Journal of Experimental Child Psychology 177 (2019) 265–281

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Developmental evidence for a link between the inherence bias in explanation and psychological essentialism

Shelbie L. Sutherland [a,*], Andrei Cimpian [b]

[a] Department of Psychology, University of Toronto, Toronto, Ontario M5S 3G3, Canada
[b] Department of Psychology, New York University, New York, NY 10003, USA

ARTICLE INFO

Article history:
Received 9 April 2018
Revised 7 June 2018

Keywords:
Essentialism
Explanation
Heuristics
Inherence bias
Concepts
Development

ABSTRACT

The assumption that natural and social categories have deeper "essences" is a fundamental feature of the conceptual system, with wide-ranging consequences for behavior. What are the developmental origins of this assumption? We propose that essentialism emerges in part from a bias in the process of generating explanations that leads reasoners to overuse inherent or intrinsic features. Consistent with this proposal, the inherence bias in 4-year-olds' explanations predicted the strength of their essentialist beliefs (Study 1; $N = 64$), and manipulations of the inherence bias in 4- to 7-year-olds (Studies 2 and 3; $N = 112$ each) led to subsequent changes in the essentialist beliefs of children who attended to the manipulation. These findings contribute to our understanding of the origins of essentialism.

Introduction

People commonly believe that members of many natural and social categories (e.g., lions, boys) share an internal substance or "essence" that causes their category's typical properties (e.g., *having manes* for lions, *being interested in fixing things* for boys; Dar-Nimrod & Heine, 2011; Gelman, 2003;

* Corresponding author at: Department of Psychology, University of Toronto, Toronto, ON M5S 3G3, Canada
E-mail address: shelbie.sutherland@utoronto.ca (S.L. Sutherland).

https://doi.org/10.1016/j.jecp.2018.06.002
0022-0965/

Haslam, 2017; Heiphetz, Gelman, & Young, 2017; Rhodes & Mandalaywala, 2017).[1] For instance, people might believe that a lion roars and chases antelope because of some internal, microstructural, in-principle-identifiable thing that lions share. These essentialist beliefs are in many ways useful; for instance, they license broad inferences that go beyond what our senses may reveal (e.g., if one lion sleeps midday, we expect that another lion will as well). At the same time, essences lead us to overlook important differences among category members and treat them as interchangeable copies of one another, which has a range of negative effects, particularly in the social domain (e.g., Dar-Nimrod & Heine, 2011; Rhodes & Mandalaywala, 2017). Because essentialist beliefs so deeply shape how people understand and interact with the world, it is important to establish how these beliefs form.

Here, we investigated the proposal that essentialist beliefs emerge in part from a bias in the basic, early-developing processes by which children *generate explanations*—a bias that leads children to overuse intrinsic or inherent features in their explanations (Cimpian & Salomon, 2014a, 2014b; Salomon & Cimpian, 2014). Specifically, we tested two key predictions of this proposal: (a) that individual differences in children's broader explanatory reasoning will correlate with individual differences in essentialism and (b) that manipulations of children's explanatory reasoning will in turn affect their tendency to essentialize. Before elaborating these predictions, we review previously proposed sources of essentialism and articulate how our hypothesis contributes to this literature.

What are the origins of essentialist beliefs?

One likely source of essentialist beliefs is social input. Although adults do not mention essences explicitly in conversations with children (Gelman, Coley, Rosengren, Hartman, & Pappas, 1998; Gelman, Taylor, & Nguyen, 2004), they do talk in ways that imply their presence. In particular, adults often use what are known as generic statements or generics, which are statements expressing generalizations about entire categories (e.g., "boys like fixing things"). Exposure to multiple generic statements about a category suggests to children that the category has a deeper reality that supports these generalizations, which then leads children to essentialize that category (e.g., Gelman, Ware, & Kleinberg, 2010; Rhodes, Leslie, Saunders, Dunham, & Cimpian, 2018; Rhodes, Leslie, & Tworek, 2012). However, generic statements do not in and of themselves convey the notion of an essence. A child who could not already conceive of essences would not acquire this notion via exposure to generics. Instead, generics influence *which* categories children essentialize.

How, then, do children come to posit essences in the first place? To date, the most plausible theoretical account of this process proposes that essentialism emerges at the confluence of several pre-existing cognitive abilities and biases (Gelman, 2003). For example, infants seem to understand that objects belonging to the same category are similar in nonobvious ways even if they may *look* dissimilar (e.g., Dewar & Xu, 2009; Keates & Graham, 2008). This ability to distinguish between what things are like on the outside and what they are like deep down is an important prerequisite to being able to reason about internal unobserved essences that cause category-typical features. Another developmental precursor of essentialism is the motivation to make sense of the world by formulating explanations (e.g., Baillargeon, 1994; Baillargeon & DeJong, 2017; Saxe, Tenenbaum, & Carey, 2005; Stahl & Feigenson, 2015). As early as the first year of life, infants learn not just by passively observing their environments (e.g., an object fell off a shelf) but also by actively trying to determine how the observed phenomena came about (e.g., why the object fell; e.g., Baillargeon & DeJong, 2017). Similarly, young children seem to operate with an assumption of causal determinism; that is, they assume that all events and features have a cause even when they do not know what that cause is (e.g., Bullock, Gelman, & Baillargeon, 1982; Chandler & Lalonde, 1994; Muentener & Schulz, 2014; Schulz & Sommerville, 2006). Children's causal-explanatory curiosity may lead them to wonder why dogs bark

[1] Although the term "essence" has been defined in various ways (including, e.g., as an entity's history or unique identity; Bloom, 1996; De Freitas, Cikara, Grossmann, & Schlegel, 2017), we use it to refer specifically to a causal substance or entity present within members of natural and social kinds that confers kind identity and causes their kind-typical features; this is the most common use of this term (e.g., Cimpian & Salomon, 2014a, 2014b; Gelman, 2003).

and cats meow, why girls wear pink and boys do not, and so on—questions that, on this account, are ultimately answered by positing internal category essences.

Although we favor this theoretical account of the development of essentialism, prior formulations of it have left a key issue unanswered: In seeking explanations for category features, why do children so consistently appeal to internal causes? The motivation to seek explanations does not favor one type of explanation over another, nor does the assumption of causal determinism. Explanations couched in terms of external causes would satisfy children's curiosity just as well, and in most cases they would be more accurate (e.g., pink used to be gender neutral, so its association with girls is likely a result of historical developments; Paoletti, 2012; but see Del Giudice, 2012). Infants' ability to distinguish between surface and nonobvious features does not resolve this puzzle either; the mere awareness that there are nonobvious aspects of category membership does not, in and of itself, suggest that these aspects are somehow causally responsible for category-typical properties.

Here, we tested a possible solution to this puzzle. Specifically, we tested the hypothesis that a bias in the process of generating explanations reliably guides children toward essence-like explanations. When children act on their motivation to make sense of why categories have the features they do, the basic cognitive processes they invoke to generate explanations may make it more likely that their explanations appeal to inherent or intrinsic factors rather than extrinsic or external factors, which lays part of the foundation for essentialist reasoning (Cimpian & Salomon, 2014a, 2014b).

The inherence heuristic and its proposed link to essentialism

Like many of the judgments made in everyday life, explanations are often generated heuristically—by relying predominantly on information that is easy to access and manipulate in the moment (e.g., Cimpian, 2015; Cimpian & Salomon, 2014a, 2014b; Evans & Stanovich, 2013; Hussak & Cimpian, 2018b; Kahneman, 2011; Shah & Oppenheimer, 2008). This simplification greatly increases people's ability to generate an explanation even when the "why" question being considered is incredibly complex (e.g., why are most teachers women?). The disadvantage is that the explanations generated in this way are biased; some facts come to mind more readily than others for reasons *other* than their relevance to an accurate explanation, and as a result these facts are used in explanations more often than would be appropriate.

One class of such easy-access facts is inherent or intrinsic properties of the entities in the to-be-explained observation (the "explanandum"). Briefly, an inherent property of a thing is "entirely about that thing" rather than involving relations to other entities (Lewis, 1983, p. 197; see also Barr & Caplan, 1987; Caplan & Barr, 1991; Gentner & Kurtz, 2005; Weatherson & Marshall, 2018). For instance, the image of women as inherently warm might come immediately to mind for many people and serve as a heuristic explanation for the prevalence of female teachers. Across domains, such inherent facts are retrieved more frequently and rapidly from long-term memory than extrinsic facts (e.g., historical and societal trends, situational constraints) that are also available in memory but are less accessible (Hussak & Cimpian, 2018b). As a result, extrinsic facts tend to be overlooked when generating quick, in-the-moment explanations.

Extensive research has documented an inherence bias in both children's and adults' explanations for a wide range of explananda, from scientific observations to sociopolitical phenomena to historical events and social conventions (e.g., Cimpian & Erickson, 2012; Cimpian & Markman, 2009, 2011; Hussak & Cimpian, 2015, 2018a, 2018b; Sutherland & Cimpian, 2015; Tworek & Cimpian, 2016). Thus, when children try to understand why category members have the features they do, this bias in the process of coming up with explanations might focus children's attention on possible inherent reasons and guide it away from reasons that involve relations with external entities. In turn, this narrower explanatory focus is likely to bring children one step closer to reasoning in terms of essences, which are a specific type of inherent entity—an internal, identity-defining inherent entity.

Note that this inherence bias in explanation cannot, on its own, fully account for the development of beliefs about essences. For instance, essences are just a small subset of all inherent features, so other

factors must come into play to guide children's reasoning to this particular subset. In addition, intuitions about essences emerge only in particular domains—specifically, natural kinds (e.g., lions) and categories of people (e.g., girls, intelligent people)—whereas the inherence bias is domain-general, so other factors must be present to channel children's inherent explanations toward essences in some domains but not in others. A comprehensive account of how the inherence bias in explanation interacts with domain-specific conceptual resources (e.g., early assumptions about insides; Setoh, Wu, Baillargeon, & Gelman, 2013; Taborda-Osorio & Cheries, 2017) to give rise to essentialism is beyond the scope of the current article. (For a more detailed theoretical account, see Cimpian & Salomon, 2014a, 2014b.) Our goal here was simply to test the premise that the inherence bias and essentialism are linked in young children's reasoning.

Such a link was documented among adults by Salomon and Cimpian (2014). They found that individual differences in adults' inherence bias predicted individual differences in endorsement of essentialist beliefs about social groups. Moreover, an experimental manipulation of adults' inherence bias influenced the strength of their essentialist beliefs. In their experiments, Salomon and Cimpian asked half of their participants to rate their agreement with extrinsic explanations for 10 cultural conventions (e.g., "The only reason paper, money, and books are rectangular is historical happenstance") on a scale that was skewed toward agreement (see Petrocelli, Martin, & Li, 2010). The goal was not to actually assess participants' agreement with these explanations but rather to make extrinsic information more accessible in participants' minds, thereby weakening the inherence bias in their explanations. Relative to a group of control participants (who rated their agreement with statements that simply described the same patterns; e.g., "Most books, paper, and money are rectangular in shape"), participants who had been exposed to the extrinsic explanations subsequently showed lower endorsement of essentialist beliefs about social groups (e.g., "Males share an underlying property that causes them to have many similarities"). This condition difference in participants' essentialism was mediated by a corresponding difference in their endorsement of inherent explanations for other phenomena. Thus, Salomon and Cimpian (2014) provided correlational and experimental evidence for a link between the domain-general inherence bias in participants' explanations and their essentialism. However, the fact that their participants were adults precludes conclusions about development. In principle, this explanatory bias could have come to influence essentialist reasoning *after* the latter was already in place.

The current research

In this research, we focused on the developmental window during which essentialist beliefs first emerge—namely, the preschool and early elementary school years (e.g., Gelman, 2003; Rhodes & Mandalaywala, 2017). We tested two basic predictions of the claim that the inherence bias in explanation is part of the developmental process by which essentialist beliefs are formulated. The first prediction is that children who show a stronger inherence bias will also be more likely to have formed essentialist beliefs (Study 1). The second prediction is that experimentally manipulating the strength of this bias will have downstream consequences for children's essentialism (Studies 2 and 3).

Study 1

In Study 1, we investigated whether the inherence bias in children's explanations predicts the extent to which children endorse essentialist beliefs. We tested this relationship among 4-year-olds because 4 is (roughly) the age when essentialism is first expressed consistently (e.g., Gelman, 2003; Rhodes & Mandalaywala, 2017).

Method

The raw data and analytic syntax for this study and all subsequent studies are available on the Open Science Framework (OSF) at the following link: https://osf.io/nwvbr/?view_only=8a42c897d1234c37889307bc519c2bf5.

Participants

Children (4-year-olds) from a small city in the midwestern United States (N = 64, half boys and half girls; M_{age} = 4.40 years, SD = 0.32, range = 4.01–4.99) participated in this study. Children were recruited from a database of families interested in participating in research studies and from local schools and day-care centers. Demographic information was not formally collected.

Procedure and measures

Children were tested in a quiet room at a research lab or in their school. We asked children two sets of questions, in counterbalanced order, to assess their inherence bias and their essentialist beliefs.

Measure of inherence bias. We adapted a measure of children's explanatory biases from prior work (Cimpian & Steinberg, 2014; Sutherland & Cimpian, 2015; Tworek & Cimpian, 2016). The items in the measure concerned children's explanations for cultural conventions (e.g., the fact that coins are round), whereas essentialist beliefs were measured in a different domain—animal categories. Although the inherence bias in explanation applies across a range of domains (Cimpian & Salomon, 2014a, 2014b; Hussak & Cimpian, 2018b), we purposely chose to measure it outside the domain where essentialism applies most naturally (i.e., animal kinds); by avoiding overlap in the domains targeted by the measures of explanatory biases and essentialism, we minimized the possibility of finding an association between these measures for superficial reasons (i.e., simply because we asked children about the same types of entities). In other words, we intended to provide a conservative test of whether children's explanatory biases are related to their essentialist beliefs.

For each item, children were asked four types of questions. The first question asked them to evaluate inherent and extrinsic explanations. The other three types of questions tapped intuitions that follow from inherent explanations. If a particular phenomenon is explained in inherent terms (e.g., some inherent feature of coins explains why they are round), one might also reasonably assume that this pattern is consistent over time, cannot be changed, and is necessary. We detail each of these below.

Explanation evaluation. We presented children with inherent and extrinsic explanations, in counterbalanced order, and asked children whether these explanations were "right" or "not right." Notably, these explanations (e.g., coins are round "just because they are coins" [inherent] or "just because people thought it might be a nice idea" [extrinsic]) were intentionally nonspecific in the inherent features or extrinsic facts 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 did not mention the *particular* feature or fact that they might have spontaneously thought of themselves (for a similar strategy, see Cimpian & Steinberg, 2014; Sutherland & Cimpian, 2015; Tworek & Cimpian, 2016).

After children provided their initial "right" versus "not right" answers, we followed up and asked whether the explanations were "a little [not] right" or "really [not] right." Thus, children's responses were ultimately scored on a 4-point scale (*really not right* = 1, *a little not right* = 2, *a little right* = 3, *really right* = 4).

Children's rating for the extrinsic explanation was reverse scored. The two scores were then transformed to a scale from 0 to 1 and averaged to create a single explanation evaluation score, with higher scores indicating more agreement with inherent explanations than with extrinsic ones.

The remaining three question types measured intuitions that typically accompany inherent explanations. If children explain a phenomenon via some inherent feature (e.g., that coins are round because of something inherent about coins), they might also believe (a) that the phenomenon had always been and would always be that way, (b) that the phenomenon cannot be changed (even if everyone wanted it to be changed), and (c) that the phenomenon is necessary (i.e., needed to be exactly as it is).

Temporal stability of patterns. We asked children whether they thought that 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.

Unchangeability. We asked children whether they thought that 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*), which was then converted to a scale from 0 to 1 so that it could be averaged with the other measures.

Necessity. We tested whether children believed that observed phenomena could not have been otherwise; that is, we assessed their intuitions about necessity. For instance, if children believe that we call circles "circles" because of some feature of the word or the shape (an inherent explanation), then they might also believe that no other names would have been possible for this shape. Thus, we asked children whether an object could have been called something else when it was first being named (receiving a score of 0) or whether people needed to use its current name when they were first deciding on a name for it (receiving a score of 1; for additional details, see Sutherland & Cimpian, 2015).

The four questions above were asked about two cultural conventions, selected in counterbalanced fashion out of a set of three: coins being round, school buses being yellow, and birthday cakes having candles.

Children's responses to all questions were scored on a scale from 0 to 1, where higher scores indicated a stronger inherence bias. These scores were averaged to create a composite measure of each child's explanatory tendencies (M = .61, SD = .19).

Measure of essentialist beliefs. To capture the strength of children's essentialist beliefs, we adapted four types of questions from the prior literature. As detailed below, these questions measured children's (a) category-based inductive generalizations, (b) assumptions about the stability of category features across an individual's life, (c) reasoning about the innateness of category features, and (d) reasoning about the importance of insides for determining category membership.

Inductive generalization. If children believe that a category-specific essence causes members of animal kinds to have the properties 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, 2003). To measure this inference, we showed children a picture of an animal and told them about a nonvisible property of it (e.g., "Look at this cat. This cat has something called pedicles in its tail"). We then showed children three different members of the same kind, with varying visual similarity to the target, and for each asked children whether they thought that 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 [pointing to the original cat picture]?"; adapted from Gelman & Markman, 1987). If children said "yes," they were assigned a score of 1; if they said "no," they were assigned a score of 0.

Stability of individual traits. Another consequence of having essentialist beliefs about the source of an animal's properties is the belief that these properties are stable and unchanging.[2] 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 (or maintain) the property as long as the individual is a member of its category (e.g., Gelman, Heyman, & Legare, 2007). To test this belief, we asked children (a) whether a current property of an animal said to be 8 years old (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 (b) whether it would always have the same property (e.g., "Do you think this cat will always do this thing called 'flehming'?"; Gelman & Heyman, 1999; Gelman et al., 2007). For each question, children were assigned a score of 1 if they agreed or 0 if they disagreed.

Innateness. A commonly measured outcome of essentialism is the belief that biological (and arguably essence-related) causes of properties are more influential than environmental causes (e.g., Gelman et al., 2007; Gelman & Wellman, 1991; Taylor, Rhodes, & Gelman, 2009). 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

[2] The measure of temporal stability used in the inherence bias scale asked about an *entire artifact category* maintaining a property over the category's existence (e.g., have birthday cakes always had candles?). In contrast, the essentialism question asks about the stability of a feature within an *individual* belonging to an essentialized category.

that the dog had a baby and that 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 that they understood the story, we asked whether they thought that the baby (now 6 years old) has the property of the birth mother or of the adoptive mother (e.g., "Do you think that the baby eats bones like this dog, or does it eat grass like this horse?"; answer options were counterbalanced). If children chose the birth mother (an essentialist response), they received a score of 1, and if they chose the adoptive mother, they received a score of 0.

Importance of insides for determining category membership. Essentialism generally involves the belief that the category essence is internal (e.g., Gelman & Wellman, 1991); thus, insides are often taken to be more indicative of category membership. To test this aspect of essentialism, we used a modified version of Gelman and Wellman's (1991) test of whether the nonvisible insides or the visible properties of an animal are more important for determining its category membership. In the version we used, we told children that two teams 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 about the team's creation, they were asked memory questions to ensure understanding and were then asked whether or not the creation was a member of the relevant animal category (e.g., "Is what this team made a dog or not a dog?"; the order of the answer options was counterbalanced across children). If they said that it was a member of the intended animal kind, they were given a score of 1; if they did not, they were given a score of 0. The other team of scientists was said to make something with all of the right outsides but none of the right insides. We asked children follow-up memory and category identity questions for this team as well, but we reversed the scoring such that children received a score of 0 if they judged the team's creation with proper outsides but wrong insides to be a member of the relevant category.

These four types of essentialism questions all were asked about animal kinds. Two animal kinds were selected in counterbalanced fashion out of a set of three (cats, dogs, and birds) and were presented to each child as the targets of the questions above. Note, again, that we measured children's explanatory biases and essentialism with stimuli from different domains (cultural conventions vs. animal kinds, respectively). This way, any link we might find between these constructs cannot be attributed to superficial similarities in the content of the stimuli.

The scoring of this measure was similar to that of children's inherence bias; scores of 0 to 1 for each question were averaged into an essentialism composite ($M = .60$, $SD = .17$), with higher scores indicating more essentialist responses.

Results and discussion

If the inherence bias in children's explanations promotes the development of essentialist beliefs, we should see a positive correlation between measures of these two constructs at the age when essentialist beliefs are typically first formulated. Indeed, as illustrated in Fig. 1, these two measures were significantly correlated among 4-year-olds, $r(62) = .38$, $p = .002$.

Because both the inherence bias in explanation and children's essentialist beliefs change in strength over development (e.g., Cimpian & Steinberg, 2014; Taylor et al., 2009), it is possible that the observed correlation is simply a byproduct of such age-related changes. To test this possibility, we adjusted for children's age in a multiple regression; that is, we predicted the essentialism variable jointly on the basis of the inherence bias variable and children's exact age on the day they were tested (e.g., 4.40 years). Children's inherence bias was a significant predictor of their essentialism even when adjusting for age, $b = .34$, $SE = .10$, $p = .001$. (Age, however, did not predict essentialist beliefs, $b = -.10$, $SE = .06$, $p = .117$). Thus, the relationship between children's inherence and essentialist biases is not an artifact of simultaneous changes with age in these dimensions.

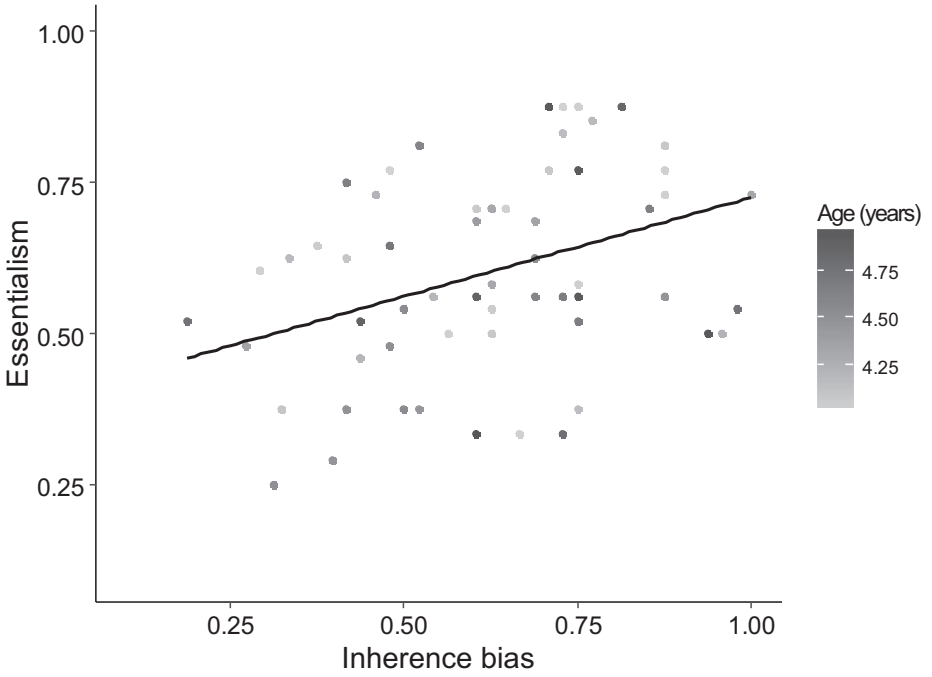

Fig. 1. Relationship between children's scores on the inherence bias and essentialism measures in Study 1. The color of the point indicates the child's age.

Study 2

Study 1 suggests a relationship between children's broader explanatory tendencies and their essentialism, but it does not identify the direction of this relationship. Does the inherence bias in children's explanations have a causal influence over children's endorsement of essentialist beliefs? If so, then manipulating children's explanations should lead to corresponding changes in their essentialist beliefs. This is the prediction we tested in Study 2.

Method

Participants

In this study, we recruited a group of children aged 4 and 5 years ($n = 56$, half boys and half girls; $M_{age} = 4.98$ years, $SD = 0.62$, range = 4.08–5.96) and a group of children aged 6 and 7 years ($n = 56$, half boys and half girls; $M_{age} = 6.89$ years, $SD = 0.57$, range = 6.02–7.98). We expanded the age range beyond 4-year-olds (Study 1) for two reasons. First, the broader age range allowed us to explore whether the link between the inherence bias and essentialism varies in strength with age. Second, the manipulation was relatively demanding from an information-processing standpoint; thus, it was not clear to us a priori that children as young as 4 years would be sensitive to it. Children were recruited from the same population as in Study 1 but had not participated in that study.

Overview of procedure

To manipulate the inherence bias in children's explanations, we read children a book intended to either promote or discourage inherent explanations of cultural conventions. Following the logic of Study 1, we kept the domain in which we manipulated the inherence bias (cultural conventions)

separate from the domain in which we measured essentialism (animal kinds). We assessed children's endorsement of essentialist beliefs twice: once before they read the picture book (pretest) and once after (posttest).

Manipulation of inherence bias

The experimenter read a storybook to children about "why some things are the way they are." The book had four "chapters," each of which discussed a different cultural convention (e.g., orange juice for breakfast); the chapters were presented in one of two random orders. For children in the "Pro-inherence" condition, the book provided inherent explanations for the cultural conventions (e.g., orange juice has Vitamin C) as well as other inferences that often accompany inherent explanations (e.g., it makes perfect sense that orange juice is for breakfast). Alternatively, for children in the "Anti-inherence" condition, the book provided extrinsic explanations for the same conventions (e.g., people decided that they wanted it to be that way) along with additional inferences that might follow from these explanations (e.g., in the future, people might decide to drink other things for breakfast). The books are available at the OSF link provided above, and the full script is provided in Appendix S1 in the online supplementary material.

After each chapter, the children were asked to describe what they had heard in the book to a stuffed animal. These questions served as an attention check. Any mention of an element from the preceding chapter (besides simply repeating the convention itself, such as that coins are round) was considered a satisfactory response (a "pass") for the attention check for that chapter. We used children's performance on these attention checks as a moderator in our analyses; the manipulation can only be effective if children pay attention to and understand the material.

To encourage further engagement, at the end of the book the experimenter asked children to brainstorm what a new chapter in the book could be about. The experimenter made a suggestion and appropriate prompts to keep children in line with the general idea of the book (see Appendix S1).

Measure of essentialist beliefs

The questions used to measure children's essentialist beliefs were similar to those used in Study 1 except that we omitted the question about the importance of insides for determining category membership (which was the longest) to keep the sessions at a manageable length. Essentialist beliefs were measured twice: before and after the manipulation. The pretest and posttest each consisted of the three essentialism questions (induction, stability, and innateness) asked about two items, drawn in counterbalanced order from a set of four animal kinds: cats, dogs, birds, and frogs. Thus, children were asked about two animals in the pretest and about two different animals in the posttest.

Data analysis

Children's responses were analyzed with a mixed-effects multilevel model in which the fixed effects were condition (Pro-inherence vs. Anti-inherence), age group (4- and 5-year-olds vs. 6- and 7-year-olds), phase (pretest vs. posttest), attention check status (passed all attention checks vs. did not),[3] and all possible interactions among these variables. The model also included crossed random intercepts for subject, item, and essentialism question. We used bootstrapping (1000 replications) to compute standard errors, *p* values, and 95% confidence intervals (*CI*s) based on the 2.5th and 97.5th percentiles of the bootstrap distribution. All predictors were mean centered to facilitate interpretation of the fixed-effects coefficients (Anti-inherence = −.50 vs. Pro-inherence = .50; 4- and 5-year-olds = −.50 vs. 6- and 7-year-olds = .50; pretest = −.50 vs. posttest = .50; failed attention check = −.37 vs. passed attention check = .63).[4]

[3] The same pattern of results was obtained with alternative ways of including the attention checks in the analyses (e.g., as a continuous variable; see the General Discussion).

[4] The values for the mean-centered attention check variable are not symmetrical around 0 because there were uneven numbers of children in the two groups.

Results and discussion

We predicted that essentialism scores would change from pretest to posttest as a function of the book children heard (i.e., whether they were in the Pro- or Anti-inherence condition). Because these condition-dependent changes might be larger for children who passed the attention checks ($n = 41$; Pro-inherence: $n = 24$; Anti-inherence: $n = 17$) than for those who did not ($n = 71$; Pro-inherence: $n = 32$; Anti-inherence: $n = 39$), we also explored whether children's attention check status is a moderator of the Phase × Condition interaction (see Table S1 in the supplementary material for means). The two-way Phase × Condition interaction was not significant, $b = .04$, $CI = [-.03, .11]$, $SE = .04$, $p = .315$, but the three-way Phase × Condition × Attention Check Status was marginally significant, $b = .15$, $CI = [-.01, .30]$, $SE = .08$, $p = .056$ (see Table S2 in the supplementary material for the full table of fixed effects).

Next, we decomposed this three-way interaction by testing whether the two-way Phase × Condition interaction was present for the children who passed (vs. failed) the attention checks. Indeed, for children who passed the attention checks, the change in essentialist beliefs from pretest to posttest depended significantly on the condition children were in, $b = .10$, $CI = [.01, .20]$, $SE = .05$, $p = .026$. This interaction was driven mostly by the Pro-inherence condition, where the pretest-to-posttest increase was significant, $b = .09$, $CI = [.003, .17]$, $SE = .05$, $p = .042$ (see Fig. 2).[5] (The pretest-to-posttest change was nonsignificant for children in the Anti-inherence condition, $b = -.01$, $CI = [-.10, .09]$, $SE = .05$, $p = .844$.) Children who *failed* the attention checks did not update their essentialist beliefs in either condition (Phase × Condition interaction: $b = -.01$, $CI = [-.09, .07]$, $SE = .04$, $p = .846$; see also Fig. S1 in the supplementary material). Finally, no relationships involving age group reached significance.

These findings are broadly consistent with our proposal, insofar as they suggest that intervening to change children's explanatory tendencies may change the extent to which children endorse essentialist beliefs. However, the support for our prediction was somewhat limited; only children who passed the attention checks showed the predicted effects and only in one of the two conditions (Pro-inherence). To verify the robustness of these results, we conducted a replication study that differed from Study 2 only in minor respects (e.g., a shorter pretest).

Study 3

Method

Participants

As in Study 2, we recruited 56 children aged 4 and 5 years (half boys and half girls; $M_{age} = 4.89$ years, $SD = 0.64$, range = 4.00–5.99) and 56 children aged 6 and 7 years (half boys and half girls; $M_{age} = 7.02$ years, $SD = 0.56$, range = 6.04–8.06) for a total of 112 participants. Children were recruited from the same population as those in Studies 1 and 2 but had not participated in those studies.

Materials

We made several changes to the materials from Study 2. To prevent children's attention from waning, we shortened the sessions by using only one item for the essentialism pretest (the bird item was dropped from the set) and by removing the brainstorming phase at the end of the book. To help children understand and retain the information in the books, we presented the pro- and anti-inherence material in the form of short narratives (as opposed to just stating inherent and extrinsic reasons as in Study 2) and added a brief introduction and conclusion to reiterate the main take-aways from each book (see Appendix S2 in the supplementary material for the script). Finally, we made two minor changes to the essentialism measure, detailed in Table S3.

[5] The pretest-to-posttest increase was evident on all three questions measuring essentialism (inductive generalization: $M_{pre} = .33$, $M_{post} = .44$; stability of individual traits: $M_{pre} = .47$, $M_{post} = .59$; innateness: $M_{pre} = .63$, $M_{post} = .67$).

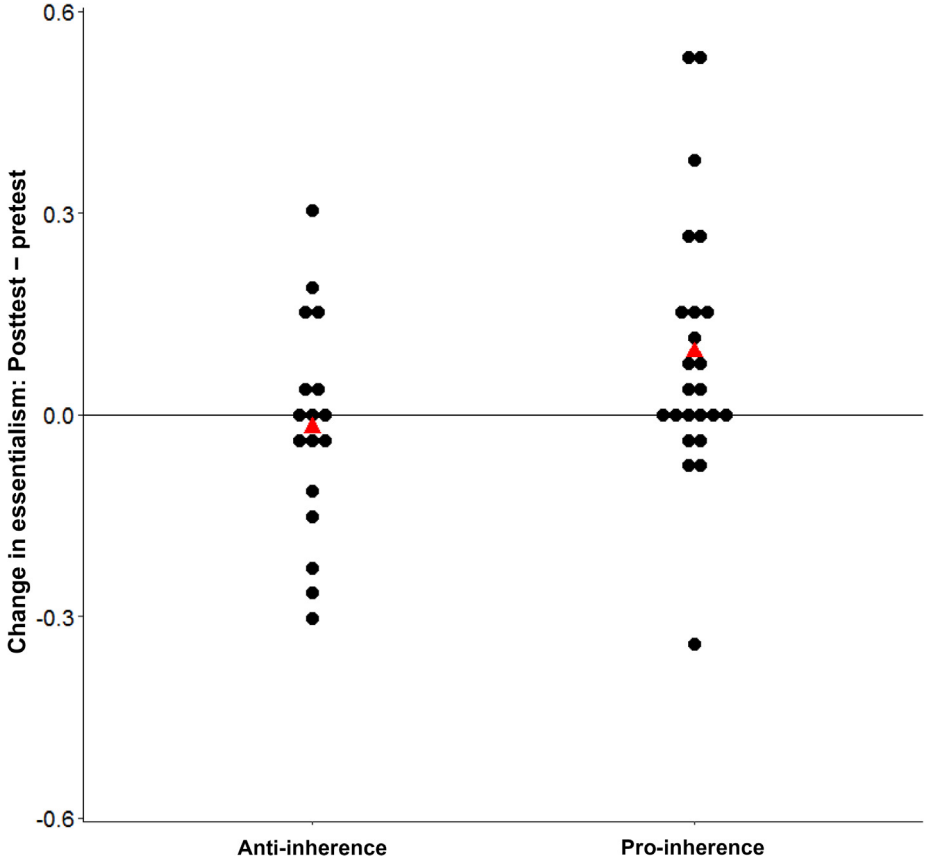

Fig. 2. Dot plot of pretest-to-posttest changes in essentialist responses (posttest minus pretest) by Anti-inherence versus Pro-inherence condition (Study 2). Only the responses of children who passed the attention checks are included in this figure. Each dot represents one child's difference score. Dot plots are similar to histograms in that they indicate the frequency of different values of the dependent measure. The red triangles represent the mean. (For interpretation of the reference to color in this figure legend, the reader is referred to the Web version of this article.)

Results and discussion

The findings replicated those of Study 2. We again found a significant three-way Phase × Condition × Attention Check Status interaction, indicating that our manipulation of children's explanatory tendencies influenced the extent to which children who passed (vs. failed) the attention checks endorsed essentialist beliefs about animal kinds, $b = .20$, $CI = [.04, .38]$, $SE = .09$, $p = .021$. This significant three-way interaction emerged because the two-way Phase × Condition interaction was significant for children who passed the attention checks ($n = 44$; Pro-inherence: $n = 18$; Anti-inherence: $n = 26$), $b = .13$, $CI = [.001, .26]$, $SE = .06$, $p = .050$, but not for those who failed some of these checks ($n = 68$; Pro-inherence: $n = 38$; Anti-inherence: $n = 30$), $b = -.01$, $CI = [-.10, .08]$, $SE = .05$, $p = .842$. As in Study 2, the Pro-inherence condition, $b = .15$, $CI = [.04, .27]$, $SE = .06$, $p = .016$,[6] was more effective than the Anti-inherence condition, $b = .02$, $CI = [-.07, .11]$, $SE = .05$, $p = .630$, in altering the essentialist beliefs of children who passed the attention checks (see Fig. 3; see also Tables S4 and S5 in the supplementary material for means and all fixed effects, respectively).

[6] The pretest-to-posttest increase was evident on all three questions measuring essentialism (inductive generalization: $M_{pre} = .43$, $M_{post} = .50$; stability of individual traits: $M_{pre} = .25$, $M_{post} = .43$; innateness: $M_{pre} = .61$, $M_{post} = .81$).

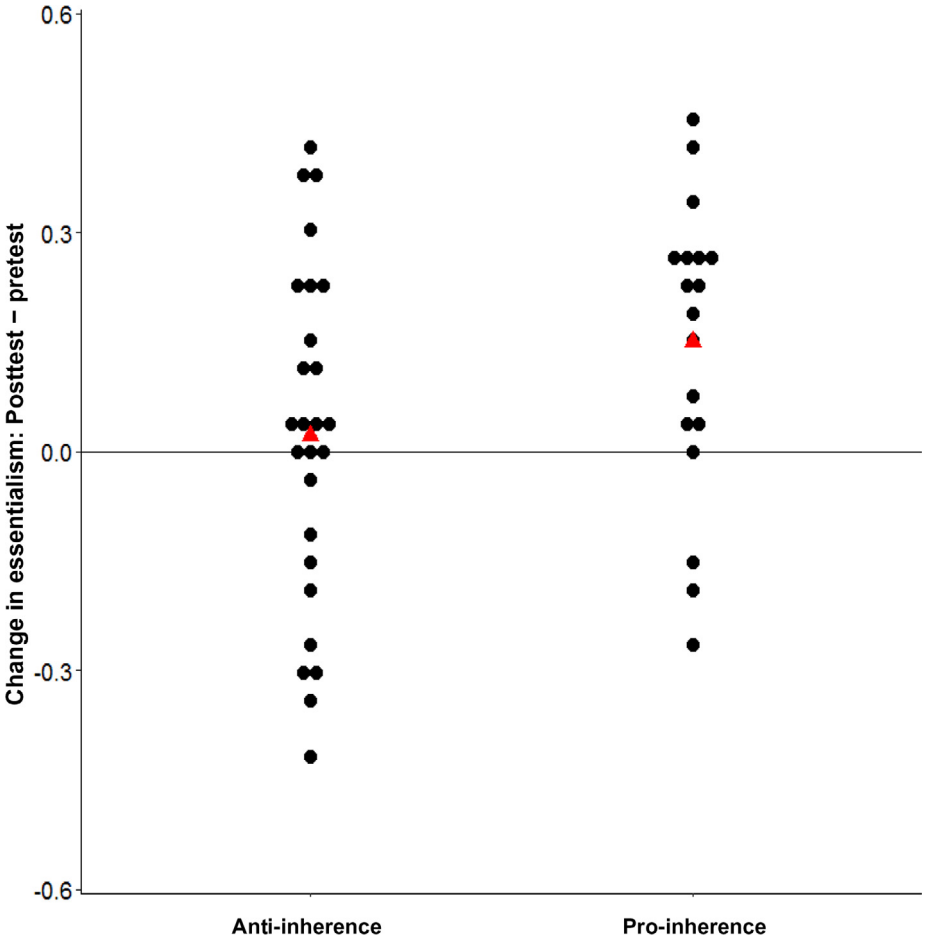

Fig. 3. Dot plot of pretest-to-posttest changes in essentialist responses (posttest minus pretest) by Anti-inherence versus Pro-inherence condition (Study 3). Only the responses of children who passed the attention checks are included in this figure. Each dot represents one child's difference score. Dot plots are similar to histograms in that they indicate the frequency of different values of the dependent measure. The red triangles represent the mean. (For interpretation of the reference to color in this figure legend, the reader is referred to the Web version of this article.)

Meta-analysis of Studies 2 and 3

To further assess the robustness of our findings, we meta-analyzed the key results across Studies 2 and 3. Such "mini meta-analyses" across the studies in a single article are increasingly recommended as a way of assessing replicability and precisely estimating coefficients (e.g., Cumming, 2014; Goh, Hall, & Rosenthal, 2016). We conducted our meta-analysis with the *metan* package in Stata 13.1 using a random-effects model, which allows the magnitude of the effects to vary across studies. Mirroring the results of Studies 2 and 3, the meta-analysis revealed a significant three-way Phase × Condition × Attention Check Status interaction, $b_{meta} = .17$, $CI = [.06, .29]$, $z = 3.00$, $p = .003$, as well as a significant two-way Phase × Condition interaction among children who passed the attention check, $b_{meta} = .11$, $CI = [.04, .19]$, $z = 2.86$, $p = .004$. Children's essentialism increased significantly from pretest to posttest in the Pro-inherence condition, $b_{meta} = .11$, $CI = [.04, .18]$, $z = 3.16$, $p = .002$, but did not decrease in the Anti-inherence condition, $b_{meta} = .004$, $CI = [-.06, .07]$, $z = 0.12$, $p = .905$. In summary,

this meta-analysis replicated the results reported for the individual studies and highlighted their replicability and robustness.

General discussion

We proposed that biases in the process by which children generate explanations promote the development of essentialist beliefs. The current findings provide support for two basic predictions of this proposal. First, Study 1 suggested that the inherence bias in children's explanations predicts the strength of their essentialist beliefs at the age when these beliefs begin to emerge. This relationship was observed even though children's explanatory reasoning was measured about phenomena that fall outside the scope of causal essentialism (namely, cultural conventions). Second, Studies 2 and 3 provided some support for a causal relationship between explanatory biases and essentialism. We found that the strength of children's essentialist beliefs was influenced by reading a book that explained various cultural conventions (e.g., why coins are round) in inherent versus extrinsic terms; this effect, however, was apparent only among children who encoded the content of the manipulation books and could repeat some of it to an experimenter. Finally, a meta-analysis of Studies 2 and 3 provided additional evidence regarding the magnitude and robustness of these experimental effects. Taken as a whole, the results presented here are consistent with the proposal that the development of essentialism is facilitated in part by a domain-general explanatory bias toward inherent features.

Theoretical contribution

Assumptions about unseen but causally powerful essences are a key aspect of concepts of natural kinds and social groups. Although hundreds of studies have documented the pervasiveness of these assumptions and their downstream consequences (for reviews, see Dar-Nimrod & Heine, 2011; Gelman, 2003), research on their origins is conspicuously scarce. With the exception of an important line of work on the role of language in promoting essentialism (e.g., Gelman & Heyman, 1999; Gelman et al., 2010; Rhodes et al., 2012), few empirical studies have set out to investigate how essentialist assumptions arise in children. This is a critical gap; development is likely to hold many crucial insights into how essentialism functions, why it is so pervasive, and what may help to combat it when needed (e.g., when it fosters stereotyping and prejudice). The current studies begin to fill this gap in the literature by documenting a link between essentialism and the inherence bias in the content of children's intuitive explanations. This research also helps to "demystify" the origins of essentialist beliefs by placing this phenomenon in the broader context of the basic cognitive processes involved in selective attention, memory retrieval, and heuristic versus analytic reasoning (e.g., Eidson & Coley, 2014; Evans & Stanovich, 2013; Hussak & Cimpian, 2018b). Although many details remain to be filled in about the developmental process by which a bias toward inherent explanations ultimately gives rise to essentialism, the evidence presented here highlights this developmental process as a promising topic for further investigation.

Is the inherence bias a precursor of essentialism or vice versa?

We proposed that the inherence bias in explanation precedes, and contributes to the development of, essentialism. Is it possible, however, that essentialism is instead the source of this inherence bias? The timing of the emergence of these two phenomena speaks against this possibility. A broad inherence bias has been documented in children as young as 4 years (e.g., Cimpian & Steinberg, 2014; Hussak & Cimpian, 2015; Sutherland & Cimpian, 2015) and potentially even among 3-year-olds (Cimpian & Tworek, 2015). If the inherence bias were to emerge as a domain-general extension of an earlier, domain-specific essentialist assumption, then one would expect to see essentialist reasoning proper (not just specific precursors) in place well before its supposed domain-general offshoot. However, this is not the case; the first reliable signs of essentialism emerge after 4 years of age (e.g., Gelman, 2003), which limits the plausibility of the reverse developmental account. In addition, the current studies and those by Salomon and Cimpian (2014) provide positive evidence of a causal

influence of the inherence bias on essentialism, which is more consistent with the former being a precursor of the latter than vice versa. Nevertheless, it is possible that once children begin to hold essentialist beliefs, these beliefs feed back into and reinforce children's more general tendency to explain via inherent features. That is, although it is unlikely that the inherence bias in explanation is an offshoot of essentialism, the emergence of essentialism could potentially strengthen this domain-general explanatory bias. This possibility awaits further empirical study.

Limitations and future directions

A limitation of this research is that the effects in Studies 2 and 3 were present in only a subset of participants—those who reliably encoded the content of the books they were read. This aspect of our results is probably due to a combination of factors. The inherence bias in children's explanations may not be easily changed given that it is an output of the basic processes involved in explaining (e.g., memory retrieval; Cimpian & Salomon, 2014a, 2014b). Moreover, our manipulation was minimal; the book read to children contained inherent or extrinsic explanations for only four observed facts, and the experimenter read this book only once. Perhaps exposing children to this information repeatedly (e.g., by providing parents with the books to read at home; see Gelman et al., 2010; Rhodes et al., 2018) would facilitate children's uptake of the crucial inherent and extrinsic ideas and, as a result, increase the magnitude of the effects. Repeated exposure may be particularly important when the information contradicts children's existing intuitions, which is the case for the extrinsic books in particular. (Consistent with this possibility, the extrinsic books did not effectively decrease children's essentialism scores in the current, one-shot setup.)

It is also interesting to consider the possibility that the attention check moderator tracked differences among children in respects *other* than their attentiveness. For example, if children did not agree with the explanations provided in a particular chapter, they may have chosen to not repeat them. If so, what we coded as "attentiveness" may have reflected in part children's prior explanatory tendencies: Children who passed all attention checks may have done so because they already agreed with all the inherent or extrinsic explanations they heard. On this alternative, our manipulation did not influence children's essentialism; rather, our effects in Studies 2 and 3 were simply correlational, with children who came in with a strong inherence bias showing more essentialism than those who preferred extrinsic explanations. However, the fact that Studies 2 and 3 had a pretest/posttest design (rather than simply comparing children's essentialism at posttest) speaks against this possibility to some extent. The finding of an effect among the children who passed the attention check necessarily reflects *changes* in their endorsement of essentialist beliefs.

Alternatively, the attention check may have tracked differences among children in their ability and willingness to learn from others. Studies 2 and 3 set up a social learning context, with knowledgeable experimenters providing their explanatory perspective to children via a book. Young children differ in their levels of epistemic trust toward unfamiliar adults (e.g., Tagar, Federico, Lyons, Ludeke, & Koenig, 2014). Perhaps such individual differences led some children more than others to believe the experimenters and adopt their explanatory preferences, which would then translate into better performance on our attention checks. In other words, perhaps our manipulation was most effective among the children who were epistemically trusting and, thus, more likely to engage in social learning. From this perspective, our findings are nicely complementary to prior work suggesting that children's sociocultural environments shape the development of their essentialist beliefs (e.g., Rhodes & Gelman, 2009).

Despite these limitations, we should reiterate that there was also evidence for the robustness of the effects of our inherence manipulation; we found these effects in both Studies 2 and 3, as well as when we meta-analyzed these two experiments. Moreover, we note that in further exploratory analyses we found the same pattern of results when we treated the attention check as a continuous variable and tested the effects of the manipulation at 1 standard deviation above and below the average attention scores. For example, in a model with both Studies 2 and 3, this analysis revealed a significant effect of the manipulation at +1 standard deviation but not at −1 standard deviation in attention check scores, consistent with the claim that children who did well on the attention checks also showed changes in their essentialism following the manipulation.

Conclusion

The current work is consistent with both the broad claim that essentialism develops out of a set of domain-general precursors (Gelman, 2003) and the narrower claim that one such precursor is the bias that skews children's explanations toward inherent features. Our studies advance scientific understanding of the origins of essentialism, an area that is understudied despite its theoretical importance and its potential to reveal effective ways of shaping the wide-reaching consequences of essentialist beliefs.

Acknowledgments

This research was funded in part by the American Psychological Foundation via the Elizabeth Munsterberg Koppitz Graduate Fellowship to Shelbie Sutherland and by research funds from the University of Illinois and New York University to Andrei Cimpian. It benefited from helpful comments and discussion from Renée Baillargeon, Lin Bian, Jilana Boston, Zachary Horne, Larisa Hussak, Tara Mandalaywala, Melis Muradoglu, Erika Salomon, Daniel Storage, and Christina Tworek. We are grateful to the participating children and their families; to the principals, teachers, and staff at the participating schools; and to the Cognitive Development Lab team for their assistance in collecting data.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jecp.2018.06.002.

References

Baillargeon, R. (1994). Physical reasoning in young infants: Seeking explanation for impossible events. *British Journal of Developmental Psychology, 12*, 9–33.
Baillargeon, R., & DeJong, G. F. (2017). Explanation-based learning in infancy. *Psychonomic Bulletin & Review, 24*, 1511–1526.
Barr, R. A., & Caplan, L. J. (1987). Category representations and their implications for category structure. *Memory & Cognition, 15*, 397–418.
Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition, 60*, 1–29.
Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 209–254). London: Academic Press.
Caplan, L. J., & Barr, R. A. (1991). The effects of feature necessity and extrinsicity on gradedness of category membership and class inclusion relations. *British Journal of Psychology, 82*, 427–440.
Chandler, M. J., & Lalonde, C. E. (1994). Surprising, magical, and miraculous turns of events: Children's reactions to violations of their early theories of mind and matter. *British Journal of Developmental Psychology, 12*, 83–95.
Cimpian, A. (2015). The inherence heuristic: Generating everyday explanations. In R. Scott & S. Kosslyn (Eds.), *Emerging trends in the social and behavioral sciences* (pp. 1–15). Hoboken, NJ: John Wiley.
Cimpian, A., & Tworek, C. (2015). Three-year-olds explain via inherent features, not essences: Evidence that essentialism develops out of an explanatory heuristic. Paper presented at the biennial meeting of the Society for Research in Child Development, Philadelphia.
Cimpian, A., & Erickson, L. C. (2012). The effect of generic statements on children's causal attributions: Questions of mechanism. *Developmental Psychology, 48*, 159–170.
Cimpian, A., & Markman, E. M. (2009). Information learned from generic language becomes central to children's biological concepts: Evidence from their open-ended explanations. *Cognition, 113*, 14–25.
Cimpian, A., & Markman, E. M. (2011). The generic/nongeneric distinction influences how children interpret new information about social others. *Child Development, 82*, 471–492.
Cimpian, A., & Salomon, E. (2014a). The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences, 37*, 461–480.
Cimpian, A., & Salomon, E. (2014b). Refining and expanding the proposal of an inherence heuristic in human understanding. *Behavioral and Brain Sciences, 37*, 506–527.
Cimpian, A., & Steinberg, O. D. (2014). The inherence heuristic across development: Systematic differences between children's and adults' explanations for everyday facts. *Cognitive Psychology, 75*, 130–154.
Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*, 7–29.
Dar-Nimrod, I., & Heine, S. J. (2011). Genetic essentialism: On the deceptive determinism of DNA. *Psychological Bulletin, 137*, 800–818.
De Freitas, J., Cikara, M., Grossmann, I., & Schlegel, R. (2017). Origins of the belief in good true selves. *Trends in Cognitive Sciences, 21*, 634–636.
Del Giudice, M. (2012). The twentieth century reversal of pink–blue gender coding: A scientific urban legend? *Archives of Sexual Behavior, 41*, 1321–1323.

Dewar, K., & Xu, F. (2009). Do early nouns refer to kinds or distinct shapes? *Psychological Science, 20*, 252–257.
Eidson, R. C., & Coley, J. D. (2014). Not so fast: Reassessing gender essentialism in young adults. *Journal of Cognition and Development, 15*, 382–392.
Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*, 223–241.
Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York: Oxford University Press.
Gelman, S. A., Coley, J. D., Rosengren, K. S., Hartman, E., & Pappas, A. (1998). Beyond labeling: The role of maternal input in the acquisition of richly structured categories. *Monographs of the Society for Research in Child Development, 63* (1, Serial No. 253).
Gelman, S. A., & Heyman, G. D. (1999). Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about social categories. *Psychological Science, 10*, 489–493.
Gelman, S. A., Heyman, G. D., & Legare, C. H. (2007). Developmental changes in the coherence of essentialist beliefs about psychological characteristics. *Child Development, 78*, 757–774.
Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58*, 1532–1541.
Gelman, S. A., Taylor, M. G., & Nguyen, S. P. (2004). Mother–child conversations about gender: Understanding the acquisition of essentialist beliefs. *Monographs of the Society of Research in Child Development, 69* (1, Serial No. 275).
Gelman, S. A., Ware, E. A., & Kleinberg, F. (2010). Effects of generic language on category content and structure. *Cognitive Psychology, 61*, 273–301.
Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition, 38*, 213–244.
Gentner, D., & Kurtz, K. J. (2005). Relational categories. In W. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff (Eds.), *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin* (pp. 151–175). Washington, DC: American Psychological Association.
Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass, 10*, 535–549.
Haslam, N. (2017). The origins of lay theories: The case of essentialist beliefs. In C. M. Zedelius, B. C. N. Muller, & J. W. Schooler (Eds.), *The science of lay theories* (pp. 3–16). Cham, Switzerland: Springer International.
Heiphetz, L., Gelman, S. A., & Young, L. L. (2017). The perceived stability and biological basis of religious beliefs, factual beliefs, and opinions. *Journal of Experimental Child Psychology, 156*, 82–98.
Hussak, L. J., & Cimpian, A. (2015). An early-emerging explanatory heuristic promotes support for the status quo. *Journal of Personality and Social Psychology, 109*, 739–752.
Hussak, L. J., & Cimpian, A. (2018b). Memory accessibility shapes explanation: Testing key claims of the inherence heuristic account. *Memory & Cognition, 46*, 68–88.
Hussak, L. J., & Cimpian, A. (2018a). Investigating the origins of political views: Biases in explanation predict conservative attitudes in children and adults. *Developmental Science, 21* e12567.
Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus, & Giroux.
Keates, J., & Graham, S. A. (2008). Category markers or attributes: Why do labels guide infants' inductive inferences. *Psychological Science, 19*, 1287–1293.
Lewis, D. (1983). Extrinsic properties. *Philosophical Studies, 44*, 197–200.
Muentener, P., & Schulz, L. (2014). Toddlers infer unobserved causes for spontaneous events. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.01496.
Paoletti, J. B. (2012). *Pink and blue: Telling the boys from the girls in America*. Bloomington: Indiana University Press.
Petrocelli, J. V., Martin, J. L., & Li, W. Y. (2010). Shaping behavior through malleable self-perceptions: A test of the forced-agreement scale effect (FASE). *Journal of Research in Personality, 44*, 213–221.
Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology, 59*, 244–274.
Rhodes, M., Leslie, S.-J., Saunders, K., Dunham, Y., & Cimpian, A. (2018). How does social essentialism affect the development of inter-group relations? *Developmental Science, 21*, e12509.
Rhodes, M., Leslie, S. J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 13526–13531.
Rhodes, M., & Mandalaywala, T. M. (2017). The development and developmental consequences of social essentialism. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*. https://doi.org/10.1002/wcs.1437.
Salomon, E., & Cimpian, A. (2014). The inherence heuristic as a source of essentialist thought. *Personality and Social Psychology Bulletin, 40*, 1297–1315.
Saxe, R., Tenenbaum, J. B., & Carey, S. (2005). Secret agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science, 16*, 995–1001.
Schulz, L. E., & Sommerville, J. (2006). God does not play dice: Causal determinism and preschoolers' causal inferences. *Child Development, 77*, 427–442.
Setoh, P., Wu, D., Baillargeon, R., & Gelman, R. (2013). Young infants have biological expectation about animals. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 15937–15942.
Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin, 134*, 207–222.
Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science, 348*, 91–94.
Sutherland, S. L., & Cimpian, A. (2015). An explanatory heuristic gives rise to the belief that words are well suited for their referents. *Cognition, 143*, 228–240.
Taborda-Osorio, H., & Cheries, E. W. (2017). Developmental origins of biological explanations: The case of infants' internal property bias. *Psychonomic Bulletin & Review, 24*, 1527–1537.
Tagar, M. R., Federico, C. M., Lyons, K. E., Ludeke, S., & Koenig, M. A. (2014). Heralding the authoritarian? Orientation toward authority in early childhood. *Psychological Science, 25*, 883–892.
Taylor, M. G., Rhodes, M., & Gelman, S. A. (2009). Boys will be boys; cows will be cows: Children's essentialist reasoning about gender categories and animal species. *Child Development, 80*, 461–481.

Tworek, C. M., & Cimpian, A. (2016). Why do people tend to infer "ought" from "is"? The role of biases in explanation. *Psychological Science, 27*, 1109–1122.

Weatherson, B., & Marshall, D. (2018). Intrinsic vs. extrinsic properties. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Stanford, CA: Center for the Study of Language and Information. https://plato.stanford.edu/entries/intrinsic-extrinsic.