because people decided [*emphasized included in reading*] that they wanted it to be that way. But as long as you can use them to buy things, it doesn’t matter what coins look like.

In the future, people might decide to make coins square or triangle shaped, or maybe even star shaped. People get to decide.

So, can you tell Feppy why coins are round? [*if child does not answer:*] Let’s think about it together. What did we just read about that we could tell Feppy? [*if child still does not answer:*] Well maybe we could tell Feppy that coins are round because it is what people decided.
Sample scripts from the Introduction, Coins chapter, and Conclusion of the books used in Study 9

This book is called “Why?” It talks about why some things are the way they are!
So you know how we clap when someone does a good job? And you know how fire trucks are red, and coins are round? And you know how we drink orange juice at breakfast? Something like these are the way they are because...

<table>
<thead>
<tr>
<th>Pro-Inherence book</th>
<th>Anti-Inherence book</th>
</tr>
</thead>
<tbody>
<tr>
<td>...they make perfect sense. Some things just go together. Maybe they go together because of the way the things look, or maybe they do together because of the way the things work. So, today, some things are the way they are because it makes perfect sense for them to be that way. They really work best that way. Let’s talk about coins. You know how coins are round, right? A long time ago people just used coins as their money. Lots of people kept the coins in their pockets or in small bags. When they reached into their pockets or bags to get them out, the round shape of the coin worked really well. The coins were round and so it was smooth in people’s hands. And the round shape made the coins really easy to hold. So, just like fire trucks are red because it makes perfect sense for it to be that way, that’s why coins are round, because it makes perfect sense for them to be that way. There is an important reason that they are that way. Can you tell Feppy why coins are round? What did we read about? [if child mentions anything from the chapter beyond coins being round:] Yeah! That’s right! [if the child does not mention something from the chapter beyond coins being round:] Maybe we can tell him it worked really well for them to be like that. [Conclusion] We talked about why some things are the way they are. Something are the way they are because it makes perfect sense for them to be that way. They just go together because of the way they things look, or because of the way the things work. So, clapping after people do a good job, fire trucks being red, coins being round, and drinking orange juice for breakfast are all just because it makes perfect sense for them to be that way. And probably lots of other things today are the way they are because it just makes perfect sense for them to be that way.</td>
<td>...of people a long time ago. Many people many years ago just decided to make them that way, or maybe something happened that made them that way. So, today, some things are the way they are just because of people many years ago. And people could have easily made them another way. Let’s talk about coins. You know how coins are round, right? Well, a long time ago there were places that used coins that were all different shapes. Many years ago, lots of people from the place with round coins moved to all different places all over the world. After they moved, they got the people in their new town to use round coins instead. Then, slowly countries all over the world started to use round coins. So, just like fire trucks are red because of people a long time ago, that’s why we have round coins, because people a long time ago just happened to make it that way. There is no important reason that it is that way. Can you tell Feppy why coins are round? What did we read about? [if child mentions anything from the chapter beyond coins being round:] Yeah! That’s right! [if the child does not mention something from the chapter beyond coins being round:] Maybe we can tell him people just happened to start copying other towns. [Conclusion] We talked about why some things are the way they are. Some things are the way they are because of people long ago, but people could have easily make them another way. People might have had an idea and decided for it to be that way, or something might have happened to make it that way. So, clapping after people do a good job, fire trucks being red, coins being round, and drinking orange juice for breakfast are all just because of people long ago. And probably lots of other things today are the way they are just because of people long ago.</td>
</tr>
</tbody>
</table>