# The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism

**Andrei Cimpian**
*Department of Psychology, University of Illinois, Champaign, IL 61820*
**acimpian@psych.illinois.edu**
**http://psychology.illinois.edu/people/acimpian**

**Erika Salomon**
*Department of Psychology, University of Illinois, Champaign, IL 61820*
**salomon3@illinois.edu**
**http://www.erikasalomon.com**

**Abstract:** We propose that human reasoning relies on an *inherence heuristic*, an implicit cognitive process that leads people to explain observed patterns (e.g., girls wear pink) predominantly in terms of the inherent features of their constituents (e.g., pink is a delicate color). We then demonstrate how this proposed heuristic can provide a unified account for a broad set of findings spanning areas of research that might at first appear unrelated (e.g., system justification, nominal realism, is–ought errors in moral reasoning). By revealing the deep commonalities among the diverse phenomena that fall under its scope, our account is able to generate new insights into these phenomena, as well as new empirical predictions. A second main goal of this article, aside from introducing the inherence heuristic, is to articulate the proposal that the heuristic serves as a foundation for the development of psychological essentialism. More specifically, we propose that *essentialism* – which is the common belief that natural and social categories are underlain by hidden, causally powerful essences – emerges over the first few years of life as an elaboration of the earlier, and more open-ended, intuitions supplied by the inherence heuristic. In the final part of the report, we distinguish our proposal from competing accounts (e.g., Strevens's K-laws) and clarify the relationship between the inherence heuristic and related cognitive tendencies (e.g., the correspondence bias). In sum, this article illuminates a basic cognitive process that emerges early in life and is likely to have profound effects on many aspects of human psychology.

**Keywords:** correspondence bias; development; essentialism; explanation; inherence heuristic; is–ought problem; nominal realism; system justification

## 1. Introduction

The ability to identify and exploit the predictable aspects of a complex environment is, without doubt, part of what makes humans such a successful species (e.g., Murphy 2004; Saffran et al. 1996; Zhao et al. 2013). Even the youngest members of our species are able to detect the broad patterns that characterize their world: that boys wear blue and girls wear pink, that orange juice is consumed for breakfast, that giraffes are called *giraffes*, and so on. In the present article, we propose that people often make sense of such regularities via a simple rule of thumb – the *inherence heuristic*. This fast, intuitive heuristic leads people to explain many observed patterns[1] in terms of the inherent features of the things that instantiate these patterns. For example, one might infer that girls wear pink because pink is a delicate, inherently feminine color, or that orange juice is consumed for breakfast because its inherent qualities make it suitable for that time of day. As is the case with the output of any heuristic, such inferences can be – and often are – mistaken. Many of the patterns

that currently structure our world are the products of complex chains of historical causes rather than being

ANDREI CIMPIAN is Associate Professor of Psychology and a 2012–2013 Fellow of the Center for Advanced Study at the University of Illinois at Urbana–Champaign. He received his Ph.D. from Stanford University in 2008. His research focuses on the development of human concepts about natural and social categories, as well as on the development of beliefs about ability and achievement.

ERIKA SALOMON is a Ph.D. student in the Department of Psychology at the University of Illinois at Urbana–Champaign. She studies the cognitive roots of psychological essentialism and the social origins of environmental attitudes. She is a recipient of the Clean Energy Education Fellowship from the Graduate College at the University of Illinois.

simply a function of the inherent features of the entities involved. The human mind, however, may be prone to ignore this possibility. If the present proposal is correct, people often understand the regularities in their environments as inevitable reflections of the true nature of the world rather than as end points of event chains whose outcomes could have been different.

Consider the color/gender mapping example. Although pink and blue are now unmistakably gendered, during the nineteenth and early twentieth centuries those colors were actually viewed as interchangeable "nursery colors" that symbolized the young age of the children who wore them, not their gender (Paoletti 2012, Ch. 5). What's more, on the occasions when particular colors were suggested for one gender or the other, the recommended mapping was often the *opposite* of what it is today. For instance, in the November 1890 issue of the *Ladies' Home Journal*, readers are advised in no uncertain terms that they should use "blue for girls and pink for boys, when a color is wished" (Hooper 1890). Therefore, the current color/gender mapping (pink for girls, blue for boys) is best explained not by the inherent perceptual properties of pink and blue but rather by the confluence of now-forgotten historical developments (e.g., marketing campaigns by department stores and clothing manufacturers [Paoletti 2012]). And yet, for many people living today, the pink/girl and blue/boy mappings feel natural and inevitable, as if there were something inherently feminine about pink and inherently masculine about blue (e.g., Hurlbert & Ling 2007). The thought of dressing their boys in an all-pink outfit, let alone a pink dress, would make many parents feel uncomfortable and would probably draw vehement protests from the children themselves – at least from those old enough to have detected the relevant regularity (e.g., LoBue & DeLoache 2011).

A similar argument applies to the case of orange juice being consumed for breakfast. Contrary to what our inherence-based intuitions may lead us to believe, the fact that we currently drink OJ for breakfast is largely a matter of historical accident – in particular, an extensive marketing campaign by the California Fruit Growers Exchange, which was saddled with a consistent glut of citrus fruit in the early 1900s and was looking for new ways to market its product (Laszlo 2007, Ch. 7). To illustrate, the *Annual Report of the General Manager of the California Fruit Growers Exchange* for the year 1929 boasts a magazine marketing campaign "with a total of 310,964,842 impressions" that "reached 28½ million homes" and featured, among other things, "orange juice for breakfast" (p. 18). The campaign appears to have been an instant and smashing success, enabling orange growers to sell even portions of their crop that were "formerly considered undesirable." Thus, orange juice went from being a novelty drink to a breakfast staple largely because of concerted efforts by orange producers to make it so, and not because its inherent properties made it an obvious choice for breakfast.

If, as hypothesized, people explain observed patterns mostly in terms of the inherent features of their constituents, this perspective cannot be an automatic consequence of the brute statistical facts. For example, the mere existence of a pattern whereby girls wear pink is not, in and of itself, informative about the reason for such a pattern. In principle, thinking that girls wear pink for reasons
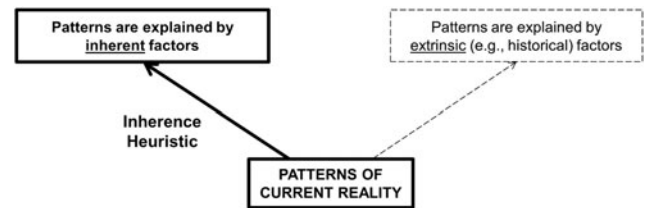


Figure 1.   Schematic representation of the inherence heuristic.

*extrinsic* to both girls and pink (e.g., it's just a convention) would be as legitimate as thinking that girls wear pink because of something *inherent* to pink or girls (e.g., pink is a delicate color; girls have a "hardwired" attraction to pink). Therefore, a consistent preference for thinking that observed patterns are explained by inherent, rather than extrinsic, factors may speak to the rules of thumb that guide how people make sense of the world (see Fig. 1). More specifically, such a preference may speak to the operation of the hypothesized inherence heuristic, which leads people to interpret many broad facts about the world as being the by-products of inherent factors. In this article, we will argue that this heuristic is a pervasive feature of human cognition. At some point or another, humans have reasoned inherently about all sorts of patterns that actually arose from mutable historical forces (e.g., caste systems, child labor, women's confinement to the home) – just as people today tend to explain many of the regularities that structure their lives as the by-products of inherent features (e.g., Jost & Banaji 1994).

The rest of this article is structured as follows. In section 2, we explain how the hypothesized inherence heuristic may work: for example, what sort of mental process it is, what its inputs and outputs are, what other inferences it licenses, and how it may be overcome. In section 3, we argue that the inherence heuristic can provide a unified explanation for a number of disparate psychological phenomena. Then, in section 4, we propose that the inherence heuristic is a necessary ingredient in the process by which humans construct beliefs about the existence of physical, internal essences that define and explain how the natural and social worlds are carved up into kinds (e.g., Gelman 2003; Haslam et al. 2000; Medin & Ortony 1989). Our proposal that the inherence heuristic lays the foundation on which these so-called *essentialist* beliefs are constructed may shed new light on their origins, which are currently something of a mystery. Finally, section 5 clarifies the relationship between our proposal and other hypotheses that seek to account for some of the same phenomena (Prasada & Dillingham 2006; 2009; Strevens 2000), as well as related cognitive biases (e.g., the correspondence bias).

## 2. The inherence heuristic: What is it, and how does it work?

This section spells out how the inherence heuristic is hypothesized to operate. We begin by clarifying our use of the term *heuristic*, whose multiple senses may otherwise obscure what sort of cognitive process we have in mind.

## 2.1. Two types of heuristics: Deliberate and intuitive[2]

The term *heuristic* applies to two distinct classes of mental processes (e.g., Evans 2009; Frederick 2002; Gilovich & Griffin 2002). Some heuristics are deliberate strategies or procedures that we use to simplify complex problems. For example, when looking to buy a new car, one could decide to avoid the trouble of visiting multiple dealerships and to instead purchase the car that has the best online reviews within one's price range. Such voluntary searches for simple solutions are termed *deliberate* heuristics and can be contrasted with heuristics that operate at a more implicit level – *intuitive* heuristics. For example, the choice of a new car may also be swayed by impressions that pop into one's mind spontaneously, without any apparent effort (e.g., a hybrid would be nice). These easy intuitions also help narrow down answers to problems that would otherwise be complex and time-consuming, just as deliberate heuristics do; however, the two operate via different processes. Rather than being the output of a conscious decision-making process aimed at saving effort, the answers suggested by intuitive heuristics are the result of fast implicit processes that are automatically triggered by the problem under consideration (e.g., Frederick 2002; Kahneman 2011).

We now go on to explain the process by which the inherence heuristic is proposed to operate. In doing so, we draw on the compelling account synthesized by Kahneman (2011) out of four decades of empirical research on intuitive heuristics.

## 2.2. How does the inherence heuristic work?

The process underlying the inherence heuristic is set in motion when people seek to explain observed patterns (see Fig. 2). One of the fundamental conclusions of modern psychology is that humans have a powerful drive to make sense of their environments, a drive that prompts them to seek explanations spontaneously, on a routine basis, and from the earliest ages (e.g., Gopnik 1998; Gopnik et al. 2004; Lipton 2004; Murphy & Medin 1985; Premack & Premack 1996; Ross 1977; Weiner 1985). Even infants seem to posit unseen causal mechanisms to explain the evidence gathered from their interactions with the world (e.g., Gweon & Schulz 2011; Saxe et al.

2005; Schulz 2012). Although infants (and laypeople in general) may not approach the task of generating explanations in a terribly systematic and rigorous manner, they nevertheless show a deep-seated motivation to uncover the underlying structure of reality.

Once this explanatory drive is targeted at a particular pattern (e.g., Why do we drink orange juice for breakfast? Why do girls wear pink?), the next stage of the heuristic process is activated (see Fig. 2). Adapting one of Kahneman's (2011) terms, we call this stage the *mental shotgun*: the process of quickly activating any easily accessible information that might be relevant to answering the question at hand (see also Evans 2006; Stanovich 1999; 2011).[3,4] In the case of the inherence heuristic, then, the mental shotgun stage is likely to consist of a fast, shallow search for information that might be applicable (Higgins 1996) to the task of constructing an explanation for the pattern under consideration. In the rare cases when a specific answer is already known, the process terminates here. However, under most circumstances, the shotgun search will culminate not in the retrieval of a stored answer but rather in the generation of an assortment of facts that are potentially relevant to finding an answer.

Although the content generated by the mental shotgun will undoubtedly vary depending on the pattern to be explained, this content may nevertheless be structured along predictable lines. Because the shotgun prioritizes speed and ease of access, on most occasions it will start its search with the entities that happen to be most prominent in our minds at the point when the heuristic process is triggered. In the case of the inherence heuristic, these entities are typically the constituents of the pattern we are trying to explain. For instance, because people are already thinking about OJ and breakfast when they start wondering what explains their pairing, the mental shotgun will probably target those focal objects first. To be more specific, the shotgun is likely to activate any information that it has easy access to regarding these focal objects. What information might this be? Given that an object's representation in semantic memory often consists of information about its stable, inherent characteristics (e.g., McRae & Jones 2013), we expect that the output of the shotgun will correspondingly be dominated by the stable, inherent features of the participants in the relevant pattern (e.g., OJ smells refreshing; breakfast is in the
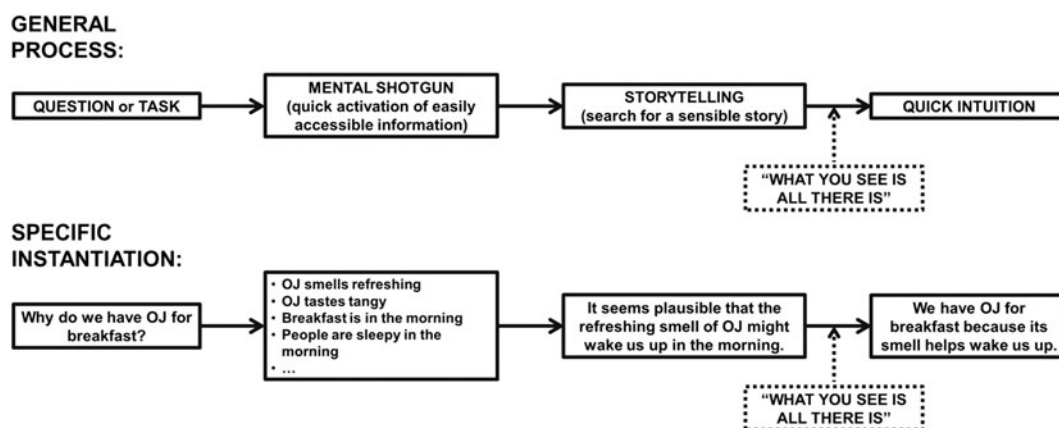


**Figure 2.** The general process involved in generating an intuitive judgment (*top*), and a specific instantiation of this process that leads to an inherence-based explanatory intuition (*bottom*).

morning). Semantic-associative information of this sort is highly accessible to implicit cognitive processes (e.g., Devine 1989; Greenwald et al. 1998; McRae et al. 1997; Rosch & Mervis 1975) and has in fact been implicated in the operation of other intuitive heuristics (e.g., Gilovich et al. 2002; Kahneman 2011; Sloman 1996).

As a side note, this description of the mental shotgun is compatible with the well-established principles that govern the process of knowledge activation, and in particular with Higgins's (1996) evidence of a "salience–applicability–knowledge activation" chain: "Salience…can impact subsequent responses by influencing which features of a stimulus event receive attention, and this in turn will influence which stored knowledge units are likely to be activated in the immediate situation" (p. 158). Because the focal entities of a pattern are typically the most salient at the point when the explanation question is posed, these entities will typically be deemed by the shotgun to provide information that is applicable (or relevant) to its task. This judged applicability will then prompt the shotgun to activate further knowledge about the focal entities – especially easily accessible knowledge about their inherent characteristics, which are tightly bound up with the representation of these entities in semantic memory.

As may already be apparent, we expect that the output of the mental shotgun will typically *fail* to include much information about past circumstances or external events pertinent to the pattern under consideration. One obvious reason for this failure is that such information (about past marketing campaigns, historical events, etc.) might be unknown to most people. However, even if such information were available, it might still not be picked up by the mental shotgun because this information is typically neither salient nor accessible. Unlike the constituents of the pattern to be explained, which loom large at the time when the inherence heuristic is triggered, the past circumstances that may have contributed to this pattern are often no longer in place and may also have no obvious physical connection with the pattern itself (for similar arguments, see Gilbert & Malone 1995). The inconspicuousness of these extrinsic factors makes it likely that they will be overlooked by the shotgun, even if knowledge about them was available. A second, related reason why extrinsic factors may not make it into the output of the mental shotgun is that information about them is often not as *accessible* as the information about inherent features, which is activated and consolidated with every additional exposure to these features (for a discussion of accessibility, see Higgins 1996; Higgins & King 1981). For example, even though somebody may have heard at some point that the color/gender mapping is currently the opposite of what it used to be, this piece of information may, without further consolidation, fail to show up on a quick shotgun search for reasons why girls wear pink.

The heuristics literature describes many examples of similar failures to retrieve relevant, but not very salient or accessible, knowledge. For instance, when told that Tom W. was "of high intelligence" but lacked "true creativity" and had "a need for order and clarity" when he was a senior in high school, people ranked the likelihood that Tom W. is currently a graduate student in computer science much higher than the likelihood that he is a graduate student in the social sciences (Kahneman & Tversky 1973). This typical, intuitive answer overlooks a crucial piece of information – namely, the relative size of the two fields. Mostly everyone knows that graduate students in the social sciences outnumber those in computer science; and yet, once provided with a description of Tom W. as a high schooler, participants quickly called up their stereotypes about computer scientists and made a decision on the basis of this easily accessible information, without retrieving the crucial base-rate information that should have been factored into their responses as well. This example illustrates the intuitive mind's tendency to make use of nothing other than the most salient and accessible information. In the case of the inherence heuristic, this information will often be about the inherent characteristics of the to-be-explained pattern's constituents.

Once the mental shotgun has completed its job, the information generated is handed over to the next stage of heuristic processing. To use another one of Kahneman's (2011) metaphors, this next stage is a *storyteller*, looking to arrange the information at its disposal into a coherent narrative whenever possible (see Fig. 2). Whenever such a narrative emerges out of the assortment of facts called up by the mental shotgun, it then percolates up to working memory in the form of an apparently effortless intuition. Of course, such an intuition only *appears* effortless. It is actually the product of vast amounts of rapid processing that implicit cognitive processes perform behind the scenes.

To reiterate, the pool of facts activated by the mental shotgun for the purpose of generating an explanation for a pattern may often be heavily biased toward the inherent characteristics of that pattern's constituents. As a result, when the storytelling part of the heuristic process takes over and attempts to make sense of the information at its disposal, it will have a rather limited number of options. That is, it will often be forced to construct a story that explains the existence of a pattern in terms of the inherent features of the entities within that pattern rather than in terms of factors external to it. However, the one-sided nature of the information delivered by the mental shotgun is not an impediment to the storytelling process. Quite the contrary – the less information is available, the easier it will be to fit it all into a coherent story (Kahneman 2011). In the case of girls wearing pink, such stories are easy to construct: For example, perhaps girls wear pink because this color is flower-like and delicate – a perfect match for girls' delicate features. Likewise, the case of orange juice being consumed for breakfast can easily be fit into a sensible narrative. As the self-styled "world's undisputed #1 expert on breakfast" speculates in a post on his website, perhaps "the odor of citrus" is "energizing, invigorating and refreshing" and thus "helps wake you up" at the time of day when you need it the most.[5] Again, the storytelling stage has settled on an explanatory story that accounts for the existence of a pattern (we drink OJ in the morning) in terms of the inherent characteristics of the entities in that pattern (it's because OJ has an energizing smell that we drink it in the morning; see Fig. 2). More generally, the outcome of this stage will often be an intuition that the pattern under consideration can be explained by *the inherent features of its constituents*. Even in cases where a specific story fails to coalesce at this point, people may nevertheless be left with a vague sense that the inherent features activated by the mental shotgun will ultimately be sufficient to explain the pattern under

consideration (e.g., some of the typical characteristics of OJ and/or breakfast will ultimately explain why we drink OJ for breakfast, although it remains to be determined which ones).

Considering the fragmentary nature of the information it typically relies on, this heuristic process falls short of any normative standard of inference. Omitting from consideration large chunks of evidence simply because they didn't come to mind right away is guaranteed to introduce bias into the inferential process. The intuitive mind, however, seems to operate by the principle of "what you see is all there is": Any information that is not activated by the mental shotgun is completely ignored for purposes of making a judgment (see Fig. 2; Kahneman 2011). Similar to a jury that issues a verdict after hearing only one side of the argument, intuitive processes such as the inherence heuristic are prone to jump to conclusions based on the scant evidence pulled up by the shotgun. This obliviousness to the possibility that there were relevant facts that didn't immediately come to mind is a defining feature of intuitive processing.[6]

But why would people endorse these heuristic intuitions, especially given the strong possibility of bias? Why would we allow these intuitive processes to hijack our judgments? The consensus answer in the reasoning literature – an answer endorsed by many social psychologists as well (e.g., Anderson et al. 1996; Chen & Chaiken 1999; Gilbert & Malone 1995; Trope & Gaunt 1999) – seems to be that the human mind prefers the path of least resistance. That is, we are often content to allow quick heuristic intuitions to have the final say in circumstances where finding an alternative answer would require controlled, effortful processing. Evans (2006), for example, provided evidence for a *satisficing* principle, which he described as a "fundamental bias" to accept the output of intuitive heuristics "unless there is good reason to give it up" (p. 379). Likewise, Stanovich (2011) concluded that members of our species are *cognitive misers* who have a "strong bias to default to the simplest cognitive mechanism" (p. 29).[7] Thus, given this pervasive tendency to conserve effort, it is not surprising that shallow intuitive heuristics end up shaping so much of our reasoning.

In broad brushstrokes, this is how the inherence heuristic is hypothesized to operate. We might also describe this heuristic process as one of *question substitution*: "When confronted with a difficult question people often answer an easier one instead, usually being unaware of the substitution" (Kahneman & Frederick 2002, p. 53; see also Kahneman 2011). In the case of the inherence heuristic, the difficult question "What explains this pattern?" is often unwittingly answered as if it were the easier question "Which inherent features explain this pattern?" Many other intuitive heuristics can be described as inadvertent substitutions of this sort: When asked to rank probabilities, for example, people end up ranking similarity to a stereotype instead (the representativeness heuristic [Tversky & Kahneman 1983]); when asked to evaluate their overall life satisfaction, people end up evaluating their current mood instead (the mood heuristic [Schwarz & Clore 1983]); and so on. Importantly, in none of these cases are people substituting one question for another purposely, as a deliberate strategy. The intuitive answer simply pops into mind and is often accepted without much suspicion or scrutiny. The same is likely to hold true in the case of the inherence heuristic.

We now go on to address a number of more specific questions about the workings of this heuristic. Among other things, we clarify the notion of an inherent feature; we highlight additional inferences licensed by the output of the heuristic; and we discuss whether the intuitions to which the inherence heuristic gives rise can be blocked or revised, as well as whether there are developmental differences in the extent of its influence.

**2.2.1. What counts as an inherent feature?** We are claiming that the cognitive process that underlies the inherence heuristic (depicted in Fig. 2) often leads to explanatory intuitions couched in terms of inherent features. Here, we provide a brief clarification regarding our use of the term *inherent*. Roughly, our account classifies features as inherent if they can be said to characterize how an entity is constituted. Thus, inherent features tend to be stable, enduring characteristics of the entities in question – characteristics of the sort that might often be associated in semantic memory with those entities (e.g., OJ has a tangy taste; pink is delicate) and thus accessible during the shotgun phase of the heuristic. Note that this is a fairly permissive definition in that it encompasses a wide range of possibilities: Inherent features can be either concrete (e.g., OJ has vitamin C) or abstract and theory-laden (e.g., OJ is healthy), either superficial (e.g., men have facial hair) or deep (e.g., men have a Y chromosome), and so on. Even though this category of inherent factors is broad, there is of course much that falls outside its scope. For example, any aspects that are external to, or removed from, the focal entity do not count as inherent, even if they are strongly associated with it (e.g., OJ comes in half-gallon cartons; OJ is kept in the refrigerator). And, importantly, also falling outside this category are any considerations involving the history of the entity, prior circumstances that affected it, and so on. For example, the fact that pink used to be worn by boys would not qualify as an inherent feature, and neither would the fact that OJ was not consumed on a wide scale before the 1920s.

**2.2.2. Patterns explained via the heuristic are seen as stable and inevitable.** The intuition that observed patterns are explained by the inherent features of their constituents may license additional inferences about these patterns. Of particular interest to us in this section are inferences about the presumed stability and inevitability of patterns explained via the inherence heuristic: If a pattern is understood as being rooted in the very nature of the things that make it up, then it becomes difficult to imagine how this pattern could be otherwise. For example, the intuition that girls wear pink because pink is an inherently delicate color seems, on the face of it, to preclude the possibility that there used to be a time when boys wore pink, or the possibility that pink has begun to dominate girls' clothing only in the very recent past. In other words, the inherence heuristic is likely to cast the girls/pink mapping as a permanent, unavoidable fixture of human life. This perspective contrasts with that promoted by explanations in terms of extrinsic–historical factors (e.g., past events). These explanatory factors specify a starting point to the existence of the relevant patterns, revealing their temporally restricted nature, and highlight the possibility that things could have turned out differently, revealing the contingent nature of these patterns.

### 2.2.3. Variability in the explanatory "stories" used by the heuristic and thus in its output.

As should be clear by now, the output of the inherence heuristic takes diverse forms. Although inherent features figure prominently in this output, the explanatory glue that binds these features together will vary quite a bit. One reason for this variability is that there are often multiple stories that can weave a plausible narrative out of the information pulled up by the mental shotgun. Which of these stories is actually incorporated into a particular heuristic judgment depends on factors such as the context in which the judgment is made and the prior knowledge that is brought to bear on it. Context matters because it can prime certain explanatory frameworks, making them temporarily more accessible to intuitive processing (e.g., Devine 1989; Higgins et al. 1977). For instance, watching an episode of *Grey's Anatomy* might increase the accessibility of biological explanations, as might having a stomachache. Similarly, context can temporarily highlight certain inherent features from among the pool of candidates, thereby influencing who the main protagonists are in the storytelling process (e.g., the features of girls vs. the features of the color pink). Context works on a broader temporal scale as well. For example, one's cultural setting can make certain explanatory constructs *chronically* accessible to heuristic processes, thereby regularly skewing the output of the heuristic in a particular direction. A person's prior knowledge may likewise boost the chronic accessibility of certain classes of explanations. The more extensive one's prior knowledge and use of a construct is, the more frequently this construct may be co-opted on subsequent occasions for the purpose of generating sensible stories (e.g., Bargh et al. 1986; Higgins et al. 1982).

The intuitions generated by the inherence heuristic vary not just within a pattern but *across* patterns as well. The source of this other type of variability is simply that the "stories" that are sensible for one observed pattern are often inadequate for another pattern. Although it may make sense, for instance, to rely on one's theories about disease prevention to explain why OJ is a morning drink (perhaps a dose of vitamin C early in the day is just what our immune systems need to ward off disease), these theories are of little use when considering why girls wear pink or why giraffes are called *giraffes*. Different explanatory notions are relevant to different types of entities in the world, and this basic fact will undoubtedly influence the storytelling step of the inherence heuristic. Depending on the regularity under consideration, the output of the inherence heuristic might appeal to a range of causal frameworks (biological, psychological, physical, etc.), as well as to value-laden, normative notions such as appropriateness, naturalness, or optimality (see especially sects. 3.3 and 3.4). For example, even though one might explain the pairing of orange juice with breakfast in terms of a specific biological mechanism (e.g., the properties of OJ boost immunity for the upcoming day), the pairing of pink with girls might perhaps be better accounted for via an optimality-based story (e.g., the features of girls and pink are an ideal match). In sum, the inherence heuristic makes opportunistic use of a variety of causal and normative explanatory notions in its attempt to cobble together a believable story.

Importantly, despite the substantial variability in the structure of the intuitions generated by the inherence heuristic, they will typically have a crucial aspect in common: They will, more often than not, explain an observed pattern as a function of its constituents' inherent features. This is the hallmark of the proposed heuristic process.

### 2.2.4. Scope of the heuristic: Patterns versus instances.

So far, we have discussed the inherence heuristic as if it pertained exclusively to how people make sense of large-scale patterns (e.g., why so many girls wear pink). This is probably an oversimplification. In all likelihood, attempts to explain general patterns and specific instances (that is, specific events, outcomes, or behaviors) trigger the same intuitive process: The mental shotgun quickly generates some facts; these facts get passed on to the storytelling stage for further processing; and so on. Thus, more comprehensively conceived, the inherence heuristic is the intuitive process that is invoked to explain a *wide range of observations* (both general and specific) and that – due to shortcuts built into its structure – leads to an overreliance on inherence-based explanations.

If the same process is called up for both patterns and instances, then what justifies the focus on patterns? The answer is that the output of this common process may often diverge for the two types of inputs. Although there are good reasons to suspect that people overweigh inherent factors even in the case of single observations (e.g., Gilbert & Malone 1995; Jones & Davis 1965; Jones & Harris 1967), this intuitive process may be *particularly* likely to generate inherence-based explanations when it targets broad patterns, as we argue next. Thus, because the departures from normativity are most dramatic in the case of patterns, we have framed the inherence heuristic as targeting primarily this class of inputs. We will continue with this framing for the rest of the article as well (except for sect. 5.3, where we discuss the correspondence bias), on the understanding that it is a simplification.

Why might inherent explanations be more common for patterns than for instances? One reason for expecting such a difference is that information about applicable extrinsic factors may be somewhat more accessible in the case of specific observations (e.g., Cimpian & Cadena 2010; Cimpian & Erickson 2012; Cimpian & Markman 2009; 2011; see also Kelley 1967; 1973). During our everyday interactions with the world, we witness mostly specific events involving specific entities (objects, people, etc.). As a result, we have extensive opportunities to discover and record in memory the extrinsic causal factors that are relevant to such events (e.g., one's wireless router might stop working after it was accidentally stepped on). Arguably, these opportunities for discovery are less plentiful in the case of broad patterns, especially if these patterns consist of instances widely dispersed over space and time. Moreover, many of the extrinsic factors that are applicable to single events would – if called up to explain broad patterns as well – be hard to fit into a story that makes sense (e.g., accidents are not frequent enough to account for the high rate of interruptions in the wireless signal). As a consequence, our explanatory intuitions may end up relying particularly heavily on inherent factors (e.g., wireless routers are unreliable) in the case of patterns.

### 2.2.5. Blocking and revising the typical output of the heuristic.

Although inherence-based intuitions exert a powerful influence on our understanding, their hold is

not inescapable. We are in fact capable of avoiding, as well as discarding, the inherence-based intuitions typically supplied by the heuristic, even if we may not do so on a regular basis. For example, on the select occasions when the mental shotgun successfully activates information about plausible extrinsic factors, this information may prevent these typical inherence-based inferences from being generated in the first place – that is, these inferences may be *blocked*. Alternatively, if an inherence-based explanation has already been generated, exposure to evidence inconsistent with it may lead to a *revision* of this explanation. Although many observed patterns may be understood initially as due to the very constitution of the things that make up these patterns, the alternative to this view is often within our grasp.

*2.2.5.1. Individual differences.* The alternative to inherent thinking may well be within our grasp, but not everyone is equally likely to reach out for this alternative. In other words, there may be individual differences in people's ability and motivation to overcome the tendency to adopt inherence-based explanations. Extensive evidence for such individual differences has accumulated with respect to intuitive heuristics more generally (e.g., Stanovich 1999; 2011; Stanovich & West 1997; 2000; see also Evans 2003; 2008; Kahneman & Frederick 2002). Extrapolating from this evidence, two dimensions of individual variability may be particularly predictive of the extent to which people rely on the typical (inherence-based) output of the inherence heuristic. The first dimension is *cognitive ability*, as measured by IQ and other similar tests (e.g., Stanovich 1999; 2011). Individuals with higher cognitive abilities possess more efficient systems for retrieving information from long-term memory, as well as more efficient mechanisms for processing the information retrieved (e.g., Cattell 1986; Lohman 1989). As a result, they may be more likely to activate (at the shotgun stage) and use (at the storytelling stage) information other than the highly available inherent features, which may in turn enable these individuals to avoid the usual output of the inherence heuristic.

Another aspect of general cognitive ability that might be relevant here is *cognitive control* – that is, the ability to regulate one's mental processes (e.g., Kane & Engle 2002; Stanovich 2011, Ch. 3). With greater cognitive control, high-ability individuals may be able to limit the influence of the "what you see is all there is" principle – that is, they may be able to avoid jumping to conclusions too quickly, based on too limited a pool of facts. Alternatively, if the inherence heuristic has run its course and an intuition has been generated, individuals with greater cognitive control may also be in a better position to scrutinize, rather than simply go along with, this heuristic intuition. That is, these individuals may be less likely to *satisfice* (Evans 2006) and thus more likely to resist the initial suggestions made by the inherence heuristic.

A second dimension of individual variability that may be predictive of reliance on the typical output of the inherence heuristic is *cognitive style* (e.g., Baron 2005; Evans 2008; Stanovich 1999; 2011). Individuals differ not only in their *ability* to perform complex cognitive computations (as already mentioned) but also in their *inclination* to perform such computations: Some people seek out tasks that challenge their abilities, whereas others gravitate

toward familiar routines; some people relish gray areas, whereas others welcome the certainty of black-and-white opinions; and so on. The more favorably one is disposed toward engaged, open-minded thinking, the less one is susceptible to the influence of intuitive heuristics – even after accounting for cognitive ability per se (e.g., Stanovich 1999; 2011). We expect the same conclusion to hold with respect to the proposed inherence heuristic. Cognitive style could affect the operation of the heuristic at multiple points: for example, by modulating how thorough the information searches are during the shotgun stage, or by influencing how wary people are about going along with "what you see is all there is" (and thus how likely they are to jump to facile conclusions). Some evidence for the claim that open-minded cognitive styles reduce the influence of the inherence heuristic, as well as for the hypothesized negative relationship between cognitive ability and reliance on the inherence heuristic, is presented in section 4.3.

*2.2.5.2. Developmental differences.* The inherence heuristic is a basic cognitive process that operates in largely the same way across development. Nevertheless, the influence of this process may be more extensive in early childhood, narrowing somewhat as children become increasingly able to block and revise the inherence-based intuitions typically generated by the heuristic. Evidence for this claim (namely, that children are particularly prone to understand the patterns of their world via the inherence heuristic) will be discussed throughout the article (see also Cimpian & Steinberg, in press). It is not the case, however, that adults' reasoning is free from the influence of the heuristic. They too reason inherently, and quite often. Our claim here is simply that adults may be in a somewhat better position than children to overcome this way of thinking, as we explain next.

Inherence-based intuitions may be more prevalent in children's thinking for several reasons. First, children's causal-historical knowledge of the world is sparser than that of adults, which means that the facts available to children for purposes of weaving a plausible story are even more biased toward typical characteristics than are the facts usually available to adults. However, as children accumulate more information about the external factors that are relevant to the origins of observed regularities (e.g., historical events, social conventions), they may become increasingly likely either to block the typical output of the heuristic or to revise their original intuitions about an inherent basis for these regularities. For example, although many children might first explain why fire trucks are red by appealing to the inherent features of fire trucks and/or the color red, exposure to evidence of systematic variability in the color of fire trucks (e.g., they are yellow in Hawaii) might lead children to discard this inherent notion in favor of an explanation based on arbitrary social conventions.

Second, developmental differences in reliance on the inherence heuristic may also occur because of developmental differences in cognitive resources. Over the course of childhood, basic cognitive operations, including inhibitory control, undergo tremendous development (e.g., Kail 1991; Williams et al. 1999). Improvements in the infrastructure of cognition are likely to have noticeable effects on the proposed heuristic process by enabling children to perform faster and more comprehensive memory searches, for example, or by enabling them to resist going with the first thought that comes to mind. With age, then, children

may be better able to rein in the influence of intuitive judgments (e.g., Kokis et al. 2002).

However, it is also important to point out that development may not always bring about a reduction in the prevalence of inherence-based intuitions. The process that underlies the inherence heuristic is subject to influence from multiple factors, and some of these factors may offset the contributions of greater knowledge and enhanced cognitive efficiency. For example, adults may be more *motivated* to reach certain inherent conclusions than children are, especially when these conclusions produce desirable psychological effects (e.g., alleviating anxiety; see sect. 3.2). In these cases, motives would steer the course of the heuristic back toward inherence-based intuitions, canceling out the influence of improved cognitive resources. Because of interactions such as these, the trajectory of inherent thinking across development may not always be a simple downward trend.

**2.2.6. Not always wrong.** The inherence heuristic does not invariably lead to error. If certain aspects of the world cluster together reliably (e.g., people wear heavy clothing in the winter), one possibility is that they do so because of something internal to the cluster itself – because some elements of this cluster explain the presence of the other elements (e.g., the winter cold explains the use of heavier clothing). In other words, the inherence heuristic is capable of generating normatively correct judgments, which is also true of intuitive heuristics as a class (e.g., Evans 2009; Stanovich 2011). We should reiterate, however, that the inherence heuristic provides merely an *approximate* solution to the problem of explaining observed patterns. Inherent features are only a subset of the explanatory factors that could be at play, and ignoring the possibility that extrinsic (e.g., historical, social-conventional) factors may also be involved is undoubtedly a source of bias.[8]

### 2.3. Interim summary

The inherence heuristic is a fast, implicit cognitive process that often leads people to explain observed patterns in terms of the inherent features of the entities that constitute these patterns. In turn, these explanations instill a certain perspective on the patterns that are being explained, making them seem stable and inevitable. Although inherent features are typically present in the output of the heuristic, the explanatory structures that bind them together are likely to vary. The storytelling step of the inherence heuristic calls on a broad spectrum of beliefs about causal mechanisms (e.g., biological, psychological), as well as on normative, value-laden notions (e.g., optimality, appropriateness). As a result, there is a fair amount of diversity in the form of the intuitions generated by the heuristic, as subsequent sections will illustrate. The ultimate fate of the intuitions typically resulting from the inherence heuristic is, in principle, also varied – inherence-based intuitions are not obligatorily endorsed. Rather, they can be blocked before they are even generated (e.g., if information about plausible extrinsic–historical factors is activated by the mental shotgun); they can be scrutinized and rejected once they are generated; and they can be revised after they have been endorsed. However, under most ordinary circumstances, we expect that the output of the heuristic

will be accepted as is and will thus influence much of how people understand the world around them.

## 3. The inherence heuristic as a unified explanation for diverse psychological phenomena

We proposed that people tend to make sense of observed patterns in terms of the inherent features of their constituents. Our goal in this section is to provide evidence for this proposal. Most of the evidence reported here was not originally collected to test our claims, so we will thus reinterpret it from our viewpoint. However, by positing the presence of an inherence heuristic, we will be able to highlight and explain the deep similarities across a broad range of findings that may at first appear unrelated. In and of itself, the fact that the inherence heuristic can provide a unified account for such a diversity of psychological phenomena highlights its worth as a scientific hypothesis.

### 3.1. Evidence from directly elicited explanations

Perhaps not surprisingly, several illustrations of the inherence heuristic can be found in research that directly elicited people's explanations for patterns. We begin with an example in which the heuristic leads to reasonably accurate (perhaps even normatively correct) explanations, followed by other examples that make more obvious the bias that the inherence heuristic can introduce into our judgments.

In one series of studies (Kelley 1967; 1973; McArthur 1972; Orvis et al. 1975), college students were asked to explain an event (e.g., why Amy laughed at Beth's joke) in the context of some additional information. Crucially, this additional information suggested either that the event in question was part of a broader pattern or that it was a one-off occurrence. Consistent with our proposal, participants were significantly more likely to invoke inherent features (e.g., aspects of Amy or Beth's personalities) when a pattern was present than when it was not. For example, if the information suggested that many other people laugh at Beth's jokes (a pattern), then the typical explanation was that Amy laughed at Beth's joke because of something inherent about Beth (e.g., she is a funny person) (Orvis et al. 1975). Similarly, if the information suggested that Amy laughs at pretty much any joke (a pattern), then the typical explanation was that Amy laughed at Beth's joke because of something inherent about Amy (e.g., she is a kind person or perhaps easily amused). In contrast, if the information suggested that the event to be explained was unique (e.g., nobody else laughed at that joke, and Amy has never laughed at Beth's jokes before), then the more common explanation appealed to "something about the particular circumstances at the time" (p. 608).[9] (To reiterate an earlier point, people may not fully escape the pull of inherent explanations even when reasoning about isolated behaviors, as suggested by the correspondence bias in person perception [e.g., Jones & Harris 1967]. However, Kelley and colleagues' evidence does seem to suggest that inherent explanations may be less prevalent for one-off behaviors.)

In simplified versions of these studies, even 4-year-old children appealed to inherent features to explain patterns of behavior (e.g., DiVitto & McArthur 1978; Higgins & Bryant 1982; Ruble et al. 1979; Seiver et al. 2013), which

is consistent with the claim that the inherence heuristic is available early in development. Evidence for inherence-based reasoning in children may also be found in the research on praise and criticism, as we explain next. Providing feedback for a specific outcome via a statement that generalizes beyond that outcome, and thus implies the existence of a pattern, often prompts the recipients of this feedback to draw conclusions about inherent qualities such as talent or aptitude (e.g., Cimpian et al. 2007; Kamins & Dweck 1999; Mueller & Dweck 1998). In turn, these inferences can undermine motivation and performance in later situations because they lead to anxiety about the level of these supposed inherent qualities (e.g., "How much talent do I have?") and about how failure might reflect on them. To illustrate, 4-year-old children who were praised for their success on a drawing task with the general praise "You are a good drawer" reacted more negatively to subsequent mistakes (e.g., displayed more negative emotions, were quicker to give up) than children who were given more specific praise, such as "You did a good job drawing" (Cimpian et al. 2007; see also Cimpian 2010; Cimpian et al. 2012). Presumably, this difference emerged because children's explanations diverged in the two cases – that is, because the praise that framed children's success as part of a broader pattern of successes was more likely to promote explanations in terms of inherent qualities than the specific praise was. This interpretation is admittedly one step removed from the data, since children were not actually asked to identify the source of their performance. Nevertheless, it seems quite likely that diverging explanations were at the root of children's behavior, especially in light of the evidence that beliefs about fixed talents (also known as "entity" theories [see Dweck 1999; 2006]) often give rise to the sort of helpless behavior that was observed after the general praise.

The conclusion that children explain regularities in inherent terms receives additional support from research on explanations for novel facts about categories and individuals (Cimpian & Cadena 2010; Cimpian & Erickson 2012; Cimpian & Markman 2009; 2011). Across a number of studies, 4- and 5-year-old children provided significantly more inherence-based explanations for the category-wide facts than for the individual-specific facts, although these two sets of facts were otherwise perfectly matched in content. For example, unfamiliar abilities that were said to characterize an entire social group (e.g., "boys are really good at a game called *gorp*") were often inferred to be due to inherent traits (e.g., "boys are smart") (Cimpian & Erickson 2012; Cimpian & Markman 2011). In contrast, children were significantly less likely to rely on inherent features when the same unfamiliar abilities were said to characterize a single individual (e.g., "there's a boy who is really good at a game called *gorp*"); children explained the abilities of individuals more often in terms of the historical process by which these abilities arose (e.g., "he practiced a lot").[10] Although such a process would probably be responsible for the abilities of a group as well (e.g., Romanians excel at gymnastics because of their rigorous training, not because they are naturally limber), children ignored this possibility and concluded instead that these patterns were due to the inherent features of their constituents. Thus, children's responses reveal the bias that the inherence heuristic often introduces into the process of generating explanations. Finally, these results also highlight

how broadly the inherence heuristic applies, skewing people's interpretation of novel, as well as familiar, patterns. Even though children had no specific knowledge about gorp or any other activity we asked them about, their explanations nevertheless gravitated toward the (presumed) inherent properties of these activities and of the social groups said to be good at them.

The evidence reviewed in this section provides support for the claim that people make sense of broad facts about the world via an inherence heuristic. In the next section, we turn to a different literature and argue that the inherence heuristic may be part of the reason why people tend to see the current state of affairs in their society as fair and legitimate.

### 3.2. Evidence from research on system justification tendencies

Many people, including many who are disenfranchised, believe that the sociopolitical systems they are part of are "good, fair, natural, desirable, and even inevitable" (Jost et al. 2004, p. 887; see also Jost & Banaji 1994; Jost & Hunyady 2005; Kay & Friesen 2011; Kay et al. 2009; Lerner 1980; Liviatan & Jost 2011; Samuelson & Zeckhauser 1988). What is at the root of this striking belief? Considerable evidence suggests that the main culprit here may be a tendency to explain current societal arrangements as being due to the inherent characteristics of the people who make up the various strata of the social hierarchy (e.g., Gabennesch 1990; Hoffman & Hurst 1990; Jost 2001; Jost & Burgess 2000; Napier et al. 2006). For example, people who are poor must be so because they lack some of the traits needed to succeed in life (e.g., they must not be very hardworking, motivated, or intelligent) and not as a result of adverse circumstances or unfair biases in the opportunity structure of one's society. Or, to take another example, the fact that men hold political power must be due to the "fact" that they possess the intellectual and personality characteristics needed to lead (e.g., they are rational and level-headed) rather than to historical events that happened to favor patriarchal societies. This tendency to justify existing societal patterns by casting them as the inevitable products of inherent features is quite robust insofar as it is present even when people reason about unfamiliar circumstances. In one series of studies, for example, participants were told about two groups of aliens "who inhabit a distant planet" and who fulfill different roles in that planet's society (e.g., Orinthians tend to work "in the free-enterprise sector," whereas Ackmians tend to work "in the research/educational sector") (Hoffman & Hurst 1990). As predicted by the inherence heuristic account, participants thought that this pattern was in place because of the inherent qualities of the aliens in the two groups, and more specifically because of the perfect match between the aliens' characteristics and the stereotypical requirements of their respective occupations (e.g., Orinthians must be more self-assured and sociable, whereas Ackmians must be more intellectual and solitary).

These system-legitimizing inferences fit well with our proposal, in that they reveal a tendency to explain current societal patterns by appeal to the inherent features of the constituents of these patterns. There is, however, a dimension to these inferences that goes beyond what our account

can explain in its current form. Specifically, people are *motivated* to adopt system-legitimizing beliefs of this sort in order to alleviate discomfort or anxiety about their place in society and in order to achieve a sense of meaning, order, and predictability in their lives (e.g., Jost & Hunyady 2005; Kay & Friesen 2011; Kay et al. 2009). If current societal structures are fair and legitimate, then one's place in them is exactly where it should be, and thus there is no need to feel frustrated (if one is disadvantaged) or guilty (if one is privileged).

According to our proposal, people's use of the inherence heuristic is not driven by the need to reduce negative affect. Rather, this heuristic is a cognitive shortcut that operates seamlessly behind the scenes, shaping our explanatory intuitions about observed patterns. It is possible, however, that motives could exert a subtle influence at several points in the heuristic process. For example, the motive to defend the status quo might lead people to keep certain information (such as negative stereotypes of disadvantaged groups) active in memory and thus easily accessible to a variety of cognitive processes (e.g., Anderson et al. 1996; Higgins 1996; Higgins & King 1981). The increased accessibility of this information might then make it more likely to be retrieved by the mental shotgun and therefore more likely to be used during the subsequent storytelling process (see Fig. 2), thereby tilting the output of the heuristic toward intuitions that legitimize the political status quo. System justification motives may also influence how people handle the output of this heuristic process. When the inherence heuristic gives rise to intuitions that satisfy defense motives, people may be even more likely than usual to satisfice – to fail to scrutinize the output of the heuristic – and perhaps also more eager to protect this output from subsequent revision (e.g., Anderson et al. 1996; Chen & Chaiken 1999; Kruglanski 1996; Pyszczynski & Greenberg 1987). In light of these considerations, the evidence that system-justifying beliefs have a motivated component does not negate the possibility that these beliefs stem from the operation of the inherence heuristic. It is still quite possible that the proposed heuristic serves as the cognitive bedrock on which people can build a motivated system-justifying ideology.

Our claim that system justification tendencies are made possible by the inherence heuristic casts this well-established phenomenon in a new light. If we are correct, system-justifying intuitions are a subset of the inherence-based intuitions that people generate to make sense of all sorts of observed patterns, many of which have nothing to do with their social or political status. Although inherence-based intuitions that legitimize the status quo may be particularly common (in part because of their palliative effects), as well as particularly consequential for how we relate to one another, the process by which they are generated is no different than the process that leads people to conclude, say, that orange juice is consumed for breakfast because of its inherent properties (e.g., its energizing smell, its vitamin C content). By identifying the cognitive underpinnings of system-justifying ideologies, this provocative hypothesis also highlights the deep similarities between this and other important social-cognitive phenomena such the correspondence bias (sect. 5.3) and essentialism (sect. 4), both of which may be rooted in the same heuristic process.

Finally, this new perspective on system justification can be used to derive surprising new predictions. For instance,

robust system-justifying tendencies should be found even in children's thinking, especially since the inherence heuristic may hold more sway at earlier ages (sect. 2.2.5.2; for some preliminary support for this prediction, see Baron & Banaji 2009). These inferences wouldn't serve the same palliative function for children as they do for adults (because children aren't typically as concerned about their place in society), but they should nevertheless be present early in development if they are in fact an output of the inherence heuristic. We might also predict that endorsement of system-justifying intuitions should be affected by the same variables that are hypothesized to affect endorsement of inherence-based intuitions more generally (e.g., cognitive ability, executive functioning, cognitive style). More broadly, our proposal might be extended to provide new insights into the cognitive foundations of political attitudes. For example, given that political conservatism is associated with both system-justifying tendencies and ambiguity-intolerant cognitive styles (e.g., Jost et al. 2003; Matthews et al. 2009), we might predict that, relative to liberals, conservatives would show greater reliance on the intuitions supplied by the inherence heuristic, even outside the realm of politics. To the extent that predictions such as these are supported by future work, they would further demonstrate the value of understanding system-justifying intuitions as emerging from the operation of the inherence heuristic.

### 3.3. Evidence from research on nominal realism

Consistent with the proposed tendency to explain observed patterns as being rooted in the characteristics of their constituents, children all over the world appear to believe that words are, to some extent, inherently compatible with the objects to which they refer. This belief is known as *nominal realism*. Of course, the vocabulary of a language is in reality the product of a long chain of historical events rather than a reflection of the inherent features of the things named. There is, for example, nothing sun-like about the English word *sun*. The pattern whereby speakers of English use this word to refer to the bright object in the sky is an arbitrary convention that arose over time to coordinate communication about that object.[11] And yet, our understanding of words as mere conventions is surprisingly fragile, especially in childhood. Piaget (1929/1967), for example, asked children whether it would have been possible for things in the world (e.g., the sun) to have different names than they currently do (e.g., *moon*). Children as old as 9 denied this possibility, citing reasons such as "Because the sun can't change, it can't become smaller" or "Because it's nothing else but the sun, it couldn't have another name" (p. 81). These responses suggest that, even for children well into their school years, names are not conventional patterns imposed on the world for the purpose of coordinating communication. Rather, names are believed to match the features of their referents – they are believed to "come from the things themselves" (p. 86). Gelman (2003) encountered the same belief in a conversation with a 9-year-old who was convinced that our names for dinosaur species could not be any different than they currently are because "they found the fossils and stuff" (p. 183) – again, as if names were inherent in the things named. Moving beyond the anecdotal, consistent evidence of childhood nominal realism was found in experimental research as well (e.g., Brook 1970;

Lockhart et al. 1977; Osherson & Markman 1975; but see Rosenblum & Pinker 1983).

Claims of nominal realism, however, might seem at odds with the developmental evidence suggesting that children understand the conventionality of linguistic forms (for a recent review, see Diesendruck & Markson 2011). This conflict is only apparent. What the evidence in this literature shows is that children expect words to be widely known to people in their linguistic community (e.g., Clark 1988; Diesendruck & Markson 2001; Graham et al. 2006). However, an expectation that words are widely known is quite different from, and does not entail, a belief that words are arbitrarily linked with their referents. After all, children expect many *nonarbitrary* facts about the world to be widely known as well (Cimpian & Scott 2012). It is thus entirely possible that children simultaneously (a) believe that words are conventional in the sense of *widely known* and (b) do not believe that words are conventional in the sense of *arbitrarily chosen*.

To return to our argument, it may not be just children who see words and objects as inherently, rather than arbitrarily, linked. Under certain circumstances, remnants of this belief can be identified in adults as well. A first piece of evidence for this idea is found in anthropological studies of nonliterate societies. Horton (1967), for example, claimed that people in these societies tend to "see a unique and intimate link between words and things" (p. 159; see also Frazer 1890/1958; Murdock et al. 1978). For many of them, words seem to be "inextricably bound up with reality" (p. 159), even to the point where a word might be used to stand in for the object it refers to (as in magic rituals, for instance). Assuming that Horton's claims are valid (cf. Scribner & Cole 1981; Tambiah 1968), the intuitions he describes here have all the telltale signs of the inherence heuristic. As people attempt to make sense of observed linguistic patterns (e.g., the pairing of *sun* with the sun), they might often arrive at the conclusion that these patterns are particularly appropriate: Perhaps, say, the sound of the word *sun* matches the features of its referent (e.g., its heat, its brightness) at some level. If people inferred such a match between the inherent features of a name and those of its referent, it would not be unreasonable for them to make the further assumption that an object's name could be used to stand in for the object itself (just as a photograph might be used to stand in for the object in it).

Traces of the belief in an inherent link between words and objects lurk beneath the surface in literate societies as well (see, for example, Rozin et al. 1986; 1990). As an illustration, we briefly describe the results of some of our own research on this question (Sutherland & Cimpian, in preparation). In these studies, our goals were to determine (a) whether we could in fact find traces of nominal realism in a sample of American adults and, if so, (b) whether individual differences in nominal realism would correlate with individual differences in more general inherence-based reasoning, as the present account might predict. To assess nominal realism, we created a scale that probed participants' ideas about the nature of word–referent links. For example, one item asked subjects whether "there is something particularly appropriate" about the name *giraffe*, or whether we "could have just as easily called this animal something else." Participants indicated their answers on Likert-type scales. (Responses to the

item above, for example, were made on a scale from 1 = "This name is particularly appropriate" to 7 = "We could have easily called this animal something else.") We also asked participants to justify their answers to these scale items. Next, to assess broader inherence-based reasoning, we created an inherence heuristic scale that required participants to rate their agreement with statements such as "It seems natural that engagement rings typically have diamonds," "It seems right that black is the color associated with funerals," or "There are good reasons why orange juice is typically consumed for breakfast."[12] None of these items were about names or language, so the two scales were nonoverlapping in content.

The results of these studies were consistent with both of our predictions. First, traces of nominal realism were indeed detectable in our sample of American adults (all of whom were college students in this study). Consider, for example, the strikingly realist intuitions expressed in some of their open-ended justifications: One person explained why she thought the name *giraffe* was particularly appropriate by saying, "The animal has a long neck, so it is a *giraffe*"; another wrote, "I think the word *giraffe* captures the height well with the use of the double *f*s." Even when subjects answered that another name would have also been suitable, which was the more typical response, their justifications occasionally revealed an undertone of realism, as when a subject wrote that "we could have called it something that goes along with its features." The implicit assumption behind these answers seems to be that a name (even at the level of orthography, as with the double *f*s) is rooted in the inherent features of its referent.

Our second prediction was also supported: Participants' nominal realism scores were significantly correlated with their scores on the inherence heuristic scale, $r(124) = 0.31$, $p < 0.001$. This result is consistent with the argument that nominal realism is a specific instantiation of a more general propensity to explain the patterns of reality in terms of inherent factors. Moreover, the correlation between participants' nominal realism and inherence heuristic scores remained significant, $r(122) = 0.30$, $p = 0.001$, even when we partialed out the variance attributable to two cognitive variables that could have conceivably influenced both of these scores: the need for certainty and orderliness (known as *need for closure* [Kruglanski et al. 1993]) and the desire to engage in effortful cognitive activity (known as *need for cognition* [Cacioppo et al. 1984]). (The link between inherent reasoning and nominal realism was replicated in two additional studies in which we asked participants about artifact names [e.g., *bottle*] instead of animal names [e.g., *giraffe*] and included a number of other control measures.)

Interestingly, the inherence-based intuitions identified in our studies seem to extend to idiomatic expressions as well. That is, people often have the intuition that idioms (e.g., *taking the bull by the horns*) are transparently rooted in the features of their referents (Keysar & Bly 1995; 1999). To many of us, for example, the idiomatic expression *taking the bull by the horns* seems uniquely suited to describe the responsible act of assuming control of a difficult situation. In reality, however, the mapping between form and meaning in idioms such as these is far more arbitrary than our (arguably inherent or realist) intuitions would lead us to believe. *Taking the bull by the horns* could have just as easily meant the opposite of what it does today – that

is, it could have referred to the irresponsible act of taking unnecessary risks such as grabbing a raging bull by its horns. The fact that we use it to describe a responsible rather than an irresponsible act is probably nothing more than an accident of history – an accident that we nevertheless understand in inherent terms. Thus, for patterns of idiom use just as much as for patterns of word use, people tend to overestimate the extent to which language matches the inherent features of the world.[13]

To summarize, the phenomena described in this section support the argument that people tend to explain observed patterns by appealing to the inherent features of the things that instantiate these patterns. More specifically, people seem prone to explain word–object mappings in terms of the match between the features of the words and the features of the objects to which they refer. As a side note, the evidence reviewed in this section also illustrates the explanatory promiscuity of the inherence heuristic (in particular, its storytelling step), insofar as people's intuitive explanations for word–object links seem to rely on normative, rather than causal, considerations (see sect. 2.2.3). That is, the intuition that names are ideally suited for their referents because they provide an optimal match for their referents' features does its explanatory work via value-laden judgments about what is suitable and optimal.

### 3.4. Evidence from research on is–ought errors

As suggested in the preceding section, the intuitions due to the inherence heuristic are often normative: They explain via value judgments about what is appropriate, ideal, optimal, or right. In this section, we argue that – when used as explanations for patterns of human behavior – these heuristic intuitions are likely to introduce considerable bias into *moral reasoning*.

The sort of bias we have in mind was first identified by David Hume (1740/2000), who famously cautioned against placing undue weight on the patterns of current reality (i.e., what *is*) in drawing conclusions about moral truths (i.e., what *ought to be*). Imagine, for example, that you were a member of a social group whose members routinely hanged those who were suspected of theft. What factors should you consider in deciding whether such punishment is morally justified? According to Hume, the current prevalence of this punishment cannot be the only basis for your decision; additional premises are needed to justify conclusions about the ethics of an act other than its being typical (see also Sober 1994; Wilson et al. 2003). In the terminology of the inherence heuristic proposal, you would be in error to proceed directly from observing a pattern (e.g., hanging is a common punishment for theft) to inferring that this pattern is due to an optimal match between the features of its constituents (e.g., the severity of hanging matches the seriousness of theft-related offenses, and thus hanging is appropriate punishment for thieves).[14]

Despite Hume's centuries-old warning, is-to-ought transitions remain prevalent in our thinking, as would be expected if they were the output of an effortless heuristic shortcut. Indeed, empirical studies suggest that modern-day undergraduates are very much inclined to take *what is* as an unerring guide to *what ought to be* (Eidelman et al. 2009; Friedrich 2005; 2010; Friedrich et al. 1989). Across a number of studies, for example, participants

rated currently observed patterns (e.g., a certain set of graduation requirements) as "good," "right," and "the way things ought to be" relative to hypothetical alternatives (Eidelman et al. 2009). Participants did so even though it was made clear that the change to these hypothetical alternatives could have some benefits and would occur at negligible cost. Here as well, then, we may see evidence for the heuristic intuition that the world is as it is, and should not be any different, because of the inherent features of the things in it: The current graduation requirements are in place because they are inherently appropriate, perhaps even optimal, and not because they are the outcome of a contingent historical process that could have just as easily followed a different path.

Consistent with the developmental predictions of our account (see sect. 2.2.5.2), children's ability to keep *is* separate from *ought* may be even more fragile than adults'. This is not to say that children imbue all social patterns with moral force: Young children are often able to differentiate among social regularities, realizing for example that some of these exist due to context-specific conventions (e.g., at our school, we sit in a circle during story time) whereas others are in place due to context-independent considerations of justice and welfare (e.g., we refrain from hitting others and stealing their toys) (for a recent review, see Helwig & Turiel 2011). However, even if children understand *some* conventional patterns as being conventional, there is also evidence of developmental change in this respect. That is, in comparison with older children and adults, younger children seem to still reify a good number of the patterns that structure their everyday lives, viewing these patterns as inherently appropriate rather than as dependent on social consensus (see Gabennesch 1990; Gelman & Kalish 1993; Kalish 1998; Kalish & Lawson 2008; Lockhart et al. 1977; Shweder et al. 1987; Smetana et al. 1993; but see Helwig et al. 1990).

To illustrate, Kalish (1998) presented preschoolers with stories whose protagonists were said to *want* to violate a social or physical regularity. For instance, one boy wanted to wear a dress (social violation), and another wanted to turn into a bird (physical violation). The participating children were then asked if the protagonists would actually perform these actions (e.g., "Will he really...?"). In a separate study, children were also asked to make predictions about the actions of protagonists who not only wanted to violate these regularities but were also ignorant of them. For instance, the boy in the story wanted to wear a dress and, not being "from around here," didn't know that boys don't wear dresses. As we would predict, the younger preschoolers in these studies appeared to treat violations of social and physical regularities as equally improbable. That is, 3½-year-olds often denied that the characters in these stories would violate the relevant regularities, even when these concerned (what are in reality) social conventions, and even when the potential violator was both ignorant of these conventions and motivated to act counter to them (see also Lockhart et al. 1977). These responses are compatible with the claim that many social patterns hold considerable force for young children, as would be expected if they interpreted these patterns via the proposed heuristic. For instance, if the features of dresses make them inherently appropriate for girls, then of course any boy – even one who has never seen a dress before – would immediately recognize that wearing such a thing would be wrong.

By offering a hypothesis about the cognitive underpinnings of is–ought reasoning, our account opens up new possibilities for experimental investigation of this important phenomenon. To date, discussion of is–ought inferences has been confined mostly to philosophy, with relatively little attention paid to their characteristics as a psychological phenomenon (for a few exceptions, see Eidelman et al. 2009; Friedrich 2005; 2010; Friedrich et al. 1989). If this type of reasoning were indeed an output of the inherence heuristic process (see Fig. 2), then its likelihood of occurring should be sensitive to the same factors that modulate endorsement of inherence-based intuitions more generally (e.g., cognitive ability, cognitive style; see sect. 2.2.5). Moreover, is–ought reasoning should pattern with other phenomena that are thought to stem from this heuristic, such as nominal realism and system justification. The presence of structural commonalities between these phenomena might also suggest that conclusions drawn about one should extend to the others. For instance, it is possible that is–ought reasoning would be influenced by defense motives, just as system-justifying reasoning seems to be. Is–ought beliefs (whose core implication is that the patterns of our world are exactly how they should be) may be adopted more often by people with a chronic need for such comforting thoughts, as well as in situations that temporarily heighten this need. Importantly, the influence of motives should not be limited to is–ought inferences regarding one's place in society, which also fall under the purview of system justification; rather, this influence should be felt with respect to a much broader range of inferences about what is right or appropriate. As these considerations illustrate, the inherence heuristic proposal is a generative source of insights into the diverse phenomena that are hypothesized to stem from this heuristic.

### 3.5. Interim summary

We set out to provide evidence for our proposal that people often understand the patterns they detect in the world as being due to the inherent features of the entities that instantiate these patterns. We found evidence for this proposal in a diverse set of findings that – although not originally collected to test our account – were all parsimoniously interpretable in terms of a single underlying process: namely, the inherence heuristic. By revealing the shared underpinnings of well-established phenomena that have previously been regarded as distinct, our proposal also provides a fresh perspective on these phenomena and a rich basis for further empirical research.

## 4. The inherence heuristic as a precursor to essentialist thinking

A major goal of this article is to outline the possibility that the inherence heuristic lays the cognitive foundation for the emergence of psychological essentialism. Specifically, we will propose that the essentialist framework emerges as an elaboration of the earlier, and more inchoate, intuitions supplied by the inherence heuristic. Before detailing this proposal, however, we explain what psychological essentialism is, highlighting its importance both as a basic feature of the human mind and as a potential source of prejudice and conflict.

### 4.1. What is psychological essentialism?

Psychological essentialism is an influential hypothesis regarding the structure of natural and social kind concepts (e.g., Gelman 2003; Haslam et al. 2000; Medin & Ortony 1989; Newman & Keil 2008; Prentice & Miller 2007; Rhodes & Gelman 2009; see also Bloom 2000; 2004). According to this hypothesis, people routinely assume that different kinds of entities in the world (e.g., lions, tulips, gay people) are underlain by different unseen, but real and causally potent, *essences*.[15] In other words, people have consistent (yet somewhat vague) intuitions about the causal source of the features that characterize various kinds: a certain internal, physical, microstructural *je ne sais quoi* that is unique to each kind and that invariably causes its members to display the full complement of typical features (barring accidents). For example, people might reason that lions look the way they do (e.g., tawny) and behave the way they do (e.g., hunting gazelles) because they all possess lion DNA or something of the sort, which is also the reason they are lions rather than donkeys or frogs. Similarly, people might reason that water looks the way it does (e.g., clear) and behaves the way it does (e.g., freezing at 32°F) because of some special fact about its microscopic structure, which is also what makes it water rather than ethanol or granite. The presence of this implicit belief in essences is supported by an impressive range of findings and has been documented in children as young as age 4 (for a comprehensive review, see Gelman 2003).

On a more practical level, because essentialist beliefs imply that the differences between categories are deep and immutable, they are also a likely source of intergroup prejudice and hostility of the sort that is common in human societies (e.g., Dar-Nimrod & Heine 2011; Haslam & Whelan, 2008; Prentice & Miller 2007). For example, the belief that social groups have essences has been associated with stronger endorsement of stereotypes (e.g., Bastian & Haslam 2006; Levy et al. 1998; see also Yzerbyt et al. 2001), greater acceptance of racial inequities (Williams & Eberhardt 2008), and more blatant prejudice toward members of minority groups (e.g., Keller 2005; see also Haslam et al. 2002).

### 4.2. The origins of essentialism: Preliminary considerations

Given the centrality of essentialist beliefs to human cognition and their role in fostering prejudice, conflict, and inequality, it seems particularly important that we understand the origins of essentialism itself. How does the human tendency to essentialize arise? To start, the consensus among scientists seems to be that the categories essentialized by humans do not in actuality have essences (e.g., Dupré 1981; Leslie 2013; Maglo 2011; Needham 2000; 2011; Sober 1994). Thus, essentialist assumptions cannot come from the world per se, for the simple reason that our world is not carved at the joints by real essences. Nor are these assumptions merely a product of socialization. Detailed examinations of parent–child conversations have found little in the way of overt essentialist talk to children (Gelman et al. 1998; 2004). At best, parental input provides subtle, indirect hints (e.g., kind-referring generic nouns) that would not be sufficient to create an essentialist

mind-set from whole cloth (e.g., Cimpian 2013; Gelman et al. 2010; Rhodes et al. 2012). It seems most probable, then, that essentialism is an endogenous feature of the human mind: either an innate module that evolved as a product of natural selection (e.g., Atran 1993; 1998; Gil-White 2001; Pinker 1994, Ch. 13) or the result of a highly constrained developmental process whereby children construct essentialist beliefs out of simpler cognitive parts, many of which may themselves be innate (Gelman 2003; but see also Bloom 2000, Ch. 6; Carey 1996; Kornblith 1995, Ch. 5).

Because the latter appears to be the more promising of the two alternatives, as argued persuasively by Gelman (2003, Ch. 11), we adopt it here as a working hypothesis.[16] On this account, the essentialist framework is elaborated out of developmentally prior abilities and biases (e.g., the ability to distinguish surface appearances from underlying reality) over the first few years of life. These preexisting cognitive capacities are clearly not essentialism per se, but they each supply distinct functional content that is necessary to the development of a mature essentialist framework. To clarify, the claim here is not that essentialism is just a collection of separate processes, without any psychological reality of its own. Although the origins of essentialism lie in prior capacities, the essentialist framework transcends their sum.

### 4.3. The inherence heuristic as a precursor to essentialism

Our proposal is that *psychological essentialism emerges as an elaboration of the intuitions supplied by the inherence heuristic*. To reiterate, the inherence heuristic often gives rise to intuitions that explain observed patterns in terms of the inherent features of their constituents. Which inherent features play the starring role in these intuitions depends to a large extent on the storytelling step of the heuristic. If, for example, people make sense of the fact that orange juice is consumed for breakfast via an explanatory story about the rousing properties of citrus smells, then these smells will be highlighted as a key source of this pattern. Alternatively, if the story used is one about the immunity-boosting properties of vitamin C, then it will be the nutritional content of orange juice that is highlighted as a source of the pattern. Or again, if no plausible stories are identified at this step, the heuristic will output a more inchoate sense that some to-be-determined combination of inherent features is responsible in some to-be-determined fashion for the observed pattern (an output that may be refined at a later time if applicable explanations are identified; see sect. 2.2). The sparser the explanatory options available to the storytelling step of the heuristic are, the more often it will be forced to produce the vague, "to-be-determined" sort of output. Given that the store of explanatory options is more limited for children than for adults, the output of the inherence heuristic may be particularly open-ended in early childhood – and thus more often in need of further elaboration.

On our proposal, the early output of the inherence heuristic is gradually elaborated into a full-blown essentialist stance over the first few years of life. This early output is not *itself* essentialism. Unlike essentialist beliefs, for instance, inherence-based intuitions do not necessarily pinpoint a single inherent feature as the source of the multiple regularities observed within a kind. Similarly, the inherent features present in the early output of the heuristic do not necessarily fit the description of an essence – they may not be thought of as necessarily physical, internal, and nonobvious, like essences are. Despite such differences in content, however, these inherence-based intuitions make it possible for children to *eventually* formulate essentialist beliefs. In other words, although children may at first understand the patterns that characterize natural and social kinds via vague intuitions couched in terms of to-be-determined inherent factors, they may soon begin to fix some of the free parameters in this equation, coming to more specific conclusions about the nature of these inherent factors. The end result of such an elaboration process, insofar as it occurs, may be the belief in an essence. More precisely, this elaboration process may lead to the belief that a *single* inherent feature explains the regularities observed within a kind – most likely a *physical* feature that is *internal* to the members of this kind and thus *nonobvious*. For example, children may eventually come to believe that *lions* and *water* display the features they do because of some unique feature of their internal structures or, later in development, perhaps because of something about their DNA and chemical composition, respectively.

How does the transition to a more essence-like notion of inherent cause occur? Are there explanatory stories or beliefs that could guide this elaboration process, and are these beliefs developmentally prior to essentialism per se? We describe two possibilities for how this process might unfold, focusing in particular on essentialism about living kinds. (Essentialist beliefs about other domains are discussed later in this section.)

First, the initial output of the inherence heuristic may be elaborated on the basis of children's ideas about *insides*. Although young children know relatively few specifics about the inside contents of objects (e.g., Gottfried & Gelman 2005; Simons & Keil 1995), they do seem to possess an abstract expectation that insides are *causally powerful* (e.g., Gelman & Wellman 1991; Newman et al. 2008), and particularly so for animate kinds (Gelman 1990). For instance, 14-month-old infants expect novel animate objects with similar insides to exhibit similar patterns of movement and sound, even when their surface markings differ (Newman et al. 2008; see also Welder & Graham 2006). Thus, even before the onset of essentialism (Gelman 2003), children's reasoning about animate objects includes an abstract assumption that internal parts are causally responsible for these objects' behavior. Note that this early causal bias is not tantamount to a belief in a category essence: Initially, the insides are seen simply as an enabler of movement and behavior, not as the identity-conferring source of all properties that characterize the relevant category. Even so, such a skeletal causal bias might be sufficient to guide the initial output of the inherence heuristic toward the notion of an internal, nonobvious (i.e., essence-like) inherent cause. Whereas children might start out agnostic concerning which inherent features explain why the members of, say, the *lion* kind exhibit all the regularities they do (e.g., being tawny, having manes, hunting gazelles), the presence of an early bias to endow internal parts with causal powers might soon enable children to refine this view: Whatever it is that makes lions what they are, it probably has to do with their insides. Moreover, since lions display a set of regularities that is distinct from that of

any other kind, this internal *je ne sais quoi* that lions have must also be distinct from that of any other kind.[17]

Children's early reasoning about animate objects also features the construct of *internal energy*, which could provide an alternative means of elaborating the initial output of the inherence heuristic. That is, young children seem to assume that animate beings possess some sort of vital substance that is responsible for maintaining life (e.g., Hatano & Inagaki 1994; Morris et al. 2000) and for providing the causal force behind movement and growth (e.g., Gelman & Gottfried 1996; Gottfried & Gelman 2005). The roots of this causal belief may be present even in infancy. For example, 5- to 6-month-olds' reasoning about the behavior of self-propelled objects suggests that infants attribute to these objects a source of internal energy that makes it possible for them to behave independently of external physical forces and thus, say, resist movement when pushed or even float when released in midair without adequate support (Luo et al. 2009). Again, this early notion of an internal, nonobvious source of energy that is causally responsible for the observable behaviors of animate objects does not amount to a belief in category essences. Yet, this notion may still provide an adequate means of refining the inchoate early output of the inherence heuristic. It is clear, though, that the essentialist belief constructed via this refinement process must transcend children's original ideas about internal energy, both because the essence is *specific to a kind* rather than being a general causal force and because the essence is causally responsible not only for dynamic features such as movement and growth but for *static* features as well.

To review, at least two sorts of explanatory beliefs (beliefs about insides and beliefs about internal energy) could in principle guide children from an inchoate sense that the patterns encountered within a living kind are due to some combination of inherent factors to more concrete ideas about these factors. The result of this elaboration process, we propose, could function as the category essence. It is important to note, however, that on either of the alternative paths sketched above, the end point will not be terribly concrete in an absolute sense; it will simply be more concrete relative to the starting point. That is, we suspect that children typically narrow down the possibilities only to the level of a *type* of inherent factor: something physical, internal, nonobvious. It is rare even for adults to go beyond this level of (non)specificity in their thinking about essences, which is why essences are characterized more accurately as *essence placeholders* (e.g., Gelman 2003; Medin & Ortony 1989): Although people believe that essences exist, they often cannot say much about them, perhaps other than that they are physical, internal entities that are in principle knowable at a detailed level (e.g., by the appropriate experts). Thus, the elaboration process that we proposed to account for the origin of essentialism probably has as its terminus point an essence placeholder rather than a specific essence.

Although we have so far focused mostly on essentialism about living kinds, the present proposal can also account for an important characteristic of essentialist beliefs – namely, the substantial cross-domain variability in their content (e.g., Bloom 2000; Gelman 2003; Haslam et al. 2000; Newman & Keil 2008). Such variability is a natural by-product of the mechanism hypothesized here. If essentialist beliefs are constructed by elaborating the open-ended output of the inherence heuristic on the basis of plausible explanatory stories, and if these stories differ depending on the entities under consideration (e.g., lions vs. New Yorkers vs. corkscrews), it follows that the end product of this elaboration process (that is, essentialist beliefs) should differ as well. We flesh out this argument via several examples.

First, children seem to endorse a more localized view of the essence for animate than for inanimate natural kinds (e.g., lions vs. gold) (see Newman & Keil, 2008, Experiment 2). Such differences in essentialist beliefs are consistent with our elaboration account insofar as there are also differences in the explanatory stories available to children in the two domains. As evidence of such domain differences, consider children's beliefs about the causal powers of insides. These beliefs appear to pertain specifically to the domain of animates (e.g., Gelman 1990; Gelman & Wellman 1991), with no evidence that children believe inanimate natural kinds such as *gold* to also have discrete internal parts or organs with causal properties. Thus, if the hypothesized elaboration process relies on different explanatory beliefs in these two domains, this process will most likely lead to different assumptions about where the essence is located. With respect to animates, the notion of insides as a causal source may license the inference that the essence is found where the vital insides are rather than being distributed homogeneously throughout. Depending on the explanatory notions used with respect to inanimates, children may either remain agnostic about the location of the essence or else adopt a distributed view.

Second, variability in the explanatory stories used during the elaboration process may also lead to variability in the content of people's beliefs about *what* the essence is (not just *where* it is). Some evidence for such variability is found within the social domain (e.g., Haslam et al. 2000). Although certain social groups (e.g., men, African-Americans) are thought to share an underlying biological essence, essentialist beliefs about other groups (e.g., New Yorkers, hippies) do not include a biological component but nevertheless cast these groups as coherent and homogeneous. Speculatively, it is possible that explanatory stories involving biological factors are deemed more plausible for social groups that have distinguishing biological markers (e.g., men) than for groups that do not have such markers (e.g., New Yorkers). Put another way, if some of the regularities that set a group apart from others include observable biological features (e.g., facial hair, deep voices), then a story that appeals to underlying biological factors unique to that group will seem more plausible and may thus be used more often in the elaboration process. Although the regularities detected within the social groups that do not have such observable biological markers are also likely to be interpreted via the proposed heuristic, the inherent features assumed to be responsible for these regularities may remain unspecified or may be fleshed out in psychological, rather than biological, terms (e.g., the regularities in New Yorkers' behavior may be due to something about their inherent personality traits).

Third, variability in the explanatory stories available to the elaboration process may also explain whether essentialist beliefs about a given category emerge at all or whether this process takes the initial output of the inherence heuristic in an altogether different direction. To take a salient

example, artifact concepts do not exhibit the hallmarks of psychological essentialism – or at least not the canonical forms of it (Gelman 2003; but see Bloom 1996; 2000; Hood & Bloom 2008). That is, artifacts (e.g., coins) are not typically assumed to possess an internal, microstructural category essence that is the causal source of their features. And yet, the regularities pertaining to artifact categories are, we suspect, interpreted via the same heuristic process that is used to make sense of the regularities pertaining to natural and social kinds. Why, then, do children's inherence-based intuitions about artifacts follow a different course from their intuitions about living things and natural objects? On our account, there are two main reasons for this divergence. The first reason is that there are no artifact counterparts to the early causal notions about insides and internal energy, which pertain exclusively to animate beings. As a result, there are fewer opportunities (if any) in this domain for children to flesh out the early output of the heuristic into more specific beliefs about causally powerful, internal essences. The second reason for this divergence is that the early output of the inherence heuristic with respect to many artifact categories is often overridden rather than elaborated. Children may *initially* understand the regularities observed within many familiar artifact categories in inherent terms, much as they do the regularities observed within natural and social kinds. For instance, young children may believe that coins are round, that fire trucks are red, and so on, because of some to-be-determined combination of inherent factors (Cimpian & Steinberg, in press). Relatively early on, however, children also start to understand the crucial role that people (an *extrinsic* cause) play in the emergence of these regularities (e.g., Gelman & Bloom 2000; Gelman & Gottfried 1996; Rhodes & Gelman 2009). As evidence accumulates in favor of this explanatory story, the appeal of earlier inherence-based intuitions may diminish to the point where children may actually abandon them. (As argued before, the intuitions supplied by the inherence heuristic are open to later revision if information comes along that is contradictory to these initial intuitions.) In sum, we argue that essentialism does not take hold in the artifact domain in part because (a) there are no plausible candidates for fleshing out the initial, open-ended output of the inherent heuristic into beliefs about causal essences, and (b) the initial output of the inherence heuristic is often revised or overridden on the basis of children's early-developing appreciation for the extrinsic sources of artifact properties.

In the final part of this section, we lay out a few predictions that follow uniquely from the proposal that the inherence heuristic provides the foundation for psychological essentialism. To begin, if the inherence heuristic is a precursor to essentialism, it would have to be available to children before the first signs of essentialism emerge. The developmental evidence available so far (see sects. 3.1, 3.3, and 3.4) suggests that children as young as age 4 exhibit reasoning patterns that bear the signature of the inherence heuristic. However, since 4 is also the age when children show the first clear signs of essentialist thought (Gelman 2003), evidence is needed about the extent to which children younger than 4 rely on the inherence heuristic. In addition, to minimize the possibility that the explanations of these younger children would stem from essentialism per se, it seems best to explore their inherent reasoning with

respect to domains that are not typically thought to fall under the scope of essentialism even when it does eventually emerge (e.g., artifacts). For example, do 3-year-olds think that fire trucks are red for inherent reasons and not because people decided at some point that they should be red? Do 3-year-olds think that coins have always been and will always be round, as would be expected if they understood this pattern in inherent terms (see sect. 2.2.2)? If inherence-based reasoning of this sort were present in 3-year-olds, and we predict it should be, then it could plausibly serve as a foundation for later essentialist thought.

Beyond simply being present early on, the inherence heuristic should play a direct role in the subsequent emergence of essentialism. One way of examining this link would be to test whether individual differences in inherent thinking at age 3 correlate with individual differences in later essentialist thinking (say, at age 4). If essentialism is elaborated out of the intuitions generated by the inherence heuristic, as we argue here, then the strength of children's inherence-based intuitions at earlier ages should be predictive of the strength of their essentialist beliefs at later ages (but not vice versa). In fact, a similar prediction could be made with respect to the levels of essentialism observed in *adults*. There is considerable variability in the extent to which adults endorse essentialist beliefs, particularly those about social groups (e.g., Bastian & Haslam 2006; Keller 2005; Williams & Eberhardt 2008). According to our account, the people who are particularly likely to endorse essentialist beliefs should also be the people who are particularly likely to go along with, rather than scrutinize, the output of the inherence heuristic.

This last prediction was explored in a recent study from our lab (Salomon & Cimpian 2014). To measure reliance on the inherence heuristic, we used the scale described earlier, in section 3.3. This scale required participants to rate their agreement with a series of statements such as "There are good reasons orange juice is typically consumed for breakfast" and "It seems natural to use red in a traffic light to mean *stop*." To measure endorsement of essentialist views, we used a fairly standard scale of the sort devised by Haslam et al. (2000). The scale required participants to rate the extent to which social groups such as girls or Catholics are cohesive, have genetically determined traits, are rooted in an underlying reality, and so on. We designed these two scales so that there was minimal overlap in content between them. This feature of the study afforded a strong test of our prediction: It might seem counterintuitive to expect that people's beliefs about, say, whether it is natural to use red in a traffic light to mean *stop* would predict their beliefs about, say, whether girls are a cohesive group whose traits are determined by their genes. We also asked participants to fill out nine additional measures that enabled us to control for potentially confounding cognitive, personality, and ideological dimensions. For instance, a critic could argue that high levels of both essentialist and inherence-based reasoning might simply co-occur in people who are low in cognitive ability or whose cognitive style leads them to avoid to ambiguity and seek simple, black-or-white answers instead. The nine control scales were divided into three sets of three scales each, with each set administered to approximately a third of the participants (Ns for the three subsamples ranged from 105 to 112; total N = 323).

The results were in line with our prediction. Participants' reliance on the inherence heuristic strongly predicted their

endorsement of essentialist views about social groups, $r$ (321)=0.39, $p < 0.001$. Moreover, participants' inherence heuristic scores remained a significant predictor of their essentialism scores even when the nine cognitive and personality dimensions used as controls were taken into account, $rs > 0.33$, $ps < 0.001$. The unique predictive power of our scale reveals an intimate link between inherent thinking and endorsement of essentialist beliefs.

These results also provide insight into the factors that moderate the influence of the inherence heuristic on people's judgments. In section 2.2.5.1, we predicted that the tendency to go along with the easy intuitions supplied by the inherence heuristic should depend both on cognitive *ability* and on cognitive *style*. That is, people with higher cognitive abilities should be less likely to rely on the heuristic, as should be people whose thinking is more engaged and open-minded. With respect to cognitive ability, we indeed found that higher scores on this dimension (as measured by an abbreviated Raven's Progressive Matrices test [Raven 1960]) were associated with lower scores on the inherence heuristic scale, $r(103) = -0.28$, $p = 0.005$. Our predictions were confirmed with respect to cognitive style as well, insofar as participants' inherence heuristic scores were (a) negatively correlated with their scores on the Need for Cognition Scale (Cacioppo et al. 1984), which assesses the motivation to engage in effortful cognitive activity, $r(104) = -0.36$, $p < 0.001$, and (b) positively correlated with their scores on the Need for Closure Scale (Kruglanski et al. 1993), which assesses the motivation to arrive at simple, unambiguous judgments, $r(104) = 0.18$, $p = 0.056$. These results support the claim that the inherence heuristic provides a quick and easy means of making sense of observed patterns – a first guess that may be more easily overcome by individuals with extra cognitive resources or more reflective thinking styles.

To review, our proposal that essentialism emerges as an elaboration of the output of the inherence heuristic leads to several novel predictions. These predictions, some of which are already receiving empirical support, are also unique to our account. As a point of comparison, consider the prominent view that essentialism is an innate cognitive module that (a) was originally favored by natural selection because it "allows people to perceive and predict many important properties that link together the members of a biological species" (Atran 1998, p. 550; see also Pinker 1994, Ch. 13) and that (b) was co-opted, later in our evolutionary history, for reasoning about social categories such as ethnic groups, which shared certain characteristics with biological species (Gil-White 2001). As currently described, the modularity account provides little reason to expect that young children would have a broad tendency to view the patterns of their world, regardless of what sorts of entities they involve, in inherent terms. It also provides little reason to expect that this tendency to reason inherently would precede, and predict, the emergence of essentialism about natural and social kinds (the putative targets of the innate essentialist module). In addition to facilitating new insights into the roots of essentialist thinking, the present proposal can account quite naturally for some of the established findings in the literature, such as the domain differences in the content of essentialist thought. This is, again, an area where existing proposals struggle. For example, people's tendency to essentialize nonliving natural kinds such as gold or social groups such as

New Yorkers is difficult to accommodate within the framework of a modularity account because these sorts of categories are not species-like in any straightforward way and thus fail to meet the input conditions of the hypothesized essentialist module.

### 4.4. Interim summary

This section outlined a new perspective on the process that leads to the emergence of essentialist views. Briefly, we proposed that the essentialist framework is constructed as a means of elaborating the inchoate intuitions supplied by the inherence heuristic. Although it is beyond doubt that there is theoretical and empirical work left to be done, the argument laid out here nevertheless holds considerable promise because it spells out a detailed, testable mechanistic model for the development of essentialism.

## 5. Distinguishing the inherence heuristic from related proposals and processes

This section is intended to clarify the scope and contribution of our proposal. To do so, we first differentiate it from two other hypotheses (concerning so-called K-laws and formal explanations) that seek to explain some of the same phenomena. We then go on to discuss the relation between the inherence heuristic and the correspondence bias, a cognitive process that leads to an overreliance on inherent explanations of human behavior. As we argue below, the correspondence bias may be viewed as a specific instantiation of the inherence heuristic.

### 5.1. K-laws

Strevens (2000; 2001) argued that people endorse what he termed *K-laws* (where the *K* stands for *kind*). According to such K-laws, membership in a natural kind is causally responsible for the observable features displayed by the members of that kind. For example, a tiger has stripes because something about being a tiger caused it to have stripes.

To begin, there are notable differences in the content of K-law-based and inherence-based explanations, both in terms of the explanatory entities that figure in the two types of explanations and in terms of the explanatory relations they use. On the first point, inherent features always do the explanatory work in the output of the heuristic. In contrast, at least early in development, K-laws are neutral about what aspects of category membership are causally efficacious (and thus explanatory). That is, children "might have no opinion about what does the causing" or might even think that "it is just a brute fact about the world" that membership in a category gives rise to category-typical features (Strevens 2000, p. 154). Thus, the K-law account makes no commitment to features of any sort serving as explanations for observed patterns. There are also differences in the explanatory relations used by the two accounts. The inherence heuristic is opportunistic: It relies on a variety of explanatory relations, both causal and normative (e.g., about what is optimal or appropriate). In contrast, K-laws are by definition causal, so they are able to explain only via causal relations. In sum, the explanatory intuitions due to K-laws are likely to differ on

multiple dimensions from the output of the inherence heuristic.

Also note that our proposal spells out in detail how the inherence heuristic operates, including why it so often leads to explanatory intuitions that rely on inherent rather than extrinsic factors. In contrast, Strevens provides no account of the process by which children come to endorse K-laws. Thus, even if the two proposals were equally successful in accounting for the phenomena they were intended to explain (but see Ahn et al. 2001; Meyer et al. 2013), the greater degree of mechanistic specificity in the inherence heuristic account may make it more fruitful as a working hypothesis.

Finally, there is a difference in the scope of the two accounts. As should be clear given the range of evidence reviewed in this article, the inherence heuristic is invoked to explain much more than the observable features of natural kinds, which is the explicit scope of K-laws. Insofar as the inherence heuristic proposal is actually successful in accounting for this broader range of evidence, its superior coverage – along with considerations of parsimony – would recommend it over the narrower K-law proposal.

### 5.2. Formal explanations of k-properties

According to Prasada and Dillingham (2006; 2009), people represent the relation between a kind and some of its properties as a part–whole relation: That is, some properties of a kind – dubbed *k-properties* – are represented as an *aspect* of being that kind of thing (e.g., having stripes is an aspect of, or a part of, being a tiger). Importantly for our purposes here, Prasada and Dillingham further argued that the part-whole relations between k-properties and their kinds license so-called *formal* explanations – that is, explanations that account for the presence of a k-property simply by appealing to the kind itself (e.g., tigers have stripes because they are tigers).

Prasada and Dillingham's proposal diverges from ours in many of the same respects that Strevens's did. For instance, there are clear differences in the content of the explanations involved. Most obviously, the proposed heuristic does not generate formal explanations of the sort Prasada and Dillingham describe; rather, its explanations rely on inherent features. One circumstance in which the two types of explanations might appear to overlap is when the heuristic's storytelling stage fails to converge on a plausible story. Although the more inchoate intuitions that emerge on these occasions sound roughly similar to both K-laws and formal explanations (e.g., girls like pink because of something about girls), the resemblance is superficial. Even when the heuristic produces this sort of vague output, the underlying assumption still is that some to-be-determined features are accounting for the relevant pattern, not the kind itself. Another major difference in the content of the two types of explanations is that inherence-based explanations are often *causal*, whereas formal explanations cannot be causal, by definition (Prasada & Dillingham 2006; 2009). However, the part–whole relations hypothesized by Prasada and Dillingham do license *normative* expectations for k-properties (e.g., if having stripes is part of being a tiger, then a tiger should have stripes), so the possibility of generating normative content is a characteristic that the inherence heuristic and the formal explanation accounts have in common.

It is also important to point out the differences in the scope of the two accounts. The inherence heuristic can be invoked to explain pretty much any observed pattern, whereas Prasada and Dillingham's formal explanations concern only those features of a kind that are represented as aspects of being that kind (namely, k-properties). In addition, even though there may be a role for noncausal, formal explanations in everyday cognition, people's reasoning about the world seems to rely heavily on causal explanations from the youngest of ages (e.g., Ahn et al. 2001; Gelman 2003; Medin & Ortony 1989; Schulz 2012; Strevens 2000). Thus, the inherence heuristic proposal is better equipped to provide a successful account of how children and adults make sense of their natural and social environments.

### 5.3. Correspondence bias

The term *correspondence bias* refers to a well-documented tendency to explain other individuals' behaviors (e.g., she is fidgeting) in terms of corresponding dispositional traits (e.g., she is an anxious person) rather than situational pressures (e.g., it's finals week) (Gilbert & Malone 1995; Jones & Harris 1967; Ross et al. 1977; Trope & Gaunt 2000; for recent reviews, see Gawronski 2004; Trope & Gaunt 2007). The clear parallels between this phenomenon and the inherence heuristic emerge, we argue, because the two share an underlying mechanism: That is, the intuition that observed behavior is due to stable psychological dispositions may stem from the same heuristic process that leads people to conclude that observed patterns are due to the inherent features of their constituents. This claim is consistent with the more comprehensive definition of the inherence heuristic we advocated for earlier (sect. 2.2.4), on which the heuristic provides an intuitive means of explaining specific instances and general patterns alike. Although we have focused mostly on patterns so far, the correspondence bias reveals that people's judgments veer away from normative rationality – and toward inherence – even in the case of specific instances.

Why would this be? With respect to behavior, one key reason is that extrinsic factors such as situations or norms are seldom on display in the actual behavior to be explained (Gilbert & Malone 1995). Although people do realize at some level that external pressures shape behavior (e.g., Gawronski 2004), there is little in the behavior itself to remind observers of these pressures. Therefore, a fast, ballistic process such as the inherence heuristic may often overlook these distal forces that influence behavior without leaving any visible clues or reminders of their influence. Instead, the heuristic process may end up (over-) relying on information about the actors themselves, since the actors are more transparently associated with the behavior to be explained than the relevant external constraints are. In particular, the actors' psychological characteristics (e.g., anxiousness) provide the heuristic with plausible causal stories for observed behavior and may thus figure prominently in people's intuitions – especially when the mental shotgun fails to activate thoughts about distal constraints. From our perspective, then, it may be reasonable to view the correspondence bias as emerging from the use of the inherence heuristic to explain particular instances of human behavior. (In fact, others have also argued that the bias is a result of heuristic, rather than systematic, processing [e.g., Trope & Gaunt 2007].)

The inherence heuristic framework we proposed here can accommodate not only the basic existence of a correspondence bias but also some of the nuances of this phenomenon. We list a few relevant findings and briefly point to the elements of our model that might account for them. First, the correspondence bias is more pronounced when participants' cognitive resources are taxed (e.g., Gilbert et al. 1988). As discussed in section 2.2.5.1, limitations in cognitive resources are also known to lead to greater reliance on intuitive heuristics (e.g., De Neys 2006; Stanovich & West 1997; 2000). Second, the correspondence bias is reduced when thoughts of situational constraints are made more accessible, either temporarily via priming (e.g., Trope & Gaunt 2000) or chronically via exposure to cultural contexts that emphasize the power of context in shaping human behavior (e.g., Morris & Peng 1994; but see Krull et al. 1999). On our proposal as well, the outcome of the heuristic's storytelling step is determined in part by the relative accessibility of the explanatory notions at its disposal (sect. 2.2.3). Factors that activate, either temporarily or chronically, thoughts of external causes should lead to a proportional decrease in inherence-based outputs. Third, the magnitude of the correspondence bias is modulated by motivational states. For example, observers who are assigned the goal of assessing the *situation* do not overuse dispositional explanations of behavior (e.g., Krull 1993). This conclusion is compatible with our inherence heuristic proposal, insofar as goals and motives can exert considerable influence over the course taken by the heuristic process (sect. 3.2).

These three examples, while not doing full justice to the richness of the literature on the correspondence bias, nevertheless highlight the plausibility of understanding this bias as a specific instantiation of a broader inherence heuristic. Looking to the future, this perspective on the correspondence bias also creates opportunities for cross-fertilization. In particular, the theoretical and methodological insights that have accumulated over 50 years of research on this bias in person perception could be used to guide new research on phenomena such as essentialism and system justification, which are relative newcomers to the psychology scene.

### 5.4. Interim summary

Our goal in this section was to clarify the relationships between the inherence heuristic and other accounts (K-laws and formal explanations) as well as other psychological tendencies (the correspondence bias) that might be thought to give rise to some of the same intuitions. In clarifying these relationships, we have also highlighted the unique theoretical contribution of the inherence heuristic account.

## 6. Conclusion

We proposed that human reasoning relies on an inherence heuristic, an implicit cognitive process that often leads people to explain observed patterns in terms of the inherent features of their constituents. This heuristic was argued to give rise to a diverse collection of intuitions and beliefs. Consider, for instance, the common belief that the political status quo is fair and natural. From the

perspective of our account, the prevalence of this belief is due, at least in part, to the fact that people make sense of observed societal patterns (e.g., income disparities between groups) in terms of the inherent characteristics of the people that instantiate these patterns (e.g., their intelligence, their motivation; see sect. 3.2 on system justification). Or, to bring up another earlier example, the heuristic tendency to explain observed patterns in inherent terms may also be at the root of the linguistic intuition that words provide a suitable match to the features of their referents (see sect. 3.3 on nominal realism). Along the same lines, we argued that the inherence heuristic is a key reason why people reify observed patterns, imbuing them with moral force (see sect. 3.4 on is–ought errors). Arguments such as these led us to the conclusion that the inherence heuristic is a powerful and pervasive cognitive process. Additionally, these arguments brought to light the deep commonalities among important phenomena that had previously been considered unrelated. By uncovering such commonalities, we were able to offer new insights into the sources of these phenomena and to generate new directions for future research. Some of these directions are already being pursued in our lab, and promising results were summarized at several points in the article to underscore the generative nature of our account.

The other major goal of this article, aside from introducing the inherence heuristic, was to propose that the heuristic serves as a precursor to psychological essentialism. This belief in hidden essences is a central feature of human concepts of natural and social kinds – a feature that has also been linked to phenomena such as stereotyping, prejudice, and intergroup conflict. Building on the arguments in the first part of the article, we hypothesized that the origin of this important aspect of human psychology can be traced back to children's reliance on the inherence heuristic. That is, essentialism was proposed to emerge as an elaboration of earlier, and less specific, intuitions generated by the inherence heuristic. The hypothesized link between the inherence heuristic and the subsequent development of psychological essentialism was able to account for well-established findings (e.g., domain differences in the content of essentialist beliefs), and it led to several unique predictions as well (e.g., relationships between essentialism proper and more general inherent thinking about other aspects of the world). In sum, our proposal spelled out a plausible mechanistic model of the development of essentialism, a model that has the additional advantage of revealing the connections between essentialism and other inherence-based phenomena (e.g., system justification, correspondence bias).

Together, the arguments and evidence in this article help uncover a fundamental new aspect of human psychology. The inherence heuristic is a powerful cognitive process that provides an easy, but also bias-prone, means of explaining why the world is as it is.

Marjorie Rhodes, Joe Robinson-Cimpian, Glenn Roisman, Brian Ross, Kristin Shutts, and the members of the Cognitive Development Lab and the Psychology of Religion, Agency, and Morality Lab at the University of Illinois at Urbana–Champaign.

## NOTES

**1.** The inherence heuristic operates only over the patterns or regularities that our minds represent as such, having either detected their presence in the world or inferred their existence via inductive generalizations (e.g., Gelman & Markman 1986; Markman 1989). The further question of which patterns are extracted from the world and which are not is complex and falls outside the scope of the current proposal.

**2.** For readers who are familiar with the dual-process literature, we intend the distinction between *deliberate* and *intuitive* processes to map roughly onto the distinction between System 2 and System 1 (e.g., Stanovich 1999) or type 2 and type 1 processes (e.g., Evans 2008). We chose not to use these other terms in the article both because their meaning is intuitively opaque and because they have misleading implications. (In particular, System 1 and System 2 are not really systems per se but rather heterogeneous collections of processes with loosely similar properties [e.g., Evans 2008; Stanovich 2011].)

**3.** This initial component has close counterparts in other theories of intuitive processing, although it usually goes by other, less colorful, names. For example, Stanovich (2011) describes a *fundamental computational bias* to "contextualize a problem with as much prior knowledge as is easily accessible" (p. 44; see also Stanovich 1999, Ch. 7). Similarly, Evans's (2006) *relevance principle* refers to "the powerful tendency to contextualize all problems with reference to prior knowledge elicited by contextual cues and the current goals that are being pursued" (p. 380).

**4.** Kahneman (2011) uses the term *mental shotgun* to refer more generally to the intuitive mind's tendency to perform a broad range of quick computations, often many more than needed, in response to a question or task. Our usage preserves some of this meaning but is also tailored to the specific process of the inherence heuristic.

**5.** http://www.mrbreakfast.com/ask.asp?askid=26 (accessed May 19, 2013).

**6.** Because intuitive heuristics use just the information that happened to be retrieved and ignore other, potentially contradictory, information that might nevertheless be available upon more careful reflection, they have been argued to introduce an element of *confirmation bias* into human judgment (Evans 2006; Kahneman 2011; for work on the confirmation bias, see Nickerson 1998; Snyder & Swann 1978; Wason 1960). Importantly, however, these heuristics do not reduce to the confirmation bias. For example, the inherence heuristic supplies explanations, whereas the confirmation bias does not – rather, the bias (in the form of the "what you see is all there is" principle) simply influences what sorts of evidence the heuristic process considers when formulating explanations.

**7.** The term *cognitive miser* has had a long tradition in social psychology as well (Fiske & Taylor 1984), although social psychologists typically use the term in a slightly different sense – as a strategic, deliberate choice to employ the simplest possible cognitive mechanisms (e.g., Bargh 1999; Gilovich & Griffin 2002) rather than as a fundamental (but defeasible) bias to default to such mechanisms.

**8.** The distinction between inherent and extrinsic explanatory factors is not always clear-cut. Some social conventions, for example, might themselves have an inherent basis (e.g., sneezing into the crook of one's arm is in fact an effective way to prevent the spread of germs). The existence of such ambiguous cases, however, does not negate the possibility that people routinely overestimate the importance of inherent explanatory factors.

**9.** Our interpretation of these responses as stemming from the operation of a fast heuristic diverges from the theoretical model that motivated Kelley et al.'s studies (Kelley 1967; 1973; Kelley

& Michela 1980). According to Kelley's model, the explanation process is analogous to the formal procedures for hypothesis testing and data analysis used by empirical scientists and is thus likely to be slow, effortful, and deliberate (see Fiske & Taylor 1984). However, there is no evidence that participants in the studies reported here employed such effortful analytic strategies. Thus, Kelley et al.'s evidence might also be accounted for by the operation of a fast intuitive heuristic that leads people to the same conclusions via a different process.

**10.** Given the research on general vs. specific praise (e.g., Cimpian et al. 2007), we speculate that inherent explanations would have been even more infrequent in these studies if we had asked children to explain a single outcome involving an individual (e.g., "there's a boy who did really well…") instead of a regularity about an individual (e.g., "there's a boy who is really good…").

**11.** Despite some traces of nonarbitrariness in word–object mappings (such as onomatopoeia [e.g., *cuckoo*] or the *bouba/kiki* phenomenon [Maurer et al. 2006; Ramachandran & Hubbard 2001]), the vast majority of words do not in fact resemble or mimic the features of their referents in any meaningful way.

**12.** The inherence heuristic scale consisted of 15 items. In a sample of 323 participants, this scale displayed satisfactory internal consistency (Cronbach's alpha=0.85, mean item–total correlation=0.48). In addition, a confirmatory factor analysis provided support for the hypothesis that it measures a unidimensional construct – namely, reliance on the output of the inherence heuristic.

**13.** An argument could be made that people also understand *grammatical gender* in inherent terms (Boroditsky et al. 2003). In reality, there is no inherent reason why certain objects have names of a certain grammatical gender. For example, it's an arbitrary, contingent fact that the name for the sun is feminine in German (*die Sonne*) – after all, the sun is masculine in Spanish (*el sol*). Nevertheless, people often seem to assume that the grammatical gender of a noun matches the inherent features of the noun's referent (e.g., that *Sonne* is feminine because there is something feminine about the sun).

**14.** We point interested readers to Roberto Unger's (e.g., 1983) critique of the modern legal system, a critique whose main thrust is similar to Hume's. Unger argues that, although laws often do little more than codify the current realities of one's society, people (including lawyers and law scholars) often view them as stemming from transcendent (moral, political, etc.) principles. From a psychological perspective, this tendency to "sanctify the actual" (p. 571) by believing that laws reflect an absolute reality rather than the contingencies of one's time may be another instance of the inherence heuristic at work.

**15.** To be clear, psychological essentialism is a claim about the mind, not about the world. It posits that people believe in the existence of essences, not that essences actually exist in the world.

**16.** Briefly, Gelman's (2003) argument rests on two types of considerations. First, she argues that current versions of the modularity proposal cannot explain the full range of evidence regarding the development and scope of essentialist reasoning. Second, she argues that, given the existence of cognitive capacities that are developmentally prior and functionally related to essentialism, any proposal that suggests essentialism is an entirely separate module that does not draw on these existing capacities seems unparsimonious.

**17.** The question might also arise of whether the elaboration process proposed here leads to expectations about a single essence like inherent cause or multiple such causes for each category. Assuming that children make use of the same, most relevant and accessible, explanatory story to refine many of their inherence-based intuitions within a category, it is likely that they will also end up pinpointing the same cause (e.g., internal parts). The resulting single-cause structure is, in fact, typical of essentialist reasoning (e.g., Ahn et al. 2001).

# Open Peer Commentary

## Is the inherence heuristic simply WEIRD?

Andrew Scott Baron

*Department of Psychology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.*
abaron@psych.ubc.ca
http://www.psych.ubc.ca/~abaron

**Abstract:** Although many studies suggest that children and adults focus more on internal causes rather than situational causes to explain observed patterns, such findings may be more limited to WEIRD populations (*w*estern, *e*ducated, *i*ndustrialized, *r*ich, *d*emocratic) samples. Evidence from cross-cultural studies may point to several distinct attribution mechanisms with their culturally specific deployment reflecting both a developmental achievement as well as a possible signal of group boundaries.

Cimpian & Salomon (C&S) argue that human reasoning is endowed with a cognitive bias that emphasizes inherence as a primary causal mechanism to explain observed phenomena (e.g., why gold is yellow, why orange juice is served at breakfast, why European-Americans disproportionately hold America's economic wealth relative to other ethnic groups). The inherence heuristic is argued to be domain-general and developmentally primitive. Although C&S do not specifically formulate an evolutionary argument, they do identify key properties of this bias that are consistent with a more nativist claim about the origins of this heuristic. Although a broad array of data from social, cognitive, and developmental studies of inductive and causal reasoning can be accounted for by the proposed inherence heuristic framework, there is insufficient evidence to support the strong claim that this reasoning bias is itself innate, or even a human universal.

Researchers have recently noted that psychological studies often draw on WEIRD (*w*estern, *e*ducated, *i*ndustrialized, *r*ich, *d*emocratic) samples that can reason quite differently compared with more traditional societies (Henrich et al. 2010). Most studies of induction and causal attribution, especially among developmental populations, have focused on these WEIRD samples. According to C&S, although cross-cultural differences in the output of the inherence heuristic may vary substantially, the basic processes outlined in their attribution model should apply universally. However, there may be good reason to doubt this claim.

An inherence bias in attribution is ironic, considering humans are unique among all creatures on Earth in that they have powerful mechanisms evolved for cultural transmission – the capacity both to be influenced by and to shape others in their local community (Csibra & Gergely 2011; Herrmann et al. 2007; Richerson & Boyd 2005). As such, anthropologists and psychologists have pointed to the importance of culture as both a mechanism for transmission and a mechanism for explanation (Henrich & Broesch 2011; Henrich & McElreath 2003; Miyamoto & Kitayama 2002; Morris & Peng 1994; Norenzayan et al. 2002; Richerson & Boyd 2005). To this end, cultural and evolutionary psychologists have argued for multiple mechanisms available to humans to facilitate causal reasoning (Astuti et al. 2004; Atran 1998; Bloch et al. 2001; Moya et al., in press). These mechanisms can be broadly classified as genetic transmission (e.g., folkbiological reasoning) and cultural transmission (e.g., folksociological reasoning). *Genetic transmission* does not posit that people actually represent anything like modern biology's conception of genetic inheritance. Rather, it refers to the tendency to think that, at least for living natural kinds, an individual has a property it does because this was passed on at birth from the biological parent. By contrast, *cultural transmission* posits that external forces are responsible for the way something is (including horizontal and vertical transmission). Any observed cultural differences in the use of these reasoning strategies is thought to be the result of local differences in the deployment of these strategies, not in their availability. In contrast, C&S argue that a single process gives rise to outputs that appeal to inherent causes or to situational causes, but has a natural tendency to point toward (the easier to make) internal causes.

What is the evidence that humans may instead have multiple distinct mechanisms that support both genetic transmission (a close approximation of the inherence heuristic) and cultural transmission rather than a single process suggested by C&S? A recent cross-cultural examination of children and adults from Peru, Fiji, and the U.S. explored judgments about the transmission of traits for social groups including beliefs, skills, personality, group identity, and physical properties (Moya et al., in press). This work examined the use of folkbiology, folksociology, and domain-general structured learning mechanisms to make causal attributions about the transmission of these traits. These mechanisms differ in terms of their predictions of cross-cultural convergence and divergence in transmission judgments and in the sensitivity of these judgments to local differences (e.g., patterns of group migration and permeability of group boundaries). Crucial to the present argument, there is substantial cross-cultural variability in children's pattern of inference and in how their reasoning aligns with adults from their local communities, with children in Peru showing the opposite pattern as their adult counterparts. They also report that children in Fiji show a completely undifferentiated pattern, suggesting that they privilege neither genetic nor cultural transmission explanations when reasoning in the social domain. Evidence that children reason differently from their adult counterparts, especially in a manner that does not emphasize inherence, may pose a problem for the strong claim that the inherence heuristic represents a universal cognitive bias in reasoning.

C&S propose a single mechanism to help explain a broad pattern of phenomena in psychological reasoning – the tendency to imbue a causal explanation for an observed pattern in terms of features inherent to that entity under observation. However, a careful examination of ethnographic evidence and cross-cultural comparisons would suggest that C&S may be guilty of employing their own heuristic to explain the nature of that heuristic. Although an inherence heuristic may very well explain large domain-general patterns of causal attributions among WEIRD populations, cross-cultural comparisons point to the early divergence of mechanisms that emphasize cultural transmission from those that emphasize genetic (or inherent) forces.

The culturally appropriate deployment of these various reasoning biases may be part of an enculturation process that serves to reinforce group boundaries by helping to distinguish in-groups from out-groups. Indeed, an interesting proposal to explore is whether humans rely on culturally normative patterns of explanation as a marker of group boundaries just as they have been shown to do with other normative behaviors (e.g., language, dress, music, and food preferences). Surely, this is the case in terms of individual beliefs that circumscribe individual roles. (For example, prosecutors and opposing defense attorneys are in part differentiated by their belief in the underlying explanation for some behavior – was it the defendant's reckless abandon for the law that led her to rob a store or was she left with no choice in order to feed her family and pay the rent?) Whether broader patterns in the selective deployment of inherence and situational causal frameworks to explain regularities observed in the physical and social world serve as a prominent signal of groupness and one's cultural affiliation remains to be seen.

## Why we assume it's all good: The role of theory of mind in early inherent feature inferences

doi:10.1017/S0140525X13003646

Karen Bartsch and David Estes

*Psychology Department, University of Wyoming, Laramie, WY 82071-3415.*
bartsch@uwyo.edu
http://www.uwyo.edu/psychology/faculty/bartsch.html
estes@uwyo.edu
http://www.uwyo.edu/psychology/faculty/estes.html

**Abstract:** Cimpian & Salomon's (C&S's) characterization of a domain-general inherence heuristic, available to young children, underplays the importance of our early interest in and recognition of agency, intentionality, and mental life. A consideration of the centrality of desires, goals, and agency in our earliest reasoning suggests an alternative, perhaps complementary, account of our tendency to be satisfied with the status quo.

According to Cimpian & Salomon (C&S) our earliest inferences arise from a cognitive heuristic that predisposes us to attribute observed correspondences to inherent qualities of the involved entities. C&S propose that this inherence heuristic paves the way for the development of psychological essentialism, but they describe the heuristic itself as influencing reasoning indiscriminately across a wide range of topics and domains. In casting such a broad net, their characterization of early inferences ignores (or at least underplays) a distinctive element of childhood thought – specifically, children's grasp of agency, intentionality, and the life of the mind. We suspect this element deserves more attention in attempts to characterize early causal reasoning.

For several decades now, research on developing social cognition has attributed to young children some capacity to interpret events in terms of agency, intentionality, and (eventually) mental representations such as desires and beliefs (e.g., Bartsch & Wellman 1995; Wellman 1990). Evidence on children's developing *theory of mind* suggests that, in events involving humans, the fodder for children's "mental shotgun" (Kahneman's 2011 term, as invoked by C&S) likely includes the intentional states of humans – their goals, emotions, "pro-attitudes" (Astington 2001), intentions, and, eventually, beliefs and desires. From infancy (e.g., Onishi & Baillargeon 2005), we seem poised to view phenomena involving humans and other animates as implicating goal-focused agency. Even toddlers can relate specific subjective desires to different individuals (Repacholi & Gopnik 1997).

Indeed, there is reason to suppose that, from early on, we not only *can* view events from the perspective of an "intentional stance," to use Dennett's (1987) term, but we *prefer* to do so. Long ago, Piaget (1929) noted children's tendency to credit even inanimate objects, on occasion, with some sort of intentionality, such that they explained cloud movements by attributing to clouds the desire to follow a child. Young children prefer to characterize actions in terms of mental/intentional states rather than in terms of behavioral descriptions (Lillard & Flavell 1990).

Yet our predisposition for an intentional stance is downplayed in C&S's treatise. Indeed, in providing examples of inherence-based reasoning, the authors seem to go out of their way to avoid intention-related inferences, even when such inferences are quite plausible. For instance, C&S submit that a conclusion following the observation that girls often wear pink might be that "pink is an inherently delicate color" (sect. 1, para. 1; sect. 2.2.2, para. 1). We think that a more natural inference would be that girls like pink. The difference in these inferences is not trivial: The first attributes an inherent quality to the color pink, whereas the second attributes an (arguably) inherent quality to girls – specifically, a psychological quality of girls' intentional stance on colors.

Other examples offered by C&S can similarly be reinterpreted from a person-centered focus. For instance, we suspect that, in pondering why orange juice is typically consumed in the morning, most of us would assume a pivotal role for human agency, desires, and goals. Thus, the inference that orange juice has "energizing" qualities (suggested by C&S [sect. 2.2, para. 7; sect. 3.2, para. 4]) likely presumes there is a human actor who wants to be energized in the morning. When we infer that people wear heavier clothing in the winter because the weather is cold, we are surely assuming that *people* want to stay warm, and not that heavy clothing and cold weather are conjoined in some world apart from human desires and intentions.

The idea that young children bring an intentional orientation to their observations of correspondences – especially correspondences that have anything to do with human (or animate) activities – opens the door to alternative characterizations of the role of the *inherence heuristic* in development. It may, for instance, explain why inherent qualities are broadly inferred in childhood reasoning. Maybe our tendency to view animate agents as possessing internal qualities such as goals and desires begets the assumption that there are inherent qualities even in the physical world, such that eventually we are crediting correspondences to DNA and similar inherent constructs. A consideration of our tendency to bring a person-centered orientation to our first inferences may also help to explain our apparent satisfaction with the way things are: If the way things are is the way people *want* them to be, then maybe it's all good. So we assume a *rightness* about heavy coats being worn in the winter (we assume someone wanted it that way, maybe needed it that way, to achieve a desired goal), and similarly we assume justification for other correspondences (e.g., separate drinking fountains for those with white skin), at least until those arrangements are questioned, because we start with a cognitive framework in which human desire operates to achieve our goals.

We think some such alternative characterization – one that acknowledges our early recognition of agency and goal-directedness – better accommodates evidence that early cognition recognizes a domain distinction between physical and psychological/social/biological phenomena, consistent with developmental accounts such as Gopnik and Wellman's (1994) "theory-theory." Moreover, this sort of alternative characterization, with its straightforward link to human narrative that necessarily assumes a rational network of desires, beliefs, intentions, and actions affords an easy path to the storytelling that, according to C&S, must be the end product of pattern detection.

## Can the inherence heuristic explain vitalistic reasoning?

doi:10.1017/S0140525X13003658

Brock Bastian

*School of Psychology, University of Queensland, St Lucia, QLD 4072, Australia.*
b.bastian@uq.edu.au
www.brockbastian.com

**Abstract:** Inherence is an important component of psychological essentialism. By drawing on vitalism as a way in which to explain this link, however, the authors appear to conflate causal explanations based on fixed features with those based on general causal forces. The disjuncture between these two types of explanatory principles highlights potential new avenues for the inherence heuristic.

Cimpian & Salomon (C&S) argue that a reliance on inherent explanations may be a precursor to psychological essentialism and especially in the case of living kinds. They suggest that essentialist beliefs arise through a refinement process where ideas about general causal forces (such as internal energy, commonly referred to as vitalism [Inagaki & Hatano 2004]) translate the inchoate early output of the inherence heuristic into a kind specific essence that is responsible for static as well as dynamic features.

Inherence is an important component of essentialist beliefs (e.g., Bastian & Haslam 2007) and is most often characterized by a belief in the biological basis of natural kinds (Bastian & Haslam 2008). However, a belief in underlying biological characteristics that cause tigers to look and behave in similar ways, or that make pine trees different from willow trees, is not the same as a belief in a generalized causal force (internal energy) that makes tigers move or trees grow. Children do not refer to stable inherent features when they posit the existence of internal energy. Rather, they refer to causal forces that may even be transferable or exchangeable between entities themselves. For example, energy transmission may be used to explain why one ball moves when hit by another ball (Morris et al. 2000). Within the biological domain, conceptions of energy may be more akin to a general life force as opposed to something that defines the shared attributes of category members.

How do ideas about energy, which are not specific to kinds but explain growth or movement in general, translate the output of the inherence heuristic into full-blown essentialist accounts? According to C&S, the output of the inherence heuristic is a focus on "stable, enduring characteristics of the entities in question" (sect. 2.2.1, para. 1) Essentialist beliefs share this focus on stable or fixed features that distinguish different kinds. Ideas about internal energy, however, do not distinguish between kinds and are not linked to fixed characteristics, making their role in this process somewhat unclear. As C&S note, "the essentialist belief constructed via this refinement process must transcend children's original ideas about internal energy" (sect. 4.3, para. 5).

This disjuncture between a belief in vital energy and a tendency to attribute fixed essences to kinds is not only apparent in early childhood but also evident within adult thinking. Beliefs in general vital forces, energies, power, or spirits are evident within many traditional belief systems (Frazer 1890/1959; Mauss 1902/1972; Tylor 1871/1974) and in adult superstitious thinking (Lindeman & Saher 2007; Nemeroff & Rozin 1994), and feature in theories such as astrology, feng shui, or homeopathy (Fisher 2002). Perhaps most striking is the observation that causal explanations regarding agent intentionality rely on forms of vitalistic reasoning (Cashmore 2010). Agent intentionality, at its core, relies on a belief in free will – that some internal causal property not determined by biological processes (e.g., the soul) is responsible for individual action.

Just like the concept of internal energy, the concept of souls is not clearly attached to stable or fixed features. Souls can be tarnished, cleansed, converted, corrupted, or lost. One can sell his or her soul to the devil or engage in reprehensible behaviors that contaminate the soul. Luckily, however, souls can also be rehabilitated and purified (Bastian et al. 2011; Zhong & Liljenquist 2006). Souls may also be easily influenced by external factors. Consider the reasoning of televangelist Pat Robertson when he said that the earthquake-ravaged Haiti has been "cursed" by a "pact to the devil" (Shea 2010). This causal explanation draws on both external events (a historical pack with an external entity) and internal qualities (the souls of Haitians). Souls provide excellent causal explanations, but like energy or other vital forces, they are not linked to stable or fixed features and they do not serve to distinguish between different kinds.

Of course, it is clear that C&S did not set out to explain a belief in souls or free will. These examples do, however, raise the question of how the inherence heuristic may be related to vitalism (as opposed to essentialism). Recent work has identified that people rely on vitalistic beliefs within the moral domain (Bastian et al. 2013). Specifically, some people endorse a belief in forces of good and evil, implying that they see these forces as real and agentic. Such beliefs appear to have a range of implications for social cognition, shaping how people understand ideological (as opposed to strictly biological) differences among people. Moreover, these beliefs are not attached to entities, but capture how people think about good and evil in general. This demonstrates

that vitalistic beliefs appear to be largely distinct from essentialist beliefs – a distinction that not only makes it hard to understand how vitalistic beliefs enable the development of essentialism from the output of the inherence heuristic, but also raises the question of how the inherence heuristic may apply in these cases. Can the inherence heuristic explain a reliance on vitalistic reasoning in both children and adults? Put differently, can the inherence heuristic explain the many ways in which people seek to make intuitive sense of the world by relying on causal properties that are not themselves fixed characteristics linked to kinds?

One would assume that the same mechanism underlying the inherence heuristic – the mental shotgun – would be responsible for promoting vitalistic reasoning in more complex domains, such as human morality. For example, the mental shot gun would likely support character-based explanations for criminal behavior (a bad soul) as opposed to a range of behavior determining developmental and environmental factors. This is particularly apparent in the case of more heinous crimes, where explanations based on external factors seem less fitting: we are more likely to see Hitler as evil than as a victim of hard times.

The inherence heuristic serves to integrate a number of well-known, influential theories. Examining how the heuristic may extend to intuitive explanations based on fluid and shifting internal properties (vitalistic beliefs) may broaden its influence further still.

## The inherence heuristic: A key theoretical addition to understanding social stereotyping and prejudice

Rebecca S. Bigler and Caitlin Clark
*Department of Psychology, University of Texas, Austin, TX 78712.*
**rebeccabigler28@gmail.com**     **caitlinclark11@gmail.com**

**Abstract:** Prior work has detailed the constructivist processes that lead individuals to categorize others along particular dimensions (e.g., gender) and generate the content (e.g., stereotypes) and affect (e.g., prejudices) associated with social groups. The inherence heuristic is a novel mechanism that appears to shape the content and rigidity of children's social stereotypes and prejudices.

Much progress has been made in mapping the mechanisms by which individuals develop stereotypes (i.e., ascribing attributes to others on the basis of membership in a particular social group) and prejudices (i.e., positive or negative affective reactions to others on the basis of membership in a particular social group). Nonetheless, aspects of the complex processes remain unexplained. Cimpian and Salomon's (C&S's) proposal that human thought is characterized by a bias to explain observed patterns in terms of the inherent properties of their components makes an important contribution to our understanding of the ontogeny of social stereotyping and prejudice. We review some of what is known about the causes of stereotyping and prejudice and describe novel insights that stem from C&S's work.

Stereotypes and prejudice are founded on categorization. There are myriad bases on which humans might be parsed into groups, and thus one or more mechanisms lead children to attend to some (and ignore other) bases for classifying people. Bigler and Liben's (2006; 2007) *developmental intergroup theory* (DIT) posited that evolution led to a flexible cognitive system that motivates and equips children to infer from environmental data which bases of classification are important within a given context. DIT proposes that the psychological salience of grouping criteria (e.g., gender, age) increases when adults label groups needlessly (e.g., "Good morning, girls and boys" rather than "Good morning, students") or in the service of organizing the environment (e.g., assigning desks by gender).

In addition to the explicit marking of attributes, DIT posits that *implicit* sorting of humans by particular dimensions (e.g., race, age) increases the psychological salience of such group criteria. One powerful example is de facto segregation. Although segregation affects stereotyping and prejudice via its effect on familiarity (Pettigrew & Tropp 2000), DIT proposes an additional, inferential, constructivist process in which children notice perceptual similarities among those who live, work, and socialize together and then *infer* that these observed social divisions are caused by meaningful, inherent differences among groups (Bigler & Liben 2006). The inherence heuristic is just such an inferential process and adds a novel component: it leads children to view the patterns of group sorting as natural, legitimate, and good.

Bigler and Liben (2006) argued that once a particular basis of classification becomes salient and is used to parse humans into groups, multiple mechanisms contribute to the development of social stereotypes and prejudice. In addition to constructivist processes (e.g., essentialist thought, illusory correlations, in-group bias), DIT proposes that children detect veridical correlates (e.g., activities, roles) associated with particular social groups (e.g., women and cheerleading; men and football). The inherence heuristic expands this account by arguing that children interpret (rather than merely learn) such links. Specifically, children construct schemata about social groups that reflect their belief that the observed correlates (e.g., roles) are inherent to the groups. Simultaneously, correlates are interpreted through stereotypic views of the groups with which they are linked. (For example, medical professions are viewed as more nurturing when performed by women than men.)

We provide two illustrations of the utility of the inherence heuristic for understanding stereotyping and prejudice. The first case concerns the development of heterosexist attitudes (Clark & Bigler 2013). Heterosexual couples and parents appear more often in children's environments than their nonheterosexual counterparts, as a result both of the greater statistical frequency of the former and societal prejudice against the latter (Martin 2009). Thus, children repeatedly observe that romantic relationships include one man and one woman, an implicit sorting of the type hypothesized by DIT to facilitate attention to gender. According to C&S, children should use the inherence heuristic to explain the pattern of cross-sex romantic relationships and, as a consequence, view heterosexual romantic pairings as natural, inevitable, and legitimate. Such romantic pairings should also be viewed as a product of the inherent qualities of maleness and femaleness: to be a member of one group requires romantically desiring members of the other. Deviance from the pattern should be viewed as wrong and unnatural. In an ongoing study, we present children with images of straight, gay, and lesbian couples, and ask them to rate their liking of the targets (among other questions). Preliminary findings are consistent with the operation of the inherence heuristic: 5- to 7-year-old children's responses to advertisements featuring same-sex couples often include rejection and disgust (e.g., "The two people cannot go together; it's just not good" and "They're two girls. They're not supposed to marry each other. Eww, eww, eww").

A second illustration comes from three studies, conducted before President Barack Obama's election, that examined 5- to 10-year-olds' knowledge of and explanations for the lack of female, African American, and Latino U.S. presidents (Bigler et al. 2008). Most children knew that these three groups had been excluded (i.e., they had detected the patterns linking the presidency exclusively to white men). When asked to *explain* the absence of women and people of color from the role, children's answers frequently reflected the operation of the inherence heuristic. So, for example, most children failed to note past circumstances or external constraints (e.g., legal restrictions, discrimination) and instead reasoned that the excluded individuals were ill-suited for or uninterested in the role (e.g., "Women aren't as smart as men" and "Latinos would rather do other things, like

drive trucks"). They also frequently relied on precedent to justify their belief that women, Latinos, and African-Americans *should* be excluded (e.g., "The presidents are always boys").

Finally, the inherence heuristic is a promising theoretical base for the design of intervention programs aimed at reducing stereotyping and prejudice. Many extant interventions have proven ineffective or impractical for reducing prejudice. For example, counterstereotypic models are relatively ineffective because they are outnumbered by traditional, stereotypic models and are often misremembered. Explicitly teaching children about discrimination – an external attribution for observed group differences – might reduce reliance on the inherence heuristic, thereby reducing intergroup biases. The recommendation to discuss intergroup biases explicitly with children is consistent with the DIT and assessments of the benefits versus risks of such lessons (Bigler & Liben 2007; Bigler & Wright 2014).

# The psychology of inherence is self-referential (and that is a good thing)

Fred L. Bookstein

*Department of Statistics, University of Washington, Seattle, WA 98195; and Faculty of Life Sciences, University of Vienna, 1010 Vienna, Austria.*

flb@stat.washington.edu        fred.bookstein@univie.ac.at

**Abstract:** Cimpian & Salomon (C&S) appear to characterize the inherence heuristic and essentialism as unwise or childish aspects of human reasoning. But actually, these cognitive modes lie at the core of statistical analysis across all of the quantitative sciences, including the developmental cognitive psychology in which the argument here is couched. Their whole argument is as much an example of its topic as an analysis of it.

It seems an implicit axiom of Cimpian & Salomon (C&S) that the form of reasoning its authors are studying is childish and unwise. The authors have very little good to say about the inherence heuristic or the essentialism to which it leads. It is characteristic of children as young as 4 years of age, of persons needing reassurance, and of people with a low cognitive capacity or under conditions of high cognitive load. It leads to "errors in moral reasoning" (cf. the target article abstract), "the correspondence bias" (sect. 5.4), and so forth. It is responsible for prejudice and perhaps for religious warfare. Even Kahneman's great essay of 2011 was not so negative. There this form of reasoning was explicated not as childish but as "System 1," the making of a decision speedily or without concentrating overmuch on the weighing of evidence. In other words, Kahneman used the category to characterize not people but episodes of reasoning.

And yet the inherence heuristic and essentialism that have been derogated in this wise characterize the core of contemporary statistical data analysis across the sciences, including the developmental cognitive sciences that are the domain of C&S's argument. I mean this literally: the authors' assertion that "this heuristic is a pervasive feature of human cognition" (sect. 1, para. 4) is actually an example of that selfsame essentialism.

Consider, for instance, the statistical notion of an average, which runs inexorably through all these reports of what is found in studies of children aged 3, 21, or in between. In any report of an empirical psychological investigation, averages are always to be found. Within any sampling frame, the optimality (maximum likelihood) of an average is indistinguishable from the (very simple) theory entailed in that computation: that there exists an inherent *true value* (mean) characterizing measurements made on the members of that group, such that deviations from that value are distributed in a bell curve (a Gaussian distribution). This argument was first made by Carl Friedrich Gauss himself in

this context – hence the distribution's name. In other words, anyone who computes an average is already working inside the inherence heuristic. Similarly, a regression coefficient is the weighted average of a collection of separate estimates of the same (sub) sample-specific slope, which thereby becomes an "inherent" characteristic of a sample right alongside its average. This point was made by the British statistician Francis Y. Edgeworth in the 1890s (Stigler 1986). A correlation, if required, is just a small arithmetical modification of this same averaged slope. Further out on this thread of statistical reductionism are the path coefficients generalizing multiple regressions into interlinked linear causal laws (Wright 1968; 1969; 1977; 1978), the advances in econometric method that are two-stage and three-stage least squares (Malinvaud 1966) and the combination of either of these with factor analysis in the formalism of "structural equation models" (Wold & Joreskog 1982). All such coefficients are instances of essentialism.

Indeed, everywhere we look in the statistical sciences, whether physical sciences, natural sciences, psychological sciences, or social sciences, we encounter this same urge toward the inherency of summary numerical parameters. Some contemporary statisticians place this drive at the very center of the statistical mindset, as, for instance, when Cox (2006, p. 7) states that statistical inference "takes the family of models as given and aims to give intervals or in general sets of values within which [the parameter] is likely to lie, [and] to assess the consistency of the data with a particular parameter value." The consilience of statistical parameters that lies at the core of all of quantitative science (Bookstein 2014) is at root little more than the extension of C&S's ideas of inherence to the objective subworlds of controlled observations that are the object of scientific study discipline by discipline. What could be more "essentialist" than the periodic table of the elements, which is posted on the wall of every high-school chemistry classroom? The authors inadvertently make the point that essentialism and the inherence heuristic likewise characterize academic psychology.

Even more generally, the inherence heuristic may interact closely with the "Ionian enchantment" (Wilson 1998, pp. 4–5), the trust that our understanding of the patterns of this world rests at least as much upon some fundamental structures in that world as upon cognitive customs of our minds. Philosophers (e.g., Peirce 1934; Russell 1967, p. 3) have often agreed with physicists (e.g., Wigner 1960) that the role of numbers in describing our world is ultimately transcendental. For this to be inherent in human reasoning by the age of 4, even if limited to the aspects of numeracy that deal with size, count, and direction, must be an aspect of that "unreasonable effectiveness of mathematics" that Wigner refers to.

Perhaps the authors' objection is not really to essentialism and the inherency heuristic per se but instead to a mildly age-graded category error governing its typical applications. The child likely lacks understanding of the difference between aspects of the objects in our world that are appropriately taken as inherent or essentialist, such as the properties of water or our usually pentadactyl limbs, and aspects such as personality that are properly studied only after careful, controlled observation. But even in the academy, these are not distinctions of mind but only of subject matter.

To this extent, the child's essentialism is not to be deprecated or deplored, but celebrated. Recourse to inherence and essentialism, far from being a pejorative characteristic of mainly the youngest children, the intellectually challenged, those of a System 1 cognitive style, and the like, appears to be an inherent theme of human ratiocination at all levels from earliest childhood through the halls of academe. Our authors should be reveling in their discovery that ordinary human childhood prepares us in part to be scientists, not just large, unusually fecund, eusocial mammals. Meanwhile, from its very first sentence to its last, the article by C&S serves as much as an example of its subject as it does an analysis.

# Not so fast, and not so easy: Essentialism doesn't emerge from a simple heuristic

Nick Braisby

*School of Psychology, Social Work and Human Sciences, University of West London, London W5 5RF, United Kingdom.*
nick.braisby@uwl.ac.uk
http://www.uwl.ac.uk/about-us/how-university-works/biographies/professor-nick-braisby

**Abstract:** Cimpian & Salomon's (C&S's) proposal comes unstuck on precisely the claim that inherence is an heuristic, able to deliver simple, shallow outputs that are right most of the time. Instead, the inherence heuristic delivers outputs that imply it is not an heuristic after all, and is simply too fast and too easy a mechanism to do the job of explaining categorisations.

Cimpian & Salomon (C&S) posit an *inherence heuristic* to explicate how humans explain patterns of co-occurrence. Their proposal is that the heuristic requires explanations to be couched in terms of the inherent features of the objects that figure in the co-occurrence pattern. So, for example, that lions co-occur with lion properties is to be explained in terms of the inherent features of lions and their properties. C&S argue that explanations of this sort are precursors to *psychological essentialism*, the view that a category's properties are determined by its essence, and that people's beliefs and behaviours are consistent with this.

However, C&S's proposal appears to come unstuck on precisely the claim that inherence is an heuristic. Heuristics are fast and frugal cognitive mechanisms (Gigerenzer et al., 1999) that deliver simple and shallow outputs – outputs that are right most of the time (cf. bounded rationality [Simon 1957]). These properties provide a plausible basis for heuristics' adaptive value: Heuristics have evolved because they are simple and fast, and because they generally deliver the right answers. In effect, heuristics succeed because they transform problems that are difficult to solve and computationally complex into ones that are easier to solve and computationally simple (Kahneman & Frederick 2002; Tversky & Kahneman 1973).

However, a careful unpacking of C&S's account indicates that the inherence heuristic requires cognitive outputs to be neither quick, nor shallow, nor adaptive.

Take the co-occurrence of lions and lion properties. According to the inherence heuristic, an explanation of this co-occurrence will involve finding a story that explains why the inherent features of lions and the inherent features of lion properties co-occur. Of course, such an explanation is not obvious. Indeed, C&S acknowledge that the heuristic may deliver the outcome that the to-be-explained co-occurrence derives "in some to-be-determined fashion" from a "to-be-determined combination of inherent features" (sect. 4.3, para. 1) (consistent with the argument that categories may be associated with an essence "placeholder" [sect. 4.3, para. 6]).

Given that heuristics are thought to operate quickly, deliver shallow outputs, and offer adaptive value, we need to ask how the inherence heuristic fares on these criteria.

For lions and lion properties, the search for a co-occurrence explanation will ultimately terminate either in an explanation that is approximately true (e.g., lion DNA), or one that is false, or one that is a placeholder. The true explanation may arguably be adaptive, but would be neither quick nor shallow. The false explanation could be quick and shallow but, being false, is hardly likely to be adaptive. The placeholder explanation can be neither quick nor shallow and is of questionable adaptive value. Thus, either way you look at it, the inherence heuristic delivers outputs that imply it is not an heuristic after all.

Indeed, the inherence heuristic proposes people do very considerable cognitive work of very little value, inconsistent with the idea that heuristics turn computationally complex problems

into simpler ones. Take the very regular co-occurrence between the sun and the location of its rising: From a fixed location, the sun always seems to rise in the same place – on the East coast the sun always rises over the sea. According to the inherence heuristic, people seek an explanation for this in terms of the inherent properties of the sun and the sea (or that part of the sea where the sun breaches the horizon). But such a search would be fruitless. The sun rises where it does because of its relation to the Earth, not because of the inherent properties of either. The inherence heuristic would ultimately generate either a false or an inchoate explanation. Crucially, it could not be quick and shallow and adaptive. Take another example where relational structure is central: the very regular co-occurrence between liquids and their direction of flow – liquids flow downhill. The inherence heuristic implies people strive to explain this in terms of the inherent properties of liquids and the ground on which they flow. The heuristic entirely misses the relation (gravitation) essential to a successful explanation, and thereby must predict that people ultimately generate either false or inchoate explanations.

These kinds of example could be given computationally simpler explanations by an heuristic mechanism. A view that C&S briefly acknowledge, but ultimately eschew, is that successful explanations might advert to "brute statistical facts" (sect. 1, para. 4). Why do lions and lion properties co-occur? Why does the sun rise where it does? Why do liquids flow downhill? Because that is the way things are; because these are just brute facts. Unlike the inherence heuristic, a brute-fact heuristic would appear genuinely to be quick, shallow and adaptive. Whether brute facts exist is a matter for philosophical analysis. The point is that, whether or not they do, there is nothing to stop an heuristic device presupposing their existence in explaining co-occurrence patterns.

Strangely, C&S's proposal is not so inconsistent with brute-fact explanations. It relies on the claim that stable, inherent properties of objects are laid down in semantic memory before co-occurrence explanations are sought. It is as if the inherence heuristic actually relies on a prior stage where properties of a to-be-learned category are processed simply as if they were brute facts, no explanation required, no questions asked. But if inherence can tolerate brute facts in laying down semantic memories, it is puzzling indeed that C&S reject it in explaining co-occurrence.

Human history is replete with examples where explanations for earlier categorisations have been sought, and categorisations challenged or overturned (e.g., Eco 1999). In this, the inherent features of categories and their properties figure strongly, but people also go beyond inherent features and look to relations, too. Likewise, searching for an explanation of categorisation in terms of essential properties is likely to be neither quick nor computationally simple. The point is that these societal processes are complex. No doubt there must be intra-personal analogues, but the lesson of human history is that, to do the job, explanations will be computationally complex and deliver deep outputs. A quick and simple heuristic is just not up to the job – it is at once entirely too fast and entirely too easy.

# Inherence is an aspect of psychological essentialism

Tara C. Dennehy

*Department of Psychology, University of Massachusetts, Amherst, MA 01002.*
**tdennehy@psych.umass.edu**
**http://people.umass.edu/tdennehy**

**Abstract:** Inherence is not a distinct construct from psychological essentialism; it is one of several underlying beliefs. I propose that inherence is only one entry point to the perception of an essence and

posit that context may influence which aspects of essentialist reasoning precede inferring an essence. I also discuss how psychological essentialism can uniquely account for violations of category-based expectancies.

Cimpian & Salomon (C&S) propose that an *inherence heuristic* – people's tendency to explain statistical regularities in the environment by attributing these patterns to inherent factors – is a necessary precursor to psychological essentialism. The crux of this commentary is that inherence is not a separate construct from psychological essentialism; many of the phenomena the authors explain via the inherence heuristic can be explained by psychological essentialism (for system justification, see Keller 2005; for nominal realism, see Diesendruck & Haber 2009 and Gelman & Diesendruck 1999). Psychological essentialism also has greater explanatory power than the inherence heuristic alone (for a similar critique of Strevens' K-laws, see Ahn et al. 2001). I discuss evidence that, rather than being a precursor to essentialism, the perception of inherence belongs to a subset of essentialist beliefs. Indeed, inherence is highly correlated with essentialist beliefs (*r*s up to .74) across a number of studies (e.g., Bastian & Haslam 2007; Haslam et al. 2000) and is sensitive to manipulations of related essentialist beliefs (Haslam & Ernst 2002). Finally, I highlight how psychological essentialism can account for violations of category-based expectations.

*Psychological essentialism* is the belief that category members share a deep, underlying essence that is responsible for their internal and external attributes, as well as membership in the category. This belief can exist without evidence or even knowledge as to the nature of this essence (i.e., there can be an "essence placeholder" [Medin & Ortony 1989]). Based on theory and empirical evidence, four core beliefs are related to psychological essentialism: naturalness/biological etiology, immutability, informativeness, and discreteness (e.g., Bastian & Haslam 2006; 2008). The belief in *naturalness/biological etiology* is the idea that category membership is based in a biological reality (e.g., chromosomes determine gender). The belief in *immutability* is the idea that category membership is unchangeable; you cannot change your category membership (e.g., you cannot change genders). The belief in *informativeness* is the idea that categories are very high in inductive potential, such that membership in a category provides considerable information (e.g., knowing someone's gender is highly informative of who they are). Finally, *discreteness* is the belief that there are clear separations or distinctions between members of different categories (e.g., there is a clear dividing line between men and women). Earlier theorizing on psychological essentialism suggested that a number of additional beliefs might underlie essentialism (e.g., Haslam et al. 2000), but consequent empirical research demonstrated that such beliefs (including inherence) either lacked predictive power or were subsumed by the four core beliefs.

Inherence may be more or less relevant to the inference of an essence depending on the domain or context. Consider the example of pink and blue clothing, as discussed by C&S. Pink and blue (although particularly pink) are colors that are highly diagnostic for gender (Ben-Zeev & Dennehy 2014). An inherence judgment could lead to inferring an essence, such that if we observe a statistical regularity that more girls wear pink and more boys wear blue, we might conclude that there is something inherently different about girls and boys that relates to their color preferences, and that this difference is *essential* ("little girls are made of sugar and spice…" and therefore must like pink). However, I propose that, in the case of gender and clothing color, the inference of an essence would more likely follow from discreteness and informativeness judgments (with inherence judgments arising as a consequence of perceiving an essence). Say we observe a distinct pattern among boys and girls on the playground that boys predominantly wear blue and girls predominantly wear pink. This could lead us to judge that children in pink versus blue belong to *discrete* and *informative* groups. Discreteness and informativeness are both aspects of psychological essentialism (e.g.,

Haslam et al. 2000; Haslam & Levy 2006); simply observing or attributing discrete or informative features to groups may be sufficient information for us to assume that *something* must be causing boys to wear blue and girls to wear pink. C&S assert that we would ask the question "Which inherent features explain this pattern?" and that we would apply an inherence judgment to the color pink, concluding that it is inherently more delicate and is therefore more suited to girls. However, I propose that this step is unnecessary. Once we have observed a pattern consistent with essentialist reasoning, we can apply an essence placeholder and assume that there is an essence underlying the pattern we observe.

Psychological essentialism and inferences of essentiality can also account for violations of statistical regularities in a way that inherence-based reasoning alone does not. Of relevance, Rehder and Burnett (2005) have proposed an *underlying mechanism model*, which posits that people have a bias toward perceiving underlying mechanisms that cause category-based features (e.g., testosterone could be seen as the underlying mechanism causing facial hair in men). Similar to Medin and Ortony's (1989) concept of the "essence placeholder," people do not require concrete information about the underlying mechanism to assume its existence; it is something that is simply inferred. The underlying mechanism model is highly complementary to psychological essentialism because the mechanism may be the (known or unknown) means by which essences cause observed features. This mechanism can function either well or poorly, explaining within-category variation (e.g., not *all* girls wear pink) beyond the between-category variation (e.g., boys wear blue and girls wear pink) accounted for by the presence or absence of an essence.

To apply Rehder and Burnett's (2005) theorizing and findings to human social categories, consider the case of a boy wearing pink. Psychological essentialism in conjunction with the underlying mechanism model can account for violations of category-based expectancies. In the case of gender and clothing color, a boy dressed in pink might be perceived as having a less well-functioning mechanism that is responsible for the feature of "not-pink" given his essence of being male. A judgment of a poorly functioning mechanism could account for him still belonging to the category of male while exhibiting a feature (pink clothing) not typically associated with maleness.

# A single cognitive heuristic process meets the complexity of domain-specific moral heuristics

Veljko Dubljević and Eric Racine
*Neuroethics Research Unit, Institut de recherches cliniques de Montréal (IRCM), Montréal, QC H2W IR7, Canada.*
veljko.dubljevic@ircm.qc.ca      eric.racine@ircm.qc.ca
http://www.ircm.qc.ca/neuroethics/en

**Abstract:** The *inherence heuristic* (a) offers modest insights into the complex nature of both the *is–ought* tension in moral reasoning and moral reasoning per se, and (b) does not reflect the complexity of domain-specific moral heuristics. Formal and general in nature, we contextualize the process described as "inherence heuristic" in a web of domain-specific heuristics (e.g., agent specific; action specific; consequences specific).

Cimpian & Salomon (C&S) propose the "inherence heuristic" as an explanation of the cognitive process that leads people to explain observed patterns based on their inherent features. They propose a unified account for a diverse set of empirical findings, including "is–ought errors" in moral reasoning (sect. 3.4). We agree with the authors about the relevance of heuristic cognitive processes in illuminating moral judgment and decision making. However, we doubt the ability of the inherence heuristic alone to explain a range of phenomena, including complexity in ethical reasoning, for two major reasons: (1) the inherence heuristic offers modest insights into the complex nature of both the is–ought tension in moral reasoning and moral reasoning per se, and (2) the inherence heuristic does not reflect the complexity of domain-specific moral heuristics.

First, is–ought inferences, which Cimpian and Salomon try to explain with the inherence heuristic, require critical scrutiny. Interestingly, David Hume himself did not frame such leaps in reasoning as straightforward errors or fallacies. In fact, his *Treatise of Human Nature* (1739/1975) is an attempt to develop a science of the human nature and mind, which grants an obvious importance to the understanding of cognitive processes and emotions for ethics. G. E. Moore (1903/1971) deserves credit for having transformed Hume's observation into a "fallacy" in support of a doctrine of antinaturalism in ethics. Here lies their first potentially problematic assumption: They have interpreted the is–ought tension as both a dichotomy and a fallacy – that is, "is–ought errors." Granted, crude is–ought transitions are dubious because they rely on debatable overgeneralizations and moral conservatism, but this was exactly Hume's point of caution with respect to such transitions. Hence, the redescription of Hume's point in the language of inherence heuristic may not add explanatory power but carries the problematic overtone of antinaturalism. Such antinaturalism could lead one to reify the is–ought tension into an essence, which is exactly the problem that the authors identify. In fact, the relationship between essentialism and the is–ought tension had already been identified, for example, by Dewey as the "philosophical fallacy" – that is, the supposition that whatever is found true under certain conditions may forthwith be asserted universally or without conditions (Dewey 1922). In contemporary ethics, pragmatic and coherentist theories that recommend wide reflective equilibrium between ethical principles and factual claims provide a more generous and applicable understanding of how statements of facts inform considered judgments and moral deliberation (Beauchamp & Childress 2009; Racine 2008). Importantly, they have avoided creating essences out of normative and descriptive categories based on semantic analyses. The reason is clear: moral judgments and moral reasoning cannot avoid factual assumptions about the nature of moral problems as well as the experience of moral agents. Otherwise, one is pressed to explain where any "ought" could logically come from (Callahan 1996), unless one assumes nonnatural properties as Moore does, which the authors appear to have done.

Second, the inherence heuristic is not contextualized in a broader web of domain-specific heuristics that more explicitly relate to moral theory. This diminishes the relevance and explanatory power of the proposal because the appropriate level of explanation is not identified. The authors provide many examples where the inherence heuristic seems to be at work, but it is more plausible to assume that such a psychological process underlies a number of simple domain-specific heuristics. It could be analogous to the domain-general "attribute substitution" process (Kahneman & Frederick 2002), which is potentially implicated in a range of simple heuristics (see, e.g., Gigerenzer 2008; Sunstein 2005). However, it is certainly not a *single* heuristic that explains everything from beliefs about orange juice to ethical reasoning. The authors need to dissociate between domain-specific heuristics and domain-general "effort reduction" processes (Shah & Oppenheimer 2008). Consider the example of different types of moral intuitions and related candidate domain-specific heuristics.

Deontologists (e.g., Kant 1785/1993) claim that certain *actions* are obviously wrong (e.g., torture) or right (e.g., saving lives). They also sometimes believe that it is an inherent feature of a certain action to be wrong. Most of the time, this strategy works well, but sometimes it misfires and/or leads to problematic inferences (e.g., it is a moral imperative not to lie even to a serial killer).

However, it would be false to assume that this is a feature of only one type of moral intuitions and heuristics. Consequentialists (e.g., Sidgwick 1874/1981) use similar reasoning when they focus on *consequences* (e.g., saving five people as opposed to just one), regardless of the means used or the character traits of the agents involved. They also sometimes believe that it is an inherent feature of a certain consequence to be good, which again can misfire as a strategy and/or lead to morally abhorrent conclusions (e.g., killing one healthy person to harvest organs to save five people is rational).

Even virtue theorists (e.g., Aristotle 1890) use similar reasoning to make their case. Namely, they focus on the character – the habits, intentions, and motives of the *agent* (e.g., is this something appropriate for a virtuous person?). They also sometimes believe that it is an inherent feature of a virtuous person to react morally appropriately in every situation. Again, this might be a sound strategy most of the time, but it can certainly misfire and/or create moral outrage (e.g., if Jesus/Mohammed/Moses/Buddha/Marx responded in such a way, everyone should).

The so-called inherence heuristic could actually be a domain-general process that might inform a range of domain-specific heuristics. However, this needs to be acknowledged to ensure its integration into the rich body of theoretical and empirical moral psychology. Furthermore, it is wrong to assume that the is–ought divide is insurmountable and that fallibility *inherently* means faultiness. If the authors maintain such assumptions, they are falling prey to a process they themselves identified.

# Do we need the inherence heuristic to explain the bias towards inherent explanations?

Yarrow Dunham

*Department of Psychology, Yale University, New Haven, CT 06511.*
yarrow.dunham@yale.edu
http://scdlab.yale.edu

**Abstract:** It is clear that people often make unwarranted inherence-based explanations, but it is less clear that explaining this fact requires the inherence heuristic. Instead, it can be explained by a more general explanatory apparatus operating on the most readily available information, which, depending on the nature of that information, outputs both inherent and noninherent explanations.

We surely suffer from a kind of mental myopia, an inordinate focus on things close at hand and a corresponding difficulty taking in what's further afield. This fact underlies many foibles of human reasoning, including, no doubt, our tendency to "explain many observed patterns in terms of the inherent features of the things that instantiate these patterns" (sect. 1, para. 1). Cimpian & Salomon (C&S) artfully demonstrate just how general this tendency is, and how plausibly it is linked to a diverse range of phenomena, from system justification to essentialism and beyond. So far, so good. But need all this inherence be the result of an actual *inherence heuristic*?

Let us consider two accounts. Story 1 is C&S's, in which the bias towards inherent explanations emerges because we employ an intuitive heuristic that reflexively generates inherent explanations. The heuristic can be overridden by more deliberative processing, but because overriding requires effort, we often end up with inherence even when we shouldn't. On story 2, the same

regularities are explained by something more general – let's call it the *explanatory heuristic*. The explanatory heuristic is very much like the inherence heuristic. In fact, it follows the very same steps pictured in C&S's Figure 2. When faced with something to explain, the explanatory heuristic also conducts a fast, shallow search for relevant information (the "mental shotgun" [sect. 2.2, para. 2]). It, too, passes what it's gathered to a "storyteller" (sect. 2.2, para. 7) that looks for a satisficing explanation, and when it finds one, it too passes that explanation to consciousness, producing the same intuitive sense of *having explained*. So how is it different? Well, notice that we haven't said anything about inherence. That's because the explanatory heuristic doesn't care about inherence. It just takes what's at its fingertips and stitches it together. As C&S review, what's at its fingertips will often be the sort of stuff that leads to inherent explanations. (We too often attend to focal entities instead of their history; to enduring rather than transient properties; and so on.) In those cases, inherence is what we'll get. But if different stuff is available (the situational sort of stuff, say), it will happily output different kinds of explanations, and they'll feel just as intuitively satisfying. In short, the explanatory heuristic can also explain the bias towards inherent explanations, but it does so it via bias in the input rather than bias in the heuristic.

If the two accounts are so similar, it is fair to ask if we're splitting hairs. I don't think so. I take C&S to be proposing a genuine psychological mechanism, the kind of thing that is *in the head*, such that it can be blocked or revised, for example. If that's right, the two accounts are proposing different mechanisms and therefore different answers to the question of *why so much inherence*. Crucially, they also make different predictions regarding what will happen when our explanatory system is stressed – predictions that will be familiar from other work within dual-process frameworks. For the inherence heuristic, manipulations that tax conscious reasoning should lead to more inherence and fewer other forms of explanation, because it should disrupt the more effortful blocking process. For the explanatory heuristic, such manipulations should simply increase reliance on whatever is readily accessible to the mental shotgun, irrespective of whether that information favors dispositional or situational explanations. As far as I'm aware, the work most relevant to this issue favors the latter, more general picture. Most critically, while it's true that cognitive load sometimes leads to greater reliance on inherent or dispositional inferences (e.g., Gilbert et al. 1988), in other cases it leads to greater reliance on noninherent, situational inferences. The crucial point is not that people are capable of noninherent explanations; obviously they are, and this can be explained as the effortful blocking of the inherence heuristic. But the data suggest that such explanations often do not look like effortful blocking; in fact, they often look just as fast and automatic, just as much like an intuitive heuristic, as do inherent explanations! This happens, for example, when you are directly asked to make situational inferences (Krull & Dill 1996; Krull & Erickson 1995), or have been thinking about the dynamic nature of human attributes (Molden et al. 2006), or even when incidentally exposed to relevant information (Todd et al. 2011). All this is difficult for the inherence heuristic to account for, but is consistent with the more general explanatory heuristic: If the information within the shotgun's range includes things consistent with a noninherent explanation ("I'm supposed to be thinking about situations!" "Traits are sometimes malleable!"), then those things become, rapidly and automatically, grist for the explanatory mill.

The question I hope to have raised with these comments is whether our reliance (and overreliance) on inherent explanations requires postulating a heuristic devoted to that end, or whether it could emerge from the same dynamics of our bounded and situated rationality that also produce noninherent explanations. On this alternative account, the reason we often fail to perceive "complex chains of historical causes" (sect. 1, para. 1) is simply that they're complex and historical, and therefore difficult

for our explanatory apparatus to sink its teeth into. But whatever we think of these alternatives, C&S have done the field a service by carefully tracing the strands of inherence winding their way through so many of the phenomena we study and, in so doing, generating many promising avenues for future investigation.

# Why does the "mental shotgun" fire system-justifying bullets?

Danielle Gaucher[a] and John T. Jost[b]

[a]Department of Psychology, University of Winnipeg, Winnipeg, MB R3B 2E9, Canada; [b]Department of Psychology, New York University, New York, NY 10003.
d.gaucher@uwinnipeg.ca        john.jost@nyu.edu
http://ion.uwinnipeg.ca/~dgaucher/lab.html
http://www.psych.nyu.edu/jost/

**Abstract:** We suggest that people privilege explanations relying on inherent rather than contingent factors not only because of an innate cognitive tendency to monitor reality, but because doing so satisfies the desire to perceive the societal status quo as legitimate. In support, we describe experimental studies linking the activation of system justification motivation to the endorsement of inherence-based (essentialist) explanations.

Cimpian & Salomon (C&S) argue provocatively and convincingly for a "deep-seated motivation to uncover the underlying structure of reality" (sect. 2.2, para. 1) that leads human beings – especially children and those relatively low in cognitive ability or motivation – to explain patterns of covariation in the social (and physical) world in terms of stable, inherent, essential characteristics of subjects and objects. They propose that the "mental shotgun," which rapidly generates the most accessible associations and inferences, leads people to conclude, for instance, that women are more likely than men to stay home to raise children *because they are naturally (i.e., biologically and psychologically) well suited for caregiving activities* (cf. Eagly & Steffen 1984; Hoffman & Hurst 1990; Jost & Banaji 1994). In this way, the authors suggest that the "inherence heuristic" provides the "cognitive bedrock on which people can build a motivated system-justifying ideology" (sect. 3.2, para. 3).

There is much to admire about C&S's smooth, sophisticated, aesthetically pleasing synthesis of major strands of cognitive, developmental, social, personality, and political psychology. And we agree heartily that essentialist thinking about social groups contributes to the development and maintenance of beliefs and ideologies that justify the societal status quo (see also Jost & Hamilton 2005; Keller 2005).

It is less clear to us why a *cognitive* evolutionary adaptation (or a purely *epistemic* goal) *to perceive and understand reality accurately* would so disproportionately yield quasi-tautological, stereotypical explanations that serve inherently conservative ends – in the sense of accepting and maintaining rather than critically thinking about differences and disparities among individuals and social groups. To extract "the underlying structure of reality," it seems to us that people would need to consider both inherent (internal, stable) and contingent (external, malleable) causes of hierarchy, division of labor, social order, and so on. Indeed, the authors point out that some individuals (liberals, progressives, and those who enjoy engaging in effortful thought) are much more likely than others (conservatives and those who dislike effortful thought) to incorporate historical and situational factors when considering social inequality. But they have not really addressed the question of *why* the mental shotgun would (more often than not) fire system-justifying bullets that perpetuate rather than challenge existing systems of inequality and oppression,

such as the caste system in India (Jost & Banaji 1994; see also Blanchar & Eidelman 2013; Mahalingam 2003b). A satisfying account must incorporate contextual factors arising from the societal or systemic level of analysis, in addition to individual and group levels of analysis (Solak et al. 2012).

System justification theory holds that most individuals privilege explanations that stress inherent rather than contingent factors, not because of an innate cognitive tendency to monitor reality, but because doing so satisfies the desire to perceive the societal status quo as legitimate and stable. If dispositional characteristics of victims of environmental disasters, for instance, can be cited to explain their plight, then there is no reason to blame (or change) the social, economic, or political system (Lerner 1980; Napier et al. 2006). C&S tantalizingly assert – but have yet to demonstrate empirically – that the process of generating system-justifying attributions "is no different than the process that leads people to conclude, say, that orange juice is consumed for breakfast because of its inherent properties (e.g., its energizing smell, its vitamin C content)" (sect. 3.2, para. 4).

It strikes us as difficult, if not impossible, to construct a definitive test of the hypothesis that the psychological process of generating inherent explanations about orange juice is "no different than" the process of essentializing Untouchables (or women, etc.) to justify their discrimination or exploitation. However, it may be relevant that specific situational factors – such as feelings of system dependence (or inescapability) and exposure to system criticism – are known to affect thinking about social groups (Jost et al. 2015; Kay & Friesen 2011), and it would be surprising if these same factors were to affect thinking about orange juice.

Laurin and colleagues (2010) convinced Canadian women that it would be extremely difficult to emigrate from Canada and subsequently presented them with national statistics indicating that men's starting salaries were 20% higher on average than women's starting salaries. Compared to a control condition (in which leaving Canada was described as relatively easy), these women were more likely to attribute economic inequality between men and women to "genuine differences between men and women" and less likely to attribute it to "unfairness in society" (p. 1076). Furthermore, Brescoll et al. (2013) exposed American men and women to a passage criticizing the United States and found that it caused them to score higher on various measures of biological essentialism with respect to gender differences (see also Gaucher et al. 2013). Additional manipulations of system justification motivation (e.g., goal contagion) produced parallel results. In all of these cases, temporarily activating the motivation to defend and bolster the societal status quo caused individuals to endorse inherency-based (essentialistic) explanations for gender inequality.

Would a heightened sense of system dependence or a threat directed at the legitimacy of the social system similarly cause people to endorse inherence-based explanations for the popularity of orange juice at breakfast? If so, this would be pretty surprising, and it would lead us to wonder whether orange juice at breakfast has become as symbolically significant to the "American way of life" as, say, apple pie. In any case, we recognize that two or more psychological processes could share a common cognitive-developmental antecedent and yet possess distinctive moderators.

At the same time, we feel that much more evidence is required to determine whether and, if so, *why* – in the absence of social and motivational considerations – the mental shotgun fires system-justifying bullets. An evolutionary account based exclusively on the need for perceptual or cognitive accuracy seems incomplete at best. The evidence is fairly clear: The activation of system justification motivation affects the endorsement of those inherence-based (essentialist) explanations that preserve the legitimacy of the status quo. It seems unlikely that the ideological function of such explanations could be chalked up simply to the heuristic processing of information designed to "uncover the underlying structure of reality."

## The inherence heuristic: A basis for psychological essentialism?

Susan A. Gelman[a] and Meredith Meyer[b]

[a]*Psychology Department, University of Michigan, Ann Arbor, MI 48109-1043;*
[b]*Psychology Department, Otterbein University, Westerville, OH 43081-2004.*
**gelman@umich.edu        mmeyer@otterbein.edu**
**http://sitemaker.umich.edu/gelman.lab/home**

**Abstract:** Cimpian & Salomon (C&S) provide evidence that psychological essentialism rests on a domain-general attention to inherent causes. We suggest that the inherence heuristic may itself be undergirded by a more foundational cognitive bias, namely, *a realist assumption about environmental regularities*. In contrast, when considering *specific* representations, people may be more likely to activate attention to noninherent, contingent, and historical links.

Cimpian & Salomon (C&S) propose the *inherence heuristic* (IH), a cognitive process by which people explain observed patterns of experience by appealing to inherent aspects of the elements that make up those patterns. This original and fascinating perspective impressively unifies many disparate psychological phenomena, ranging from people's explanations of social behaviors and achievement, to system justification, to nominal realism, to reasoning errors in the moral domain.

We note two important theoretical implications of the findings that C&S provide. First, the work joins other developmental accounts in stressing the importance of theory-driven knowledge from an early age (Carey 2009; Gelman & Waxman 2009; Gopnik & Schulz 2004; Wellman 2010). Contrary to some popular models (Sloutsky & Fisher 2008), children do not merely form associative links between co-occurring features, but they construct explanatory models to account for such patterns. Second, by suggesting that essentialism emerges from a domain-general mechanism that is involved in many cognitive tasks, C&S argue that essentialism is not an evolved, modular adaptation (contra Atran 1998; Pinker 1994). Other domain-general cognitive mechanisms may underlie psychological essentialism as well, including: an appearance–reality distinction, induction from property clusters, causal determinism, tracking identity over time, and deference to experts (Gelman 2003). It remains for future work to determine how these different components might emerge in development.

C&S present a framework in which people view regularities as inherent, not contingent or resulting from historical accident. One might therefore conclude that the human mind focuses preferentially on fixed and unchanging features. Yet people are also often highly sensitive to change and nonconstancy (Woodman et al. 2012), and much of our social interactions revolve around historical conditions (e.g., ownership; inferences about thoughts, desires, and underlying motivations, which nearly always link to history). Indeed, even 2-year-olds are highly sensitive to historical features of an item, including prior ownership (Friedman et al. 2013; Gelman et al. 2012). So when do people care about history, and when do they not?

The clue, we think, lies in C&S's point that IH preferentially produces explanations for patterns or regularities. The implied converse, though not explicitly discussed, is that specific representations are more likely to activate attention to noninherent, contingent, historical links. Thus, for example, although people typically focus on inherent features to explain the general pattern that girls wear pink, we propose that they may typically assume historically contingent features when faced with specific facts (e.g., Alice is crying) (Frazier et al. 2009). On this account, different kinds of information are accessible, depending on the question at hand. People do not have a singular "shotgun" that primarily accesses inherent features, but rather a flexible conceptual system that at times selects inherent features but at times selects noninherent features (e.g., recent events).

This distinction between generic patterns and specific instances suggests that the IH may itself be undergirded by a more foundational cognitive bias – namely, *a realist assumption about environmental regularities*. It is the regularities per se that are perceived as real, stable, and nonaccidental (Gelman et al. 2010; Rhodes et al. 2012). This would then dovetail with the essentialist claim that people treat categories (which are among the most frequently encountered environmental regularities) as stable natural kinds.

This alternative interpretation of the data suggests that some of the apparatus described as part of IH may follow from realism rather than underpinning it.

Consider one of the examples that C&S sketch out: When explaining why girls wear pink, it is suggested that some inherently feminine aspect of pink (e.g., delicacy) is easily accessed and is combined with the easily activated idea that girls are also inherently delicate. Such explanations are described as a starting point for the emergence of essentialism; once children begin to assume that inherent reasons account for observed patterns, they then build up the notion of an "essence" on the basis of this assumption.

However, many explanations of patterns seem *not* to rest on a match of inherent elements of constituents, but rather on an assumption that the regularity in question has an inherent basis that generates these features. For example, we suggest that it is not the case that explaining why girls wear pink *relies* on a match between inherent features of pink and girls, but rather the reverse; by assuming that it is natural for girls to wear pink, an individual is *prompted* to search for an explanation that fits this framework (e.g., generating the inference that pink is delicate). Thus, we see judgments of a feature's inherent match with a category as a consequence of realism rather than as a building block of an emergent essentialist intuition. On this view, the process depicted in Figure 2 of the target article (in which people notice a pattern, look for causal explanations, access a mental shotgun to find inherent features, and so forth) is epiphenomenal rather than the process by which the IH comes to be.

Supporting this argument, consider that children routinely appeal to inherent bases when reasoning about entirely novel features of categories (Cimpian & Markman 2009; Gelman 2003). In such cases, the content of children's mental shotguns regarding these features is likely impoverished, if not entirely empty. C&S suggest that, in such cases, the heuristic "will output a more inchoate sense that some to-be-determined combination of inherent features is responsible in some to-be-determined fashion for the observed pattern" (sect. 4.3, para. 1). And yet, it is exactly the routine accessibility of inherent features of the constituents, and the match between these features, that are seen as necessary ingredients in *activating* the heuristic. It would seem more parsimonious to assume that, in these cases, it is not accessibility of the causal story per se that is crucial, but rather some broader assumptions about the regularities involved.

In the end, we agree with C&S's claim that "infants (and laypeople in general)…show a deep-seated motivation to uncover the underlying structure of reality" (sect. 2.2, para. 1). It is this realist assumption – that there is an underlying structure to perceived regularities – that we argue may be one of the foundations of essentialism.

## The inherence heuristic is inherent in humans

James A. Hampton

*Department of Psychology, City University London, London EC1V 0HB, United Kingdom.*
**hampton@city.ac.uk**
**http://www.staff.city.ac.uk/hampton**

**Abstract:** The inherence heuristic is too broad as a theoretical notion. The authors are at risk of applying their own heuristic in supporting itself. Nonetheless the article provides useful insight into the ways in which people overestimate the coherence and completeness of their understanding of the world.

The principle that people try to make sense of the world by assuming that observed patterns are explainable in terms of deeper structures is clearly a fundamental aspect of intelligent cognition. Any species with capacities for learning beyond simple conditioning has evolved the ability to pick up on the deeper causal structure in the world and to use it to avoid relying on simple appearances. What is interesting about Cimpian & Salomon's (C&S's) proposal is that they suggest that humans treat this as the default situation. People automatically assume there is something inherent in the nature of things that leads to observed patterns of behaviour or social practice, whether or not such a principle in fact exists. Rather than knowing the reasons for things being the way they are, people start with the known facts and then rationalise the underlying reasons. If plausible, urban myths are founded. A well-known example is the supposed meaning of the word *posh* as being an acronym printed on steamer tickets for the richer class of English folk bound for India – "port out, starboard home" – meaning a cabin on the cooler, north-facing side of the ship. It is accepted that this explanation of the word is a fabrication, and there are many others of a similar nature (O'Conner & Kellerman 2009).

As well as implicating our tendency to rationalise, there is also a clear connection between the inherence heuristic and Rozenblit and Keil's (2002) discovery of the illusion of explanatory depth. In their studies, people claimed to understand the workings of everyday mechanisms such as toilet flushes or helicopter rotors, but when challenged they had to admit to having incoherent or at best incomplete understanding. There is such a strong pull to feel that our concepts must be coherent that we easily overestimate the level of comprehension that we possess. (Students often find this out, too, when exams come around.)

The proposed heuristic, or perhaps it is a bias, explains a range of different behaviours. It also draws together many familiar characteristics of human thought – from conservatism and reification to attribution theory and psychological essentialism.

The heuristic is also perhaps reflexive – the authors are themselves attributing these observed patterns of behaviour to an inherent inherence heuristic. There is something inherent in humans, they claim, which explains why they tend to assume that kinds have essences. Perhaps as a consequence of this reflexivity, there is a risk of the heuristic lacking explanatory power. It's vaunted "explanatory promiscuity" (sect. 3.3, para. 8) is surely not a positive characteristic for any theory. As the authors again comment (sect. 5.2, para. 3) "the inherence heuristic can be invoked to explain pretty much any observed pattern," and this could be considered to be the primary weakness of the proposal. For example, differences in the tendency to essentialise natural and artefact kinds are attributed to vague notions of the causal narratives that may be available to the child, but the account, here and elsewhere, is often largely circular. The authors are subject to their own heuristic. Children develop in particular ways because of something inherent in them or in their situation. The inherence heuristic is like a first stage in scientific exploration where the researcher suspects the presence of a deeper process generating the observed patterns. The question is whether the theoretical proposal made goes beyond this.

On a more positive note, the target article does highlight the generality of the issues involved in people's need to explain and understand the world. The literature on explanation has had a strong emphasis on causal explanation of events, but a lot of everyday explanation is much weaker than this. For example, Heussen and Hampton (2008) looked at how people explain the properties of different kind concepts. Why are emeralds expensive, or why do catfish have gills? Explanations followed some familiar patterns, such as cause and effect (glass is transparent because of its molecular structure), functional (catfish have gills in order to breathe under water), teleological (axes have blades because they are used for chopping), and categorical (penguins have feathers because they are birds). But,

interestingly, many explanations were underspecified, simply explaining one property in terms of another in a relatively vague way. In several cases, explanations were considered plausible in each direction – a symmetrical explanation that raises obvious concerns of circularity. Whistles are loud because they are used for alerting people, but they are used for alerting people because they are loud. Dolphins are mammals because they give birth to live young, and they give birth to live young because they are mammals.

The circularity of these explanations speaks to the underlying homeostatic web of interlocking features that characterise our concepts (Boyd 1999; Quine 1960). Understanding a topic involves finding the relations that link the different aspects of a concept, without the need for a clear narrative moving from a primitive deep starting point through to the surface features that we observe. It is possible therefore that the *inherent* property that people appeal to when accounting for girls wearing pink or people drinking orange juice for breakfast may often be a shorthand for this unanalysed set of interlinked properties.

# The essence of essentialism?

Nick Haslam

*School of Psychological Sciences, University of Melbourne, Parkville VIC 3010, Australia.*

**nhaslam@unimelb.edu.au**

**http://www.psych.unimelb.edu.au/people/nick-haslam**

**Abstract:** As an account of the cognitive processes that support psychological essentialism, the inherence heuristic clarifies the basis of individual differences in essentialist thinking, and how they are associated with prejudice. It also illuminates the contextual variability of social essentialism, and where its conceptual boundaries should be drawn.

The idea that an intuitive heuristic underpins essentialist thinking is potentially attractive for social and personality psychologists. I am personally gratified by the focus on "inherence," having identified it as a key element of essentialist thinking in my early work on the topic (e.g., Haslam et al. 2000; 2002). Cimpian & Salomon's (C&S's) analysis of the heuristic's component processes helps to define the cognitive foundations of essentialist thinking. This process account has several implications for the study of psychological essentialism in the social domain.

C&S's process account can shed some light on individual differences in essentialist thinking, a question of obvious importance to the study of prejudice, but one that has been insufficiently theorized. The heuristic is said to involve a process of intuitive "storytelling" that explains observed patterns with reference to their inherent properties. This explanatory focus on indwelling, constitutive features is presented as a basis or precursor for psychological essentialism. Individual difference variables that influence the complexity and thoroughness of the storytelling process should therefore contribute to variations in essentialist thinking.

The target article details how cognitive abilities and styles might in theory influence the output of the inherence heuristic. It also shows empirically that the heuristic is associated with individual differences in essentialist thinking and with several measures of cognitive ability and style. These findings complement and help to explain an emerging body of work showing that cognitive styles and epistemic needs are associated with essentialist thinking in the social domain. For example, Roets and Van Hiel (2011) have demonstrated that people high in *need for closure* tend to hold more essentialist beliefs, particularly in the inherent basis of racial categories, and that this relationship is both causal and consequential. Manipulating the need influences the endorsement of essentialist beliefs, and variations in the need are associated with

racial prejudice through their mutual association with essentialist thinking. Relatedly, Tadmor et al. (2013) show that racial essentialism is negatively associated with creativity, an association that also appears to be causal rather than merely incidental.

Findings such as these add credence to the idea that essentialist thinking is linked to basic cognitive processes that generate several species of closed-mindedness. The inherence heuristic account specifies the nature of those processes in more detail than does prior work in social psychology. The target article's claim that cognitive ability should also predict differential reliance on the heuristic also opens an intriguing window on findings that general intelligence is negatively associated with prejudice (Hodson & Busseri 2012), an association that might be partially explained by essentialist thinking or use of the heuristic. Whether or not this is true, the heuristic enables a more cognitively sophisticated analysis of the essentialism–prejudice relationship.

The target article's account of the heuristic underpinnings of essentialist thinking also enables a clear understanding of the variability of psychological essentialism in the social domain. Essentialist beliefs employ diverse explanatory idioms and can be employed differently across contexts. This variability is entirely to be expected if essentialist thinking is conceptualized as a process of explanatory storytelling that works on often ill-formed understandings of inherent features. Such a sense-making process will rely on whatever explanatory resources are contextually or culturally salient, whether these be biological (e.g., blood, genes, or germs), spiritual (e.g., souls), or something else.

This view stands in stark contrast to the position that essentialist thinking involves the top-down application of a natural kind or folk-biological ontology. Similarly, if essentialist thinking, grounded in the inherence heuristic, amounts to an attempted solution to an explanatory puzzle, then we should not be surprised to see it employed flexibly, depending on the context of that puzzle. Social psychological research has made it clear that people express and hold essentialist beliefs to markedly different degrees to support particular social goals and interests. People may endorse gender essentialism only when it advantages their gender (Morton et al. 2009), vacillate between essentialist and nonessentialist understandings of an immigrant group's ethnicity depending on rhetorical purpose (Verkuyten 2003), and overcome their usual liberal reluctance to essentialize by seeing sexual orientation as biologically determined and fixed (Haslam & Levy 2006). Essentialist beliefs do not impose rigid, once-and-for-all ontological statuses on particular social categories, but can be applied malleably depending on the explanatory requirements of the context, as the inherence heuristic account would suggest.

That account may also have implications for the conceptual boundaries of psychological essentialism. The concept has been employed with varying degrees of stringency within social psychology, sometimes referring strictly to a species-like natural kind view of a social group (i.e., biological essence, discrete category boundary, immutable category membership) and sometimes more loosely to any imputation of underlying commonalities held by group members. Although the inherence heuristic is only one proposed contributor to psychological essentialism, as an important foundation it might help to delimit how far that concept can stretch. In particular, it would seem to imply that essentialism thinking must invoke inherent properties as causes. This rather minimal requirement is arguably not met by a recent argument for "belief in social determinism" as a component of essentialism (Rangel & Keller 2011), such a belief holding that people's character is shaped by extrinsic factors such as upbringing and social background. If we take inherent causal properties to be essential for essentialism, then this belief – which involves seeing a social group more as an artifact than as a natural kind – does not seem to qualify.

The inherence heuristic is a useful and potentially generative idea that awaits further study. Its definition of inherent features seems to me overinclusive, encompassing features that are superficial and secondary (e.g., the smell of citrus). The argument that inherent properties will also be seen as stable and enduring is

challenged by some of my own work (e.g., Haslam et al. 2000; 2002), which finds that beliefs in the inherence and immutability of social categories are unrelated. Nevertheless, the heuristic deserves serious consideration by psychologists who study essentialist thinking in the social domain.

# Quiddity and haecceity as distinct forms of essentialism

Bruce Hood

*School of Experimental Psychology, University of Bristol, Bristol BS8 1TU, United Kingdom.*
**Bruce.Hood@bristol.ac.uk**
**http://www.bristol.ac.uk/expsych/people/bruce-m-hood/overview.html**

**Abstract:** Psychological essentialism operates in two realms that have consequences for our attitudes towards groups and individuals. Although essentialism is more familiar in the context of biological group membership, it can also be evoked when considering unique artefacts, especially when they are emotionally significant items.

The authors have provided a plausible account of a cognitive bias operating early in development that forms an inchoate component of psychological essentialism. I would like to draw attention to two issues that I believe are pertinent to the article by Cimpian & Salomon (C&S). The first is that essentialism can actually be conceptualized by two apparently mutually exclusive principles that are related to identity. The first is the more familiar essentialist principle of *quiddity*, from the Latin *quidditas*, which translates as the "whatness" of some thing – in other words, the properties a thing shares with the other members of the category to which it belongs. Quiddity is most similar to the notion of Plato's *form*, where members of a category depart from some idealized version. In the target article, it is the essentialist principle of quiddity that is under consideration, because in order to infer an inherent feature to explain an observed pattern, one must be dealing with multiple members that constitute the group.

In contrast to quiddity, however, is the other essentialist notion of *haecceity*, from the Latin *haecceitas*, which translates as "thisness." Aristotle, critiquing his mentor Plato's notion of *form*, drew attention to the problem of reconciling the notion of unique identity with category membership. Medieval Scottish philosopher John Duns Scotus (from whose name we get the term *dunce*) introduced and championed the term haecceity. A good analogy from biology that helps to draw the distinction between these two essentialist principles would be the genome that determines which species membership category an animal belongs to (cat, dog, mouse, etc.) as the quiddity, and the DNA that is unique to the individual animal (Felix, Fido, or Mickey) as the haecceity.

These are not only distinctions in essentialism but they also manifest in naïve intuitions about nature of the world. As C&S have pointed out, essentialism is at the core of the understanding the mechanisms of inheritance, which explains why it makes evolution by natural selection so counter-intuitive. As biologist Ernst Mayr (1963) noted, if one intuits that the different species are essentially different, an identity that is passed on by inheritance, then it is problematic to accept that all life forms have a common original ancestor. Richard Dawkins (2010) calls this intuitive biological essentialism "the dead hand of Plato."

C&S have argued that artefacts do not exhibit the hallmarks of psychological essentialism, but this is not necessarily true if one considers haecceity. In the classic account attributed to Roman writer Plutarch, the *ship of Theseus* is a thought experiment that evokes haecceity. In the story, the Greek king's ship is replaced plank by plank over time so that eventually it contains none of the original material. Plutarch asked rhetorically whether the ship retained the same identity and, moreover, if one was to

reassemble all the pieces that had been removed into a second ship, which would be the ship of Theseus? The results of experimental studies with children and adults confirm that most individuals infer that individuals that are gradually replaced by component parts retain their identity (Hall 1998). This is because we infer an essential element in addition to the material composition when we are asked to consider the unique nature of things. Hall's (1998) developmental study revealed a stronger essentialist perspective for living things compared to an artefact, but we will essentialize objects that we consider significant by virtue of their unique identity if they have sentimental value (Hood & Bloom 2008). I have conjectured that this holds especially true for emotional objects such as memorabilia associated with individuals that we revere, from celebrities to religious saints, as well as so-called murderbilia, which are the items associated with murderers (Hood 2009). The authenticity of an object is similarly conferred by essentialist beliefs so much so that the value we place on objects such as artworks or collectibles is shaped by what we believe the object to be (Bloom 2010). For many, a perfect forgery indistinguishable from an original lacks some property that is difficult to articulate, consistent with the placeholder function that essentialism provides (Medin & Ortony 1989).

# The inherent bias in positing an inherence heuristic

Muhammad Ali Khalidi and Joshua Mugg

*Department of Philosophy, York University, Toronto, ON M3J 1P3, Canada.*
khalidi@yorku.ca        joshuamugg@gmail.com
http://www.yorku.ca/khalidi/
http://yorku.academia.edu/JoshuaMugg

**Abstract:** There are two problems with Cimpian & Salomon's (C&S's) claim that an innate inherence heuristic is part of our cognitive makeup. First, some of their examples of inherent features do not seem to accord with the authors' own definition of inherence. Second, rather than posit an inherence heuristic to explain why humans rely more heavily on inherent features, it may be more parsimonious to do so on the basis of aspects of the world itself and our relationship to it.

Cimpian & Salomon (C&S) present some intriguing preliminary evidence for the existence of an inherence heuristic, a basic cognitive tendency that leads people to explain patterns with reference to *inherent* features rather than *extrinsic* (i.e., relational or historical) features. While we find it plausible that people rely more heavily on inherent rather than relational properties in reasoning about many domains, we have doubts about the possibility of drawing the distinction between inherent and extrinsic properties unambiguously enough to enable us to conclude with confidence that participants are clearly tracking such a distinction in all the examples cited. But even if we set aside these doubts, the tendency that C&S are describing may not represent a cognitive bias of its own, but may instead emerge from the way the world is and our perceptual access to it.

C&S's account "classifies features as inherent if they can be said to characterize how an entity is constituted" (sect. 2.2.1, para. 1), and they add that these features tend to be stable and enduring. We take the distinction that they are tracking to be roughly that between intrinsic and extrinsic features – features that an individual object or entity has on its own and would continue to have in the absence of everything else, as opposed to those that an individual has in virtue of its relations to others. Paradigmatic cases of the former are perceptual features of an object, such as its size, mass, shape, or color. Clear examples of the latter are features that pertain to an object's location, position, relationships, or history, such as the fact that it is lying on top of the bookshelf, is located in Toronto, is my favorite toy, was manufactured in 2010, or

belongs to the public library. But though there seem to be many clear-cut cases, there are other features that may be trickier to classify in one or the other category.[1] This creates a few problems in the evidence that the authors rely upon, since a number of cases that they cite as instances of inherent features would seem in fact to be relational, extrinsic, or historical by C&S's own definition. For example, that orange juice has a tangy taste and that it is healthy, are both facts that pertain to the relation of orange juice to humans (or to some humans, since it may be unhealthy or taste differently to others). Thus, these features are arguably not inherent. Similarly, if Amy laughs at Beth's joke because Beth is funny, that fact is extrinsic to the joke (though it is an inherent feature of Beth). Meanwhile, in discussing extrinsic features of objects (such as historical features), C&S give the example of a router that stops working when accidentally stepped on. However, this episode in the history of the router presumably alters the inherent properties of the router, which is what prevents it from functioning. So it would be correct for a participant to say that the router stopped functioning due to an inherent defect, though that defect was caused by an event in its causal history.

The fact that some examples the authors cite of inherent features can be considered relational according to their own criterion, and vice versa, implies that it is hard to be sure what kinds of features participants are using in some of the experiments that are meant to support the authors' hypothesis. But if we set this worry aside and focus on those cases about which there may be little uncertainty, another concern emerges: If inherent features are roughly those that pertain to the individual or object taken in isolation, as opposed to ones pertaining to its relations, origin, history, and so on, then the former are the ones that tend to be perceptually salient to human beings and easily ascertainable. So it may not be a basic cognitive feature of humans, but rather a function of our relation to the world, that makes "inherent" features salient. Also, more often than not and in many domains, these features tend to be more explanatory than relational features. When it comes to the domain of physical objects, their motion, constancy, solidity, and so on (a domain that develops early in ontogeny), an object's length, width, shape, mass, density, material composition, texture, and other inherent features tend to be more explanatory of its patterns of behavior than its geographic location, ownership, and date of manufacture. This also holds to a large extent of the domain of living creatures. To be sure, when it comes to the domain of artifacts, extrinsic function tends to be more important than inherent features, and a chair can be made of a wide range of materials, can have various dimensions, material compositions, colors, and so on, yet remain a chair. But here, too, inherent features and function cannot drift too far apart. (How many chairs are made out of paper or are the size of a house?) Thus, given what is perceptually salient to human observers and given some broad features of the material world, it stands to reason that inherent features will be accessed more readily by cognizers and will have more explanatory power. If so, then there may be no need to posit a separate inherence heuristic to understand why cognitive agents reach first for inherent rather than extrinsic features to explain patterns in the world around them.

Finally, we cannot help entertaining the possibility that C&S fall prey to the inherence heuristic in positing an *innate* heuristic to explain certain human cognitive tendencies, rather than explaining them in terms of *relations* of human beings to the world. But then, wouldn't that be a dramatic confirmation of the very heuristic that the authors claim to observe? Not necessarily: We are arguing that, instead of a basic component of our innate cognitive endowment, our tendency to explain patterns on the basis of inherent features is instead a function of our relationship to the world and of features of the world itself.

NOTE
**1.** There is a debate in metaphysics on the proper characterization of the intrinsic–extrinsic distinction (e.g., see Langton & Lewis 1998; Lewis 1983; Vallentyne 1997). But that is not our concern here; rather,

we are concerned with whether people draw this distinction consistently enough to serve as the basis for a cognitive heuristic.

## Is it about "pink" or about "girls"? The inherence heuristic across social and nonsocial domains

doi:10.1017/S0140525X13003786

Katherine D. Kinzler and Kathleen R. Sullivan

*Department of Psychology, University of Chicago, Chicago, IL 60637.*
kinzler@uchicago.edu         krsullivan@uchicago.edu
http://dsclab.uchicago.edu

**Abstract:** The inherence heuristic provides an intriguing and novel explanation for early thought in a variety of domains. Exploring similarities and differences in inherent reasoning across social and nonsocial domains can help us understand the role that inherent thinking plays in the development of human reasoning and the process by which more elaborate essentialist reasoning develops.

Our drive to understand observed patterns in the world is pervasive and supports powerful learning throughout life. The inherence heuristic provides groundwork for this understanding both within everyday reasoning and across development, and has the potential to explain a wide range of psychological phenomena. We applaud the authors' thoughtful proposal. Yet, further specification of key aspects of the proposal – particularly regarding the domain specificity or the generality of inherence thinking – will clarify further the theoretical underpinnings of the heuristic and generate related research.

Are inherence beliefs about entities in the world (e.g., "pink as feminine") at all different from inherence beliefs about people or their psychological states (e.g., "girls like feminine things")? Cimpian & Salomon (C&S) state that both might be the case and may depend on the particular context at hand. We agree that understanding which patterns are subjected to the heuristic process is complex, yet we suggest that there may be important differences in the ways that people attend to, encode, and explain observed patterns in different domains. Are different kinds of evidence similarly susceptible to inherence reasoning? For instance, is it easier or harder to learn a new conceptualization of "pink" or of "girls," and are beliefs about people and non-person entities similarly resistant to change in the face of counterevidence? One possibility is that information about people may be particularly easily viewed as inherent, and thus it may be relatively easier to update a belief about the femininity of "pink" as compared to the femininity of "girls." Understanding how inherent reasoning is implemented across domains can be informative for understanding the development of children's reasoning about diverse concepts (e.g., people, animates, artifacts) and could also be informative about the functioning of the inherence heuristic more generally.

Relatedly, an inquiry into the domain specificity or generality of inherence thinking across social and nonsocial domains could shed light on the relationship between early inherence beliefs and later essentialist beliefs. Although intriguing, the proposal for the nature of the developmental change of essentialist reasoning as stemming from the inherence heuristic could be further specified. Is the change proposed to involve conceptual content that grows richer or conceptual content that is continuous over time yet whose exhibition requires the emergence of other supporting machinery? Social and nonsocial domains have the potential to differ not only in the way that observed patterns are weighed against existing inherent intuitions, but also in the way that inherence explanations are incorporated into essentialist explanations across development. C&S propose that inherent reasoning may indeed be overridden by counterevidence. They discuss the example of artifact categories: Children initially believe that artifact

categories derive from inherent features, but they may abandon that belief in the face of evidence that objects are constructed by people for particular functions. This example raises the general question of what guides children toward or away from inherent or essentialist beliefs across development, and whether attention to the evidence presented, the weighing process of evidence against intuitions, or both, might differ across domains. Selectivity in the patterns that are noticed and explained could play a crucial role in the workings of the inherence heuristic in and of itself and in the elaboration of some, but not all, inherent thinking into essentialist thinking.

We also question how critical to the theory is it that inherence reasoning applies more for patterns of behavior than for specific instances (or for information about groups of people rather than specific individuals). In theory, couldn't the inherence heuristic apply for both kinds of information? For example, if a child learned that someone is good at gorp, why not draw the inference that there is something intrinsic about him that allows him to gorp? Imagine an alternative pattern of results: Suppose children presented with both category-wide and individual-specific information endorsed inherent explanations – this hypothetical pattern of results could presumably also be interpreted as supporting the inherence heuristic. Yet, given the reported evidence that information about groups of people is more compelling than information about specific individuals, this finding may provide an opportunity to explore the relationship between inherence thinking and social categorization. It is plausible that persistent and coherent conceptual differences in reasoning across domains could result in differential explanations of patterns observed across people and patterns observed across objects.

Finally, the diversity of children's early environments and social experiences may have important influences on the development of the inherence heuristic across domains. The authors note that context and culture could guide the types of candidate explanations that become accessible to the heuristic. There are at least two ways in which early experience could guide inherent reasoning: through the available evidence and patterns to be explained that may differ across cultures, and through more pervasive individual differences that may vary across cultural contexts and could in turn influence the workings of the heuristics. To give one example, evidence suggests that bilingual children are more likely to see word-to-referent pairings as arbitrary (Bialystok 1988). Given the proposed link between beliefs about nominal realism and inherence, might children raised in diverse linguistic environments also be less susceptible to the inherence heuristic? It seems possible that both the process of learning two languages, as well as the diverse social experiences that accompany bilingualism or biculturalism, may influence children's inherent thinking. If so, the influence of such cultural diversity might be explored for both social and nonsocial domains of reasoning.

## Does the inherence heuristic take us to psychological essentialism?

doi:10.1017/S0140525X13003798

Anna Marmodoro,[a] Robin A. Murphy,[b] and A. G. Baker[c]

[a]*Faculty of Philosophy, Corpus Christi College, University of Oxford, Oxford OX1 4JF, United Kingdom;* [b]*Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom;* [c]*Department of Psychology, McGill University, Montreal, QC H3A 1B1, Canada.*
anna.marmodoro@philosophy.ox.ac.uk
robin.murphy@psy.ox.ac.uk         andy.baker@mcgill.ca
http://www.power-structuralism.ox.ac.uk/home
http://psy.medsci.ox.ac.uk/research/associative-learning-laboratory

**Abstract:** We argue that the claim that essence-based causal explanations emerge, hydra-like, from an inherence heuristic is incomplete.

No plausible mechanism for the transition from concrete properties, or cues, to essences is provided. Moreover, the fundamental shotgun and storytelling mechanisms of the inherence heuristic are not clearly enough specified to distinguish them, developmentally, from associative or causal networks.

Cimpian & Salomon (C&S) argue that humans generate causal/explanatory accounts of observed patterns by using the naïve reasoner's *inherent heuristic model* (IHM) and later the developmentally sophisticated *essentialist model* (EM). We worry that there is little to distinguish the IHM from existing psychological accounts relying on traditional cues to causality; further, there is no mechanism for the transition from IHM to EM.

C&S propose that a naïve reasoner selects features of objects that are *inherent* (i.e., properties of the here and now that are salient and immediately accessible or semantically associated). Using this information, the reasoner develops a causal/explanatory story about the observed pattern. Our questions are: How does the reasoner make up this causal story from these inherent properties? What makes it a *causal* story and not, simply, a sequence of juxtaposed salient properties? The IHM seems to require that the causal story preexist in the mind of the reasoner.

To bootstrap her story, the reasoner needs to use already established causal links, or something else, not included in C&S's theory. For instance, in the orange juice example, the reasoner builds a causal story associating the tanginess of the orange juice with waking. But why think that a tangy flavor rather than *any other* equally salient property of orange juice *causes* one to become alert? One way to answer this question is to assume the reasoner uses previously acquired associations to provide this causal link. But if so, building a causal story becomes an altogether different activity than the authors describe. Reasoners build causal stories from a rich repertoire of established causal links, seeking *explanatory fit* between the existing causal network and the new causal element. Before generating her story, the reasoner is already aware of many causal links between the inherent properties of the focal object. For instance, breakfast may be associated with the coldness of orange juice, its tanginess, its bright yellow color, or its sweetness. Why prefer tanginess to any other salient characteristic to build the explanation of why people drink OJ in the morning, unless the basis of this story was somehow already formed in the reasoner's mind? Furthermore, invoking semantic associations seems to allow significant room for story generation. Additionally, the IHM lacks reference to the significant body of research that illuminates how cues might acquire salience based not only on physical stimulus properties but learned and unlearned properties (e.g., Baker et al. 1996). For example, C&S reject the statistical predictiveness of a cue as particularly relevant; but research on association formation describes in detail the mechanisms related to predictiveness and to salience or associability that accompanies learning (e.g., Baetu & Baker 2009; Le Pelley et al. 2010; Mackintosh 1975; Murphy et al. 2009). An altogether different property of a cue involves preparedness (e.g., Garcia & Koelling 1966; Öhman & Mineka 2001). Certain cues go together better than others. For example, animals learn that flavor cues are much more readily associated with illness than are audiovisual cues. Is this because of an existing causal story? If not, why not? How would we distinguish it from one?

The second set of difficulties we wish to raise concerns the transition from the physical salient properties in the IHM to the essences in the EM. By *essence*, C&S mean "a certain internal, physical, microstructural *je ne sais quoi* that is unique to each kind and that invariably causes its members to display the full complement of typical features" (sect. 4.1, para. 1). We ask, how does the reasoner transition from believing that salient inherent properties to believing that internal, invisible essences have a causal/explanatory role? What relates a salient property to an essence? Possibly, the inherent features of an object picked out by the naïve reasoner are *part of* the essence that the sophisticated reasoner believes have a causal/explanatory role. Initially, C&S express this view but later claim that the part–whole relation is too formal to capture the real connection between inherent features and causal essences of things. But then, what is this connection?

C&S need to supply much more to bridge the gap between IHM and EM. The reasoner's transition from IHM to EM is not self-justifying. Why would the reasoner prefer causal explanations of observed patterns in terms of unobservable (and possibly undeterminable) essences? Is the transition from IHM to EM something reasoners find rationally compelling, or is it something they learn? What is gained by appealing to such mysterious entities as essences? Why assume that an object's relatively few essential properties explain its behavior? Why does the "inside" that accounts for the whole need to be a single property (the essence of a thing on C&S's understanding seems to be a "single inherent [internal] feature" [sect. 4.3, para. 2])? Instead, could the reasoner think that the many observable properties of a thing explain its few internal and unobservable ones? Would that still count as an EM and, if not, why not (see Oderberg 2007)?

Why is the EM a developmental advance over the IHM, when on the one hand the IHM requires more than an ability to pick out merely inherent properties, and on the other hand the EM does not tell us what distinctive and supposedly cognitively advantageous role that essences play for the reasoner when she provides causal/explanatory accounts of observed patterns?

As described by Tversky and Kahneman (1974), an heuristic is not simply a shorthand story but a testable model with a small number of variables that determine behavior. The model here is so unconstrained that any number of variables might be at work. C&S provide a descriptive phenomenology that does not elucidate the development processes involved.

## Systematic revisions to inherent notions may shape improvements in cognitive infrastructure

Rose Meleady[a] and Richard J. Crisp[b]

[a]*School of Psychology, University of East Anglia, Norwich NR4 7TJ, United Kingdom;* [b]*Department of Psychology, University of Sheffield, Sheffield S10 2TP, United Kingdom.*
**r.meleady@uea.ac.uk**        **r.crisp@sheffield.ac.uk**
**http://www.shef.ac.uk/psychology/staff/academic/richard_crisp**
**http://www.uea.ac.uk/psychology/People/Academic/Rose+Meleady**

**Abstract:** The proposed inherence heuristic centers on perceivers' failure to systematically consider external, historical factors when explaining observed patterns. We stress that this does not preclude the potential of subsequently encountered information to challenge intuitions. Drawing on models of diversity-defined social cognition, we discuss how an *updating mechanism* may reciprocally shape the cognitive infrastructure that underlies reliance on heuristic systems.

In their target article, Cimpian & Salomon (C&S) outline a new heuristic describing individuals' tendency to explain observed patterns as a function of their constituents' inherent features. C&S develop an important and interesting area for future inquest into the cognitive biases underpinning human judgement. However, we believe that certain areas of their model are underspecified. In particular, we contend that the model should be extended to include further specification of the computational mechanisms through which the heuristic system adapts in the face of information that *challenges* inherent notions.

As the authors note, it is a well-established finding in social psychology that the human mind has a tendency to prefer "the path of least resistance" (sect. 2.2, para. 10). Acting as *cognitive misers*, people are reluctant to expend cognitive resources by engaging in effortful processing and look to simplify social perception through the use of cognitive shortcuts (Fiske & Taylor 1991).

However, the appeal of heuristic efficiency is not absolute: although such systems can be adaptive when cognitive resources are limited, they leave us vulnerable to error. Accordingly, C&S acknowledge that the grip of the inherence heuristic on perceivers' judgements is not inescapable. They accept that intuitive judgements of causality can be *blocked* before they are generated. Furthermore, they outline how individual variability in cognitive style and ability, as well as developmental stage, may moderate tendencies to block the operation of the heuristic. They also accept that judgements arising from the heuristic can be *rejected* or *revised* after they have been generated – enabling potential bias and error to be corrected. However, the authors do not elaborate on *how* intuitions supplied by the heuristic can be revised in the face of challenging evidence. Here we outline a general mechanism through which systematic revisions to heuristic judgements can be made and consider how such revisions can reciprocally impact the cognitive foundations of the inherence heuristic.

Clues to the specification of the updating mechanism can be found in C&S's own application of inherence to essentialism in intergroup relations. As C&S discuss, the belief that differences between social groups are deep and immutable is associated with "stronger endorsement of stereotypes (e.g., Bastian & Haslam 2006; Levy et al. 1998, see also Yzerbyt et al. 2001), greater acceptance of racial inequities (Williams & Eberhardt 2008), and more blatant prejudice towards members of minority groups (e.g., Keller 2005; see also Haslam et al. 2002)" (sect. 4.1, para. 2). This role of inherence in forming and maintaining essentialist beliefs provides a point of synthesis with a recent, related model from the intergroup-relations domain which focuses precisely on what is lacking in C&S's model: a mechanism for how heuristic systems adapt in the face of information inconsistent with intuitive judgements.

We (Crisp & Meleady 2012) have argued that whereas people are disposed to think heuristically about social category boundaries (system 1), they must also possess the computational mechanisms to enable them to bypass this system when it is necessary to update and revise these representations (system 2). Without such a mechanism, heuristic systems would be inherently *mal*adaptive. Thus, when information that challenges extant category representations is encountered (e.g., positive and/or counterstereotypical outgroup behaviours), an *inconsistency resolution* process is engaged in which individuals inhibit the operation of category-based heuristic thinking in favour of higher-order, generative ways of thinking. This results in a creative, individuated impression of the individual encountered.

Notably, the results of studies supporting the *brain-as-muscle* metaphor demonstrate that, just as with physical exercise, repeatedly "working out" the brain literally improves its processing power. For example, tasks designed to train working memory improve fluid intelligence and are accompanied by reduced blood flow to areas necessary to complete previously cognitively demanding tasks (Jaeggi et al. 2011). Similarly, the results of longitudinal studies demonstrate improvements in inhibitory power with repeated exertion (Muraven et al. 1999). Accordingly, we argue that repeatedly encountering information that challenges existing preconceptions will train a disposition towards a more systematic, analytic cognitive style, lessening a general reliance on heuristic systems (Crisp & Meleady 2012; Crisp & Turner 2011).

A central claim by C&S is that the inherence heuristic underlies essentialist beliefs, which, in turn, are well established to contribute to intergroup biases. In line with the general principles embodied in the diversity-based model of social cognition outlined above, we contend that specification of the inherence heuristic should be extended. Specifically, the model could fruitfully incorporate a mechanism to account for how information that challenges intuitive inherent notions could lead to a lesser reliance on the inherence heuristic in the judgemental domains specified. In short, C&S tell us about individual differences in peoples' tendency to adopt this heuristic, but not how those individual differences arise. As the authors state, "The

more favorably one is disposed toward engaged, open-minded thinking; the less one is susceptible to the influence of intuitive heuristics" (sect. 2.2.5.1, para. 3). The model should therefore consider how the repeated engagement of an updating mechanism may reciprocally shape the cognitive infrastructure that underlies reliance on the inherence heuristic. An addendum to the model such as that outlined above would enable us to predict when and how generalized resistance to the inherence heuristic could lead to these individual differences.

Providing greater specification of how heuristic systems can be updated is critical because it may answer an important, yet unaddressed, question arising from C&S's proposition: How do we tackle these biases inherent to human social cognition? This is a fundamental question, because if C&S are correct in their assertion that an inherence heuristic lies at the core of multiple (problematic) biases in human judgement, we must seek ways to mitigate these biases in critical areas such as the environment, health, equality, and economics.

# Is the inherence heuristic needed to understand system-justifying tendencies among children?

Anna-Kaisa Newheiser and Kristina R. Olson
*Department of Psychology, University of Washington, Seattle, WA 98195-1525.*
**anewheiser@albany.edu      krolson@uw.edu**
**http://newheiser.socialpsychology.org/**
**https://depts.washington.edu/uwkids/krolson/**

**Abstract:** Evidence that children's system-justifying preferences track the extent of group-based status differences is consistent with the inherence heuristic account. However, evidence that children are inferring inherence per se, or that such inferences are the cause of system-justifying preferences, is missing. We note that, until direct evidence of the inherence heuristic is available, alternative models should not be ignored.

In their target article, Cimpian & Salomon (C&S) posit that people tend to explain patterns they observe based on features that are *inherent* in those patterns' constituents, and demonstrate how such an inherence heuristic may be the first step in a causal chain that leads to a variety of biases, beliefs, and attitudes observed across the field of psychology. Whereas the target article connects the inherence heuristic to many areas, we address its connection to our own area of expertise – social groups – focusing in particular on C&S's discussion of system justification (Jost & Banaji 1994). According to C&S, the inherence heuristic provides the cognitive underpinnings for, or enables, system-justifying patterns of thinking. As such, C&S note, the inherence heuristic should emerge early in childhood, and many of the cognitions resulting from the inherence heuristic by their very nature serve to perpetuate the system. Other aspects of system justification – for example its palliative (e.g., reassuring) functions – may emerge only later in development. Recent data from our lab, focusing on young children's racial attitudes, are consistent with C&S's account of early-emerging system-justifying tendencies. Here, we review our research and suggest alternative explanations for our findings, proposing that the inherence heuristic may in fact not be needed to understand system-justifying tendencies among children.

In our work, we have sought to understand how groups' relative status shapes children's race-based attitudes measured at the implicit (i.e., automatic or unconscious) level. Consistent with the inherence heuristic account, we predicted that children would be sensitive not only to the existence but more specifically to the *extent* of status differences among racial groups in their society. Focusing first on the American cultural context, we

found that although White American children aged 7–11 years showed a robust implicit preference for Whites over Blacks, Black American children on average showed a lack of implicit bias (appearing, on average, to prefer neither Blacks nor Whites; Newheiser & Olson 2012) – directly paralleling prior findings with American adults (e.g., Nosek et al. 2002). This pattern is remarkable, albeit unfortunate, insofar as it suggests that minority-group children as young as age 7 years are influenced by the low status that society attributes to their racial in-group. Furthermore, we have conceptually replicated and extended this finding cross-culturally: Turning to South Africa, a society in which race-based differentials in status and advantage are extremely pronounced – much more so than in the United States – we found that 6- to 11-year-old Black and Coloured (i.e., multiracial) South African children implicitly favored Whites, the highest-status racial group in their society, over their own in-groups (Newheiser et al. 2014). Both Black and Coloured children also implicitly preferred Coloureds, an intermediate-status group, over Blacks, representing an out-group preference among Black South African children (Dunham et al. 2014). Thus, whereas race-based status differentials canceled out implicit in-group preference among Black American children, they resulted in a full reversal to *out-group-favoring* implicit biases among South African minority-group children.

Our findings align with the inherence heuristic account insofar as they demonstrate that system-justifying cognition that tracks the degree of group-based status differences can indeed be observed among children. In terms of C&S's account, our findings might be interpreted as suggesting that as children observe a pattern indicating that members of certain racial/ethnic groups appear to be more advantaged than members of other groups, children come to infer that this pattern occurs because members of advantaged racial/ethnic groups possess characteristics that make them inherently "better" – that is, that certain racial/ethnic groups *should* have higher status by virtue of their inherent features (see also Kay et al. 2009).

Whereas our data are potentially consistent with C&S's general argument, we note that the specifics of the inherence heuristic account remain in need of direct evidence. In particular, we question whether children are in fact making inferences regarding the *inherent* features of racial/ethnic groups that differ in status. An alternative, and simpler, account relies on mere cognitive associations that may emerge as children observe the world around them. That is, as children encounter members of different racial/ethnic groups, in their everyday lives or via the media, they likely notice the covariation between race/ethnicity and status (e.g., in terms of wealth; Olson et al. 2012). Forming such associations between racial groups and status might directly result in the lack of implicit race bias we observed among Black American children, and in the implicit out-group biases we observed among South African minority-group children; an inference regarding inherent features may not be necessary at all.

To conclude, the inherence heuristic account argues that people tend to interpret patterns they observe in terms of inherent features, and that this tendency enables system-justifying cognition (among other psychological phenomena). However, we propose that merely observing relationships between group membership and a valenced characteristic (e.g., status) will result in preferences and beliefs that are consistent with those relationships. For example, observing that Group A is low in status may lead one to disfavor Group A simply because status is a highly valued characteristic; an additional inference that something inherent about Group A makes it low in status is not necessary. Moreover, even if one does eventually make such an inference about inherent features, this inference is plausibly a *consequence* rather than a cause of system-justifying beliefs and preferences – perhaps serving to justify such beliefs and preferences. It thus seems neither necessary nor sufficient to posit the inherence heuristic as a precursor to system-justifying or hierarchy-attenuating beliefs and preferences. We note that developmental research is

in a unique position to help pit these alternative accounts against each other by providing evidence of the sequence in which these, and other related, processes first emerge.

## Owning up to the role of historical information

Nicholaus S. Noles and Judith H. Danovitch

*Department of Psychological and Brain Sciences, University of Louisville, Louisville, KY 40292.*

**n.noles@louisville.edu**      **j.danovitch@louisville.edu**
**http://louisville.edu/psychology/noles**
**http://louisville.edu/psychology/danovitch**

**Abstract:** Although the inherence heuristic is a versatile cognitive process that addresses a wide range of psychological phenomena, we propose that ownership information represents an important test case for evaluating both the boundaries of Cimpian & Salomon's (C&S's) model (e.g., is the inherence heuristic meaningfully limited to only inherent factors?) and its effectiveness as a mechanism for explaining psychological essentialism.

Cimpian & Salomon (C&S) present the inherence heuristic as a surprisingly simple and intuitive overarching cognitive process that accounts for a broad array of disparate research findings. C&S describe this process as addressing any feature that "can be said to characterize how an entity is constituted" (sect. 2.1.1, para. 1), and they define these features as salient, stable, accessible, intrinsic, internal, and nonobvious. Notably, the inherence heuristic does not address extrinsic factors, including historical information, and the justification for the exclusion of these factors is that they are not very salient or accessible. C&S use historical events and social conventions as two primary examples of extrinsic factors that may be reasonably excluded from consideration. However, we propose that the concept of ownership represents an extrinsic factor that poses two critical challenges to the present formulation of the inherence heuristic.

The first challenge is that historical information plays a critical role in identifying and tracking property (Gelman et al. 2012). Ownership is abstract and nonobvious, like inherent features of entities, but, unlike historical events and social conventions, ownership is highly accessible and salient, even to very young children (Fasig 2000; Hay 2006; Levine 1983; Ross et al. 1990; Saylor et al. 2010). Whereas it may be reasonable to conclude that other extrinsic factors fail to be activated by the mental shotgun, ownership information has precisely the characteristics that should cause it to loom large in any narrative constructed by the storytelling component of the process. There are also thematic similarities between the output of intuitions about ownership and the explanations generated by the inherence heuristic. For instance, the output of both processes appears to be effortless despite resulting from a great deal of rapid, implicit cognitive processing.

The second challenge that ownership presents with respect to the inherence heuristic is that it provokes individuals to perceive the inherent, essencelike qualities of individuals in objects. At the most basic level, giving an individual a piece of property enhances that person's perception of its value (Beggan 1992; Gawronski et al. 2007; Kahneman et al. 1990; Thaler 1980). In addition, the knowledge that an object has been in contact with a specific individual or special set of circumstances often results in intuitions that the object itself is contaminated or imbued with the essence of these individuals and events (Rozin & Nemeroff 1990). In this way, art may be appreciated because of its special, historical connection to an artist (Bullot & Rebar 2013; Gelman et al. 2013), and learning that a set of old, used golf clubs once belonged to JFK might dramatically inflate an individual's attraction to and valuation of the clubs, a response that can not be explained away completely by simple associations or market demands (Newman et al. 2011). In this way, objects with no special status are granted authenticity, a nonobvious

historical property that makes certain objects special. Critically, children as young as age 4 reason about authentic objects in an essentialist manner (Frazier & Gelman 2009; Gelman et al. 2014). Because the essencelike effects of ownership appear early and persist into adulthood, it is unlikely that they are some later by-product of psychological essentialism. Thus, at least some element of historical information must be accounted for by any comprehensive explanation of psychological essentialism.

Historical information plays an important role in considerations of ownership and authenticity, a role that highlights several gray areas with respect to the inherence heuristic. First, is the separation between intrinsic and extrinsic information detailed by C&S meaningful? There are at least two effects of this dichotomy on the model: (1) a separation between internal features (e.g., DNA, insides, essences) and external features (e.g., conventions, surface details), and (2) a division between features that are salient and accessible and those that are not. These two elements travel together in the examples presented by C&S, but ownership is a case where the two might be split apart, unless one conceptualizes ownership in a very specific way (i.e., as an intrinsic feature that is historically, and therefore externally, defined – but whether this is possible is an empirical question).

Second, can historical information play a role in the output of the inherence heuristic? Historical information plays a critical role in defining property (Gelman et al. 2012) and authentic objects (Frazier & Gelman 2009). Ownership information is accessible and salient in exactly the way that other kinds of historical information, such as social conventions and historical events, are not. In its current formulation, the inherence heuristic states that the mental shotgun selects inherent features because they are salient and accessible, but there is no requirement that they be internal or intrinsic. Although it may be true that inherent features are overrepresented in human explanation, perhaps because they are salient and accessible, does the model need to be limited to inherent features?

Finally, can the inherence heuristic effectively serve as the foundation for psychological essentialism if it accounts only for essentialist thinking within a subset of domains? The inherence heuristic is a process that is proposed to explain broad patterns, and the relationship between individuals and property is consistent and systemic. More specifically, people treat property as though it is imbued with the essence of current and prior owners. Historical information is central to concepts of ownership, and historical path may be an important element of essentialism, especially with respect to artifacts (see Frazier & Gelman 2009). However, in its current formulation, the inherence heuristic does not account for authenticity and other effects of ownership. Thus, the inherence heuristic appears to be well situated to address how we think about the intrinsic properties that underlie many human intuitions, but if this process is truly the foundation for psychological essentialism, then it needs to address historical information in some contexts.

# The social aetiology of essentialist beliefs

Cliodhna O'Connor and Helene Joffe
*Division of Psychology & Language Sciences, University College London, London WC1H 0AP, United Kingdom.*
oconnor.c@ucl.ac.uk          h.joffe@ucl.ac.uk
https://iris.ucl.ac.uk/iris/browse/profile?upi=COCON80
http://www.ucl.ac.uk/psychlangsci/research/CPB/people/cpb-staff/h_joffe

**Abstract:** This commentary highlights the importance of attending to the sociocultural contexts that foster essentialist ideas. It contends that Cimpian & Salomon's (C&S's) model undervalues the extent to which the development of essentialist beliefs is contingent on social experience. The result is a restriction of the model's applicability to real-world instances of essentialism-fuelled prejudice and discrimination.

Psychological essentialism is a pervasive pattern of thinking that has been implicated in a host of unpleasant societal processes. As such, the endeavour of Cimpian & Salomon (C&S) to elucidate the roots of this psychological tendency is an important undertaking. However, we argue that the model they have developed sustains one clear void: Namely, it underplays the influence of social factors on the development of essentialist beliefs, particularly when those beliefs pertain to social groups.

C&S contend that the origins of psychological essentialism lie in the "inherence heuristic," which they present as a product of basic, potentially innate properties of the human cognitive system. Within their model of the development of essentialism, social factors are relegated to a subsidiary role. The basic tendency to make essentialist assumptions is cognitively prescribed, and social experience simply furnishes the surface content that allows these essentialist principles to be developed into culturally sensible narratives. However, this jars with the extensive body of theoretical and empirical literature that suggests that the process, as well as the content, of thought is socio-culturally shaped (Fiske et al. 1998). As children mature, they are socialised into a *community of minds* from whom they learn the appropriate ways of representing the world around them. This premise opens the possibility that patterns of essentialist thinking may be acquired, or at least formatively moulded, as a result of people's encounters with the social world. Indeed, research shows that the propensity to engage in essentialist thought (as well as the precise content of essentialist ideas) varies systematically across cultural groups, including those differentiated by religion (Birnbaum et al. 2010; Diesendruck & Haber 2009), nationality (Mahalingam & Rodriguez 2003), social status (Mahalingam & Rodriguez 2006), political conservatism (Rhodes & Gelman 2009), and race (Jayaratne et al. 2009). This suggests that social experience can affect the very impulse to think in essentialist terms, rather than merely supplying the narratives that are tacked onto the fundamental, cognitively given precepts.

C&S acknowledge the proposition that essentialist thinking is a product of socialisation, but dismiss this hypothesis as implausible. The sole empirical grounds provided for this position relates to the failure of research to detect sufficient evidence of overt transmission of essentialism within parent–child conversations. Verbalised parent–child interaction is, however, an extremely narrow proxy for the whole spectrum of social influences on individual cognition. Engagement with the social world occurs through many channels – through institutions such as the mass media, education, commercial markets, and politics, as well as immediate verbal and non-verbal interpersonal interactions. All such means of communication are viable vectors of essentialist ideas.

As an example of one such channel of social influence, research has suggested that the popular diffusion of scientific knowledge can propagate essentialist representations of particular social groups (Dar-Nimrod & Heine 2011). Recent research on media coverage of neuroscience shows that a key way in which neuroscience is communicated to the public is via emphasising that differences between social groups are biologically rooted (O'Connor et al. 2012). This neuro-essentialism of social categories is exemplified by the ubiquitous trope of "the [*adjective*] brain" (e.g., "the male brain," "the gay brain," "the criminal brain"), which expressly condenses the essence of a given category into the notion of a distinctive brain type that is universally shared by all category members. The proliferation of such messages within the public sphere may trigger essentialist thinking in those who encounter them. Experimental research has shown that exposing people to information about the purported biogenetic foundations of, for example, gender (Brescoll & LaFrance 2004; Coleman & Hong 2008) or race (Keller 2005; No et al. 2008; Williams & Eberhardt 2008) exacerbates essentialist beliefs and corresponding processes of stereotyping and discrimination. Though C&S acknowledge that biological information can be recruited into the "stories" that elaborate inherence-based assumptions, they do not consider whether exposure to

this biological information might itself cultivate essentialist beliefs, which would not otherwise have emerged.

The model's relative neglect of social influence restricts its applicability to real-world instances of essentialism-fuelled prejudice and discrimination. Although the authors acknowledge that the *content* of essentialist beliefs can deviate between category domains, they do not engage with the evidence that different human categories attract varying *levels* of essentialism (Haslam et al. 2000; Joffe & Staerklé 2007). An account of essentialism that hinges on a basic cognitive bias cannot elucidate why, for example, categories relating to eye colour do not elicit the level of essentialism that is directed at categories relating to gender or race. Without due attention to the modulating influence of societal processes, the model cannot address the key question of why particular social groups, within particular historical contexts, become constituted as homogeneous, immutable, and denigrated entities. Neither does it aid in discriminating between those contexts in which essentialist identities are internalised by marginalised groups, and cases where oppressive essentialist identities are actively challenged and resisted (Mahalingam 2007). Indeed, characterising essentialism as a product of instinctive cognitive mechanisms may lend an air of inevitability, obscuring the processes by which change can be effected in the essentialism that particular groups in society face.

Undoubtedly, essentialist beliefs are premised upon cognitive (and affective) patterns, and these merit empirical elaboration. However, social contexts shape how these basic cognitive tendencies manifest and hence their implications for individuals and society. In building an account of essentialism that foregrounds instinctual cognition to the exclusion of formative social factors, C&S's model may itself exemplify an inherence-based explanation. It thereby risks eliding the social interests and agendas that selectively channel the malignancies of essentialism towards particular sectors of society.

# Is psychological essentialism an inherent feature of human cognition?

Christopher Y. Olivola[a] and Edouard Machery[b]

[a]*Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213;* [b]*Department of History and Philosophy of Science, University of Pittsburgh, Pittsburgh, PA 15260.*

colivola@andrew.cmu.edu       machery@pitt.edu
https://sites.google.com/site/chrisolivola/
http://www.pitt.edu/~machery/

**Abstract:** Recent evidence shows that psychological essentialism is neither a universal nor stable feature of human cognition. The extent to which people report essentialist intuitions varies enormously across cultures and education levels, and is also influenced by subtle, normatively irrelevant contextual manipulations. These results challenge the notion that the human mind is "fitted" with a built-in inherence heuristic that produces essentialist intuitions.

A major goal of the article by Cimpian & Salomon (C&S), as stated by the authors themselves (see sect. 4), is to argue that their proposed inherence heuristic provides the cognitive foundation on which psychological essentialism emerges. In doing so, the authors seem to take it as a given that psychological essentialism is a fundamental feature of the human mind that needs to be explained.

However, recent evidence suggests that psychological essentialism may not be a universal or immutable human tendency. In a recent set of studies (Machery et al., submitted), we examined the stability of psychological essentialism. The participants in our studies were asked, for example, whether lemons that undergo a chemical transformation, which alters their exterior properties (e.g., taste and appearance) but conserves their genetic structure, would still be considered lemons, or

(conversely) whether a liquid discovered on a distant planet that shares all the functional and superficial properties of water, despite having a completely different molecular structure, should be called "water." An essentialist position would define a category of things (or kind) by their stable, underlying properties. Thus, according to a true essentialist, lemons that undergo a superficial transformation while retaining their genetic structure would still be lemons, whereas the alien liquid would not be considered water (since its molecular structure is different) regardless of how closely it resembles, tastes like, or fulfills the same role as water. Thus, by tallying the proportion of participants who believe the lemon remains a lemon in the chemical transformation scenario, or that the alien liquid is not water in the discovery scenario, we can estimate which proportion of a given sample holds essentialist beliefs.

The first study considered essentialist intuitions in 10 different nations spread across four continents. We found that the likelihood of holding essentialist intuitions varied dramatically (and significantly) across cultures, from less than 10% of participants in a particular country holding essentialist beliefs to more than 70% of those in another country holding these intuitions. Furthermore, these intuitions were found to vary significantly with gender and education. A second study showed that the type of scenario (transformation vs. discovery) and the category (organic/biological vs. physical/chemical) both significantly (and independently) influenced the likelihood that participants reported essentialist intuitions. Finally, a third study showed that merely varying how we framed the (same) question in a transformation scenario significantly altered the likelihood that participants would report essentialist intuitions.

Taken together, the aforementioned studies show that, far from being universal, essentialist beliefs vary dramatically across cultures, gender, and educational levels – much like many other philosophical intuitions (e.g., Machery et al. 2004; 2009). These studies also show that, far from being stable, the likelihood that people hold essentialist intuitions varies considerably with normatively irrelevant contextual factors, such as the type of scenario used to elicit beliefs or the way the question is being framed.

The results of our studies are consistent with a small, but important, literature on cross-cultural variation in psychological essentialism and on the role of social factors in the prevalence of essentialism. There is substantial variation regarding which properties are thought to be immutable as well as caused by internal features. Pfeffer et al. (1998) found that British children are more likely than Nigerian children to view criminal behavior as immutable and internally caused. Giles et al. (2008) also found that Black South African children are more likely than African-American children to view aggression as inborn and immutable. Giles et al. argue that this finding is consistent with the social role of essentialism in justifying inequalities during apartheid (see also Mahalingam 2003b). There is also substantial variation across cultures and social groups regarding which categories are essentialized. Mahalingam (2003a) has shown that the Aravanis of India (a minority of biological men who dress and live as women) essentialize the female gender, but not the male gender, holding that a male can become a woman by cross-dressing. Mahalingam (2003b) has also provided evidence suggesting that Brahmins and Dalits in India essentialize the Indian castes differently. It would thus seem that, far from being a default mode of thinking, essentialism is used strategically, in a contextual manner, often to fulfill some political goals (see also Pereira et al. 2010).

In sum, a growing of body research suggests that psychological essentialism is neither a universal nor stable feature of human cognition. Consequently, this research challenges the notion that the human mind is "fitted" with a built-in inherence heuristic that produces essentialist intuitions.

## The representation of inherent properties

Sandeep Prasada

*Department of Psychology, Hunter College, City University of New York (CUNY); New York, NY 10065.*
sprasada@hunter.cuny.edu

**Abstract:** Research on the representation of generic knowledge suggests that inherent properties can have either a principled or a causal connection to a kind. The type of connection determines whether the outcome of the storytelling process will include intuitions of inevitability and a normative dimension and whether it will ground causal explanations.

The inherence heuristic captures a deep characteristic of commonsense thought. As currently formulated, however, it is limited by an extremely cursory characterization of inherent features and patterns. A more robust characterization of the patterns that are grist for the inherence heuristic, as well as how inherent features are represented, would greatly improve the heuristic's ability to predict the types of intuitions that may be generated when explaining a specific pattern. Recent work on the representation of generic knowledge can help provide this more robust characterization.

It is noteworthy that the patterns discussed by Cimpian & Salomon (C&S) involve kinds of things such as boys, girls, and orange juice. This is not an accident. That all coins in my pocket on a given day are copper does not constitute a pattern that can be explained via the inherence heuristic. Such a pattern involves an accidental generalization and thus cannot be explained, much less explained via inherent properties of the coins in my pocket or my pocket. Thus, a minimal condition for the operation of the inherence heuristic is that the pattern be one that involves a nonaccidental generalization and thus is extendable to indefinitely many new instances of the relevant sort (Goodman 1955/1983).

Prasada et al. (2013) provide evidence that our conceptual systems distinguish at least three types of nonaccidental connections between kinds (which contain indefinitely many instances) and properties. Kinds may be characterized by properties that have a principled, causal, or statistical connection to the kind. Each connection type grounds distinct linguistic and nonlinguistic phenomena and reflects a distinct perspective from which we can think about kinds of things. Properties that have a principled or a casual connection to a kind may plausibly be considered inherent properties of the kind when thinking about kinds from a formal or material perspective, respectively, and ground distinct types of inherent thinking.

Properties that have a principled connection to a kind are properties that instances of a kind are understood to have by virtue of their being the kinds of things they are (k-properties). *K-properties* are properties (1) whose presence in instances of a kind support formal explanations – explanations by reference to the kind of thing something is (e.g., Fido has four legs because he is a dog); (2) for which we have normative expectations such that instances of the kind that lack them are judged to be defective or incomplete; and (3) are generally expected to be present in instances of the kind (Prasada & Dillingham 2006; 2009). K-properties are understood to be an aspect of being the kind of thing in question and thus are represented via a formal part–whole relation between the kind and property (Prasada & Dillingham 2009).

Attending to principled connections brings forth the formal dimension of our commonsense conceptions, and we notice formal explanatory relations, as well as the basis for certain normative and statistical expectations. In so doing, k-properties can ground key aspects of inherent thinking identified by C&S. Specifically, the intuition that the pattern is inevitable reflects our expectation that k-properties will generally be present in instances of the kind. K-properties also ground the normative dimension of much inherent thinking – dogs are supposed to have four legs

and the ones that don't have something wrong with them. Note that this is stronger than simply the intuition that something is good or beneficial. We think that it is beneficial for dogs to wear collars, but we don't think there is anything wrong with dogs that don't (Bublitz & Prasada 2013).

Attending to casual connections between kinds and properties brings forth the material dimension of our commonsense conceptions by focusing on the material constitution of the instances of kinds and thus their casual dispositions to behave in one or another manner in appropriate circumstances. Properties that merely have a casual connection to a kind do not ground an expectation that they are generally present in instances of the kind, nor do they ground normative expectations about their presence in instances of the kind (Prasada et al. 2013). As such, though properties that have a casual connection to a kind may be considered inherent properties of the kind and can ground causal explanations of patterns, the intuitions of inevitability and normativity associated with much inherent thinking cannot derive from our understanding of these properties having a casual connection to the kind.

To account for the range of characteristics associated with inherent thinking discussed in the article by C&S, the "storytelling" mechanism must crucially have access to information as to whether a property is considered an inherent property because it is represented as an aspect of being that kind of thing and thus has a principled connection to the kind or if the property is considered an inherent property of the kind because it is casually connected to its material constitution. The type of connection represented between the kind and the property is needed to determine whether the outcome of the storytelling process will include intuitions of inevitability and a normative dimension and whether it will ground causal explanations. Many properties will have both principled and causal connections to a kind (e.g., the four-leggedness of dogs)

C&S are correct that the work on k-properties does not provide an alternative to the inherence heuristic; however, if the present argument is correct, it is a necessary complement to the inherence heuristic. Incorporating the insights from the work on the representation of generic knowledge into the inherence heuristic promises to be a fruitful avenue for future research. Detailed work on the representation of different kind of "habituals" (e.g., *John takes the train to work* vs. *John prefers blonds*) will also likely help inform how the inherence heuristic functions when reasoning about instances rather than patterns. Finally, progress on specifying the scope of the inherence heuristic can likely be made by detailed investigations of the formal characteristics that distinguish kind representations from representations of other types of multiplicities (Prasada et al. 2012).

## The developmental and evolutionary origins of psychological essentialism lie in sortal object individuation

Hannes Rakoczy[a] and Trix Cacchione[b]

*[a]Department of Developmental Psychology, University of Göttingen, D-37073 Göttingen, Germany; [b]Institute of Psychology, University of Bern, CH-3012 Bern, Switzerland.*
hannes.rakoczy@psych.uni-goettingen.de      trix.cacchione@psy.unibe.ch
www.psych.uni-goettingen.de/en/development/team/rakoczy-hannes
http://www.psy.unibe.ch/entwicklung/content/fvp/team/cacchione/index_eng.html

**Abstract:** Cimpian & Salomon (C&S) present promising steps towards understanding the cognitive underpinnings of adult essentialism. However, their approach is less convincing regarding ontogenetic and evolutionary aspects. In contrast to C&S's claim, the so-called inherence

heuristic, though perhaps vital in adult reasoning, seems an implausible candidate for the developmental and evolutionary foundations of psychological essentialism. A more plausible candidate is kind-based object individuation that already embodies essentialist modes of thinking and that is present in infants and nonhuman primates.

The approach described by Cimpian & Salomon (C&S) presents very promising steps towards a better understanding of the cognitive underpinnings of and interindividual differences in intuitive reasoning and explanation seeking by adults. However, it is less convincing as an explanatory approach for the developmental and evolutionary foundations of essentialist reasoning. One of C&S's main claims – that the so-called inherence heuristic (IH) is the developmental foundation for psychological essentialism – has little plausibility for at least two reasons: *First*, the IH is built on a complex inferential machinery (e.g., the storyteller) that might not be available early in development (and evolution), among other reasons because it appears to rely heavily on linguistic capacities. *Second*, C&S consider psychological essentialism to be a late-developing phenomenon. In their view, essentialism first appears around age 4, when children master verbal tasks that require them to distinguish between essential and superficial features of animals and other objects. In classical transformation and adoption vignettes ("Will a squirrel painted like a raccoon and growing up among raccoons turn out to be a squirrel or a raccoon?"), children from this age base their judgment of the identity of animals exclusively on essential features and neglect superficial ones (Keil 1989).

Unfortunately, however, C&S fail to take notice of earlier cognitive capacities that share some of the essential properties of essentialist reasoning and should thus be considered the ontogenetically and evolutionarily primary forms of psychological essentialism. In particular, even human infants and nonhuman primates engage in a form of object individuation – kind-based or sortal object individuation – that embodies essentialist modes of thinking. Conceptually, *sortal object individuation* is the capacity to conceive of objects as objects of certain kinds X, Y, Z (dogs, cats, bananas…) by using sortal concepts ("X," "Y," "Z" – "dog," "cat"…) that supply criteria for individuation ("How many Xs are here?") and identity ("Is this the same X that I saw before?"). Empirically, this capacity has been studied by confronting infants and apes with, say, a box into which an X enters at time 1, followed at time 2 by a Y coming out of the box. Subjects' numerical expectations as to how many objects are in the box are then measured, as indexed in looking and searching behavior: If they think of the objects as objects of distinct kinds that cannot turn into each other, they should expect that there must be (at least) one object, the X, still in the box.

Basic versions of such tasks in which an X (e.g., a banana piece) and a Y (e.g., a carrot piece) are used that differ both in essential and in superficial properties are mastered by human infants from around 12 months (Xu 2007) and by nonhuman primates (Mendes et al. 2008; 2011; Phillips & Santos 2007). Because these findings – due to the confound between essential and superficial feature differences – remain somewhat inconclusive regarding the question of whether infants and primates really solve these tasks by using sortal concepts and essentialist reasoning, we recently designed a modified version of individuation tasks. This version was inspired by the classical verbal essentialism tests and enabled us to deconfound essential and superficial property differences: Infants and nonhuman primates saw one object with surface features SF1 enter into a box at time 1 and then at time 2 an object with different surface features SF2 coming out of the box. In fact, however, the object with SF1 was identical to the object with SF2. (For example, in the infant studies, there were stuffed toy animals that could be turned inside out, appearing as a bunny in one form and as a carrot in the other.) The crucial variation was whether the subjects were aware of this. The results showed that those infants and apes that were unaware of the dual identity of the object took the superficial feature differences as diagnostic for their numerical expectations:

They thought that there still must be an object in the box – as indicated by their significantly longer searching in the box as compared with events where the superficially identical object with SF1 appeared at time 2. In contrast, those infants and apes that knew about the object's dual identity disregarded the superficial feature differences, did not expect a second object in the box, and did not search differently in the two events. That is, given the requisite background knowledge, infants and apes disregarded the superficial feature differences in very much the same way that older children disregard the superficial feature differences between a normal squirrel at time 1 and a raccoon-looking squirrel at time 2 (after it has been painted etc.) when it comes to the question of the animal's identity (Cacchione et al. 2013; submitted).

So, what these studies – together with other infant studies (Newman et al. 2008) and primate work (Phillips et al. 2010) – suggest is that basic forms of psychological essentialism appear to be present much earlier in development and evolution than assumed by C&S. And this has important implications for the type of account C&S are putting forward – suggesting the following slightly different picture:

1. The developmentally and evolutionarily primary forms of psychological essentialism lie in sortal object individuation: Distinguishing between essential properties of an object that determine criteria of identity and countability, and merely superficial features that do not, is already the simplest form of essentialist reasoning. From an evolutionary perspective, such a basic distinction between deep and superficial properties seems to be a fundamental design feature of higher cognition given the need to distinguish between differences in essential (and therefore identity-preserving) properties and merely superficial properties in so many domains (e.g., regarding natural food items that constantly change form and color, or regarding conspecifics that change appearances by growing and decaying).

2. Out of this general capacity, then, more specific forms of conceiving of the (essential vs. merely superficial) properties that are relevant for a given kind of object emerge, possibly based on some domain-specific sensitivities to which kinds of properties might be relevant for which type of object.

3. Over development, subsequently, the IH – as a broader and more complex cognitive tendency– might emerge out of and hook onto this developmental basis.

# Inherence-based views of social categories

Marjorie Rhodes

*Department of Psychology, New York University, New York, NY 10003.*
**marjorie.rhodes@nyu.edu**
**http://www.psych.nyu.edu/rhodes/**

**Abstract:** Children adopt an inherence-based view of some social categories, viewing certain social categories as reflecting the inherent features of their members. Thinking of social categories in these terms contributes to prejudice and intergroup conflict. Thus, understanding what leads children to apply inherence-based views to particular categories could provide new direction for efforts to reduce these negative social phenomena.

Cimpian & Salomon (C&S) present a compelling description of how a fundamental cognitive bias – to explain observed patterns as resulting from the inherent features of the involved entities – underlies an impressive range of cognitive and social phenomena. From a developmental perspective, it is easy to see how such a heuristic might be useful in early conceptual development. By allowing children to expect stability in their environment, the inherence heuristic simplifies children's immense learning challenge. Yet, as discussed by the authors, the inherence heuristic may also have deleterious consequences, particularly when

learners use it to explain social phenomena. Here, I consider the implications of the inherence heuristic for the development of social categorization.

Preschool-age children are rampant social categorizers – they readily categorize people based on a wide range of criteria (e.g., gender, race, age, languages, shirt colors, and labels) (Dunham et al. 2011; Kinzler et al. 2010) and use these categories to explain and predict social behavior (Diesendruck & HaLevi 2006; Gelman et al. 1986; Taylor et al. 2009; Waxman 2010). Prior research on the development of social categorization has examined a component of children's social categories that closely resembles – and indeed may reflect – the inherence heuristic: a belief that certain categories reflect the objective natural structure of the world. This view of social categories suggests that certain categories reflect the inherent features of their members (e.g., that boys and girls are categorized separately because they are inherently different from one another).

Of all the possible ways that people can be categorized, children appear to adopt this inherence-based view to explain an important subset of social categories. Illustrating this component of children's categories, Rhodes and Gelman (2009) asked children to consider whether a pairing composed of a boy and a girl, for example, could be considered the "same kind of person." Even though children could readily bring to mind many similarities between the individuals – shared age, race, facial expressions, and so on – they overwhelmingly rejected these categories. Children in this study viewed the decision to classify people by gender as reflecting an underlying natural reality (the inherent nature of the categorized entities) instead of as the consequence of social histories, personal choices, or convention. Thus, children took an inherence-based view of the structure of social categories.

Whether a particular social category elicits these inherence-based beliefs varies across development and cultural contexts. In the article by Rhodes and Gelman (2009), for example, young children viewed gender categories, but not racial categories, as reflecting an objective natural reality. In this work, older children's beliefs varied by cultural context; older children growing up in an ethnically diverse and liberal community did not view either gender or race in these terms, whereas older children growing up in an ethnically homogeneous and conservative environment viewed both categories in this manner. Further, Diesendruck and colleagues (2013) found that the belief that religious groups reflect objective categories developed earlier among children growing up in Israel than in the United States. These patterns suggest that some form of culturally embedded learning influences whether and when in development children rely on the inherence heuristic to explain the social groupings they encounter (Rhodes et al. 2012).

Thinking of categories as marking people who are inherently different from one another has long been theorized to contribute to prejudice (Allport 1954). Indeed, recent research in my lab has found that as preschool-age children learn about new social groups, those who develop the belief that the categories mark people who are inherently different from one another develop more negative attitudes about the group (Rhodes & Leslie, in preparation). Another possible consequence of taking an inherence-based view of social categories is a tendency to view category-based behaviors as inevitable and consistent with prescriptive norms. By at least age 3, children have robust expectations that social categories shape their members' social interactions; in particular, children predict that harmful interactions will occur more often between members of different categories than among members of the same group (Rhodes 2012). Children's explanations suggest that they take an inherence-based approach to these patterns – preschool-age children explain instances of intergroup harm by referencing stable category memberships but intragroup harm by referencing more transient situational factors (Rhodes 2014). Further, children appear to view patterns of intergroup harm as consistent with prescriptive norms – they

view intragroup harm as necessarily prohibited, but view the permissibility of intergroup harm as dependent on the external context (Rhodes & Chalik 2013).

Thus, inherence-based thinking appears to contribute to two key negative consequences of social categorization – social prejudice and the cognitive processes that underlie intergroup conflict. Understanding how the inherence heuristic develops, therefore, should provide useful guidance as to how these phenomena could be prevented or undone. C&S note that inherence-based views can be undone if people encounter evidence inconsistent with these beliefs. Yet, prompting people to overturn social beliefs in the face of new evidence is notoriously challenging. Given that young children adopt an inherence-based view of only a subset of possible social categories – with the particular categories in this set varying across cultures – a more promising approach might be to identify the process that leads children to apply these beliefs to particular categories in the first place. This proposal highlights the importance of the early childhood years – when inherence-based thinking and social categorization are emerging – as an important potential time for intervention. Understanding more about why certain patterns trigger inherence-based thinking, or the features of children's experiences that lead them to rely on these beliefs more for some categories than others, could provide new insight into how to approach these intractable social problems.

# The causes of characteristic properties: Insides versus categories

doi:10.1017/S0140525X13003889

Michael Strevens
*Department of Philosophy, New York University, New York, NY 10003.*
**strevens@nyu.edu**
**http://www.strevens.org**

**Abstract:** Cimpian & Salomon (C&S) propose that the inherence heuristic, a tendency to explain the behavior and other properties of things in terms of their intrinsic characteristics, precedes and explains "essentialist thinking" about natural kinds. This commentary reviews evidence that it is rather essentialism (or something like it) that precedes the assumption of inherence, and suggests that essentialism can do without the inherence heuristic altogether.

In human reasoning about the characteristic appearances and behaviors of natural kinds such as folk genera and chemical substances, the cognitive timeline, according to Cimpian & Salomon (C&S), is as follows: First comes the inherence heuristic and thus an assumption that the causes of those appearances and behaviors – the tiger's stripes, water's transparency – is something internal. Then comes essentialism and so the further suppositions that (a) this internal property is for each appearance or behavior the same, so that a *single* internal property explains all characteristic appearances and behaviors, and (b) the property in question is an essence, necessary and sufficient for kind membership.

Essentialism in its sparest form does not attribute to the thinker the belief that the essence is something wholly internal; it is allowed, in particular, that the thinker is agnostic about the essence's location. The hypothesis that essentialism emerges from a prior commitment to inherence can be tested, then, by asking whether the essence is, from the first moment that the child begins to think essentially, represented as something definitely inside the animal, plant, or substance.[1]

An original aim of psychological essentialism was to explain the results of Keil's (1989) "discovery" and "transformation" experiments. In the biological version of the discovery experiment, a creature with the external characteristics of one species is discovered to have the insides and lineage of another – an animal that looks and acts like a horse, for example, is found to have the insides and ancestry of a cow. In the biological version of the transformation experiment, a creature is cosmetically transformed to have the appearance and behavior of another species – a raccoon, for example, is made up to look and smell like a skunk. In both experiments, subjects are asked to classify the animal in question: Horse or cow? Raccoon or skunk?

Keil and his collaborators found that younger children tend to classify in accordance with appearances, answering "horse" and "skunk" when presented with the scenarios described in the previous paragraph, whereas older children and adults tend to classify in accordance with deep properties ("cow" and "raccoon"). Crucially, the older children sooner and more confidently make "deep" classifications in the transformation cases than in the discovery cases. This is the datum that tells against the inherence explanation of essentialism.

How so? If essences are thought to be wholly internal properties, then an animal with cow insides must possess the cow essence and so must be a cow. An essentialist thinker, then, provided that they have enough courage in their convictions to classify in contradiction to external properties, should be confident that the animal in the discovery experiment is a cow. Suppose that the advocates of psychological essentialism are correct in thinking that the results of the transformation experiments are to be explained by essentialist thinking. Then, any subject who gives the "deep" answer in the transformation experiment is an essentialist thinker and so ought at the same time to give the deep answer with equal confidence in the discovery experiment. That is not what Keil found: On the whole, deep answers in the discovery experiments come later and with less certainty.

Let me make the same point in a different way. In the transformation experiment, subjects are told that the animal begins as a raccoon. In the discovery experiment, they are not told that the animal begins as a cow, but only that it has cow insides. But if C&S are correct in supposing that essences are from the first believed to be internal, then it should be straightforward for subjects to reason from insides to essence – from having cow innards to being a cow. The two cases ought therefore be on a par. (Indeed, the deep answer ought if anything to come more easily in the discovery cases.) Quite the contrary effect is observed; this shows, I suggest, that many young essentialist thinkers (and quite possibly many adults, too) are agnostic about the location of the essence.

More generally, where we see "essentializing" in human thought we are not always seeing internalizing. New Yorkers are expected to behave the same way wherever they go; does this show that naive reasoners believe that New Yorkers have some internal property that causes their famous characteristics? Not necessarily; it would be enough that the reasoners believe that being a New Yorker causes those characteristics and that being a New Yorker, like being a raccoon, is a persistent state of affairs. Expectations about New Yorkers are explained, on this approach, by attributing to the reasoner a belief in the persistence and causal efficacy of category membership.

Such an attribution is at the core of the psychological essentialist's explanatory strategy.[2] Essentialism does not rely on a commitment to inherence at all: What is persistent need not be internal, and what is internal need not be persistent. Thus, the inherence theorist's explanatory schema is more a rival to than a complement to the essentialist's explanatory schema: Whereas the essentialist emphasizes category membership, the inherence theorist emphasizes the physical constitution of individual category members. The two can live together, as they do in an "internal essentialism," according to which category membership is determined by something physically internal. But the evidence suggests that, psychologically, often enough they do not.

**NOTES**
**1.** What follows are arguments drawn from Strevens (2000), presented there as reasons to favor "pure essentialism" over "insides essentialism," the difference between the two being precisely the question of whether essences are internal properties of the organisms and substances that possess them.
**2.** In Strevens (2000), I argue that the strategy does not require us to attribute to the reasoner a belief in essences, and that it is in fact better not to do so; the persistence and efficacy of category membership are quite enough. What's needed, then, is a "lite" essentialism that dispenses with essences, a view I have called *causal minimalism*. But this is an argument for another time.

# System-justifying motives can lead to both the acceptance and the rejection of innate explanations for group differences

Eric Luis Uhlmann,[a] Luke Zhu,[b] Victoria L. Brescoll,[c] and George E. Newman[c]
[a]*HEC Paris – School of Management, Management and Human Resources Department, 78351 Jouy-en-Josas, France;* [b]*Sauder School of Business, University of British Columbia, Vancouver, BC V6Z 3B7, Canada;* [c]*School of Management, Yale University, New Haven, CT 06511.*
eric.luis.uhlmann@gmail.com luke.zhu@sauder.ubc.ca
victoria.brescoll@yale.edu george.newman@yale.edu
http://www.socialjudgments.com
http://www.sauder.ubc.ca/Programs/PhD_in_Business_Administration/Current_Students/Zhu_Luke
http://som.yale.edu/victoria-l-brescoll
http://som.yale.edu/george-e-newman

**Abstract:** Recent experimental evidence indicates that intuitions about inherence and system justification are distinct psychological processes, and that the inherence heuristic supplies important explanatory frameworks that are accepted or rejected based on their consistency with one's motivation to justify the system.

In the target article, Cimpian & Salomon (C&S) make a compelling and persuasive case that a wide range of psychological phenomena, from essentialism to correspondent inferences to system justification, are not only closely related but reflect a fundamental bias toward explaining observed patterns in terms of inherent features. The present commentary focuses on one aspect of this theory in particular: the relationship between the inherence heuristic and system justification.

System justification theory (Jost & Banaji 1994; Jost et al. 2004) proposes that people have a fundamental motive to view their social system as just, fair, and good, and will thereby engage in a number of strategies to rationalize prevailing social arrangements. C&S suggest two possible relationships between the inherence heuristic and system justification. One is that the tendency to regard existing social arrangements as fair and just may result directly from the more general tendency to view various aspects of the social environment as inherent features of that environment. In other words, the inherence heuristic may largely explain or even subsume system justification. Another possibility is that intuitions about inherence may provide important inputs into biased reasoning aimed at rationalizing and justifying the prevailing social order, but that ultimately, system justification and the inherence heuristic are in fact distinct psychological processes.

This commentary reviews empirical findings favoring this latter model, in which the inherence heuristic and system justification are distinct psychological processes that interact with each other in complex ways. First, past research on system justification suggests that the phenomenon itself is a motivated process that can be activated through various types of experimental manipulations (e.g., Cutright et al. 2011; Jost et al. 2005; Kay et al. 2005; Laurin

et al. 2010). For example, in a frequently used manipulation, participants read either that their country has reached a low point in terms of social, economic, and political conditions (system threat), or that things are relatively fine (system affirmation) (Jost et al. 2005; Kay et al. 2005). Consistent with prior work on motivated reasoning (Dunning et al. 1995; Kunda 1990), people tend to react against threatening information by supporting and bolstering their social system through greater endorsement of prevailing ideologies, preferences, and stereotypes of social groups (Cutright et al. 2011; Jost et al. 2005; Kay et al. 2005; Lau et al. 2008).

Second, further empirical findings suggest that rather than the inherence heuristic leading directly to system justification, it instead appears to provide important inputs to reasoning aimed at satisfying system-justifying motives. In one recent investigation Zhu et al. (2013) manipulated both system-justifying motives (Laurin et al. 2010) and the accessibility of beliefs related to meritocracy and egalitarianism (two important but often opposing American values) through a priming manipulation. Activating motives to justify the system led participants to support whichever value had been made temporarily accessible through priming. In accord with C&S's view, this result suggests that people view whatever ideologies are accessible in a given social environment as inherent properties of that environment, but only seize on and use those ideologies when motivated to support the system. As a result, system-justifying motives can lead individuals to endorse not only hierarchy-enhancing beliefs (e.g., meritocracy) but even hierarchy-attenuating beliefs (e.g., egalitarianism).

Finally, especially strong evidence for the independent nature of system justification and the inherence heuristic comes from recent studies demonstrating that system-justifying motives can lead to not only the acceptance but also the *rejection* of innate explanations for group differences (Brescoll et al. 2013). Specifically, a threat to the system leads people to endorse innate explanations for gender differences when these differences are portrayed as immutable but reject such explanations when they are portrayed as mutable. That is because one way of justifying the system and existing status hierarchies, such as those between men and women, may be to explain group differences as immutable. For example, if one perceives the existing social structure as an inevitable "fact" that is unlikely to change over time, then it limits the criticisms that can be made of the status quo and lends support to the current system. In other words, innate explanations imply that existing social structures (such as differences between social groups) are fundamental, not likely to change, and therefore "right," which can be used as a means of reaffirming the status quo. Indeed, Brescoll et al. (2013) find that motivations to justify the system can lead people to seek out innate explanations for gender differences because such explanations imply the system is stable and incapable of being changed. However, when those same innate explanations are portrayed as *mutable*, individuals experiencing a system threat reject those same innate explanations for gender differences.

Additionally, under system threat, people also *reject* innate explanations for socially stigmatized behavior such as homosexuality and obesity (Brescoll & Uhlmann 2013). Thus, system-justifying motives can lead people to radically flip-flop on whether they endorse the belief that human behavior is driven by inherent/innate factors, depending on the implications of those explanations for system justification.

In sum, it seems clear that the tendency to explain social arrangements in terms of inherent factors is fundamental to human cognition and also plays an important role in many phenomena, including system justification. Recent research indicates that when considering the relationship between the inherence heuristic and system justification in particular, these appear to be distinct psychological processes with the inherence heuristic supplying important explanatory frameworks that are accepted or rejected based on their consistency with one's motive to justify the system.

# Generalizing a model beyond the inherence heuristic and applying it to beliefs about objective value

Graham Wood

*School of Humanities, University of Tasmania, Launceston, 7250 Tasmania, Australia.*
**graham.wood@utas.edu.au**
**http://www.utas.edu.au/humanities/people/philosophy-profiles/Graham-Wood**

**Abstract:** The inherence heuristic is characterized as part of an instantiation of a more general model that describes the interaction between undeveloped intuitions, produced by System 1 heuristics, and developed beliefs, constructed by System 2 reasoning. The general model is described and illustrated by examining another instantiation of the process that constructs belief in objective moral value.

Cimpian & Salomon (C&S) describe a process by which inchoate intuitions outputted by the inherence heuristic are developed by other cognitive processes into beliefs about psychological essentialism. There is much of interest in their proposal, but here I discuss the potential to move beyond this specific instantiation to a general model that may have wider application. To begin, I identify what I take to be the central insight of the authors, combine that with other insights drawn from cognitive science, and describe the features of the general model. To illustrate the general model, I apply it to belief in objective value. And I conclude by highlighting an implication of the general model that is relevant to comments made by the authors about moral reasoning.

The central insight of C&S is that psychological essentialism "emerges as an elaboration of the earlier, and more inchoate, intuitions supplied by the inherence heuristic" (sect. 4, para. 1). In claiming this, I assume the authors accept: (1) the existence of two systems of cognition as described by Stanovich and West (2000), (2) that System 1 (*intuition*) outputs inchoate (I take that to mean "undeveloped") intuitions, and (3) that System 2 (*reasoning*) elaborates upon those to produce developed beliefs associated with psychological essentialism. But I suggest that the process they describe is only one instantiation of a more general process of interaction between undeveloped intuitions generated by System 1 heuristics and developed beliefs endorsed by System 2.

To fill out this general model of interaction between System 1 and System 2 further, let me identify some other insights. Kahneman (2002) describes intuition (and here let us assume the inherence heuristic is one source of such intuition) as "occupy[ing] a position – perhaps corresponding to evolutionary history – between the automatic operations of perception and the deliberate operations of reasoning" (p. 450). This comment prompts a question: Is intuition (e.g., the outputs of the inherence heuristic) closer to perception or reason? From what they say, I take C&S to assume that the functioning of the heuristic is closer to what we call reasoning. But what if its functioning is much closer to what we call perception?

C&S distance themselves from the possibility that the inherence heuristic could be an "innate module," preferring to characterize the associated process as a "highly constrained developmental process" that generates "essentialist beliefs out of simpler cognitive parts" (sect. 4.2, para. 1). But what if these "simpler cognitive parts" are the outputs of a module? There are features of modules that fit well into the general model that I am advancing. For example, Fodor (1983) describes the outputs of peripheral modules as mandatory and contends that the manipulations made inside these modules are relatively inaccessible to introspection (p. 55). The Müller–Lyer illusion is a good example of the operation of such a module (or set of modules). No detailed neurophysiological explanation of the illusion has achieved broad acceptance (Bertulis & Bulatov 2001, p. 5), but presumably the retina receives information that corresponds to two parallel lines of equal length, and somewhere between the retina and the conscious experience of the

illusion the information representing the lines is manipulated to represent lines of unequal length.

I invite C&S to reflect on the insights of Kahneman and Fodor when considering their proposal. Perhaps the functioning of the inherence heuristic is closer to perception than to reason, and its functioning is more inaccessible to introspection than they have assumed. If so, then they may agree that the inherence heuristic is one instantiation of a more general model of interaction between System 1 and System 2.

In this general model, undeveloped intuitions about represented objects, events, or states of affairs are produced by one (or more) System 1 heuristic. These undeveloped intuitions are then available to System 2, and System 2 generates developed beliefs about these objects, events, or states of affairs. Elsewhere, I have described this process as System 1 providing System 2 *"with the concepts with which it thinks"* and illustrated it by claiming that System 1 provides System 2 with pieces of a jigsaw, and System 2 then does what it can to construct a coherent picture with those jigsaw pieces (Wood 2012, p. 77). I suggest the inherence heuristic can be understood as an instantiation of this general model.

To give another instantiation of the general model (I call it the "jigsaw-piece" model) consider moral judgement. Mackie (1977) correctly observes that "ordinary moral judgements include a claim to objectivity, an assumption that there are objective values" (p. 35). Furthermore, Mackie claims that "if there were objective principles of right and wrong, any wrong (possible) course of action would have not-to-be-doneness somehow built into it" (p. 40). Now, the question of the existence of *objective* value is beyond the scope of this discussion, but what is of interest here is Mackie's identification of "not-to-be-doneness" because the nature and origin of this "objective, intrinsic, prescriptivity" (p. 35) is considered mysterious by many philosophers. But the jigsaw-piece model removes the mystery and offers a way to understand the relationship between "not-to-be-doneness" and belief in objective value. An undeveloped "not-to-be-done" intuition is outputted from a System 1 heuristic (like the way the inherence heuristic outputs a inchoate intuition), and System 2 constructs a developed belief in objective value (again, like the way System 2 constructs a developed belief in psychological essentialism). And, the "not-to-be-doneness" intuition could be innate or set by social referencing for particular activities (just like the way the inherence heuristic works).

In conclusion, C&S claim that understanding the inherence heuristic will help avoid "is–ought errors" in moral reasoning (sect. 3.4). But if the general model described here is correct, and if it includes, as an instantiation of it, the process that produces belief in objective value, then it has implications for moral reasoning that challenges C&S's claim. Indeed, the general model challenges moral reasoning itself because the "not-to-be-doneness" intuition – that is the foundation of our belief in objective value and thus is behind the central assumption underlying ordinary moral judgement – may be an illusion, like the apparent objectivity of the different lengths of the lines in the Müller–Lyer illusion.

# Inherence heuristic versus essentialism: Issues of antecedence and cognitive mechanism

Vincent Y. Yzerbyt and Stéphanie Demoulin

*Institute for Research in Psychological Sciences, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium.*

**vincent.yzerbyt@uclouvain.be**          **stephanie.demoulin@uclouvain.be**

**Abstract:** We argue (1) that the empirical evidence offered by the authors is insufficient to sustain their claim; (2) that, beyond methodological problems, the proposed underlying cognitive mechanism is largely speculative and that a reverse, more motivational, path is equally plausible; and (3) that the distinction and antecedence of inherence intuitions with respect to essentialist beliefs remain to be demonstrated.

Cimpian & Salomon (C&S) make two innovative moves: First, they highlight the commonalities among seemingly distinct psychological phenomena. Although not necessarily new (e.g., see Yzerbyt et al. 1997), such an endeavor is valuable because it enables connecting different domains and possibly developing a more parsimonious understanding of human functioning. Second, they argue that these commonalities all trace back to a single underlying cognitive mechanism – that is, the inherence heuristic. Doing so, they fall into the trap of the exact same heuristic that they describe: They assume that the similarities between different psychological phenomena ought to be the reflection of some inherent human cognitive functioning. Sensible as this idea may be, we argue (1) that the empirical evidence offered by the authors is insufficient to sustain their claim; (2) that, beyond identifiable methodological problems, the direction of their proposed underlying cognitive mechanism is largely speculative and that a reverse, more motivational, path is equally plausible; and (3) that the distinction and antecedence of inherence intuitions with respect to essentialist beliefs remain to be demonstrated.

Throughout their article, C&S present a series of arguments along with allegedly supportive evidence in favor of the idea that human beings show a deep and irrepressible inclination to "make sense of observed patterns in terms of the inherent features of their constituents" (sect. 3, para. 1). Interestingly, they hasten to soothe their message and acknowledge that products of the explanatory process are not always consistent with the inherence heuristic: Next to individual and developmental differences, people are not only more prone to make dispositional attributions when observing others (or members of other groups) than when explaining their own behaviors (or the behaviors of close others), but the valence of the observed behaviors would also seem to be an important moderator of this effect (e.g., Hewstone 1990). According to C&S, inconsistencies emerge when individuals *block* or *revise* the inherence heuristic process, enabling more extrinsic explanations to emerge. If this were the case, one should expect that when *blocking* or *revising* is hindered, say, because participants are under cognitive load, intuitions based on the inherence heuristic should pop up again. Such empirical evidence is not offered by C&S. Also, their strong view has some trouble with findings showing that cultural and pragmatic factors moderate the correspondence bias (e.g., Leyens et al. 1996).

In addition, those studies that most directly test the authors' hypotheses are not without limitations. To mention only one, the inherence heuristic scale should be thoroughly tested for its discriminant validity with respect to closely related constructs – for example, adherence to cultural norms or conservative tendencies. If, as we suspect, the scale measures norm adherence, it is likely that children would report lower absolute levels on this scale than would adults, an outcome that would be incompatible with C&S's argument that the inherence heuristic is more prevalent during childhood.

More critically, C&S provide no clear evidence regarding the causal direction of the hypothesized inherence heuristic. They propose that information about the targets is activated first, that this information is biased toward inherence, and that the activated inherent features then give rise to inherent-type explanations for the (observed) pattern. Because of the correlational nature of the data, a reverse causal path is equally plausible, one that would start with the activation of an inherence explanation, which would lead to the subsequent preferential reliance on the inherent characteristics of the pattern's constituents (see Yzerbyt et al. 2001). It is well known that human beings have a strong motivation to develop a sense that the world is coherent and predictable (Fiske & Taylor 2008). Because inherent accounts are especially well suited to answer humans' need for coherence and predictability, these accounts should be preferentially selected as working

hypotheses about events' co-occurrence. In this model also, inherent features of the constituents are activated, but, rather than being precursors of the inherence intuition, they are conceived of as its consequences, via a hypothesis confirmation bias (Snyder 1984). Concretely, when observing girls' affinity with pink, people initially develop the hypothesis that there must be "something" underneath that explains the color preference (i.e., the inherence intuition) before elaborating on what this "something" could be (i.e., the constituents' inherent features – girls' feminine nature).

This reverse causal path is entirely consistent with the observation that people often develop inherence intuitions despite their inability to pinpoint directly what the inherent-type explanation may be. That is, intuitions develop before people gain access to the exact features on which they later develop their account. The reverse path also explains why inherence intuitions tend to persist even if external (e.g., historical, socioconventional) explanations are otherwise accessible. As people test the inherence hypothesis, they search for information that confirms (inherent features) rather than questions (external constraints) their a priori beliefs. Moreover, although C&S draw a clear-cut line between inherence explanations and historical/socioconventional ones, we suggest that the two types of accounts are not necessarily mutually exclusive or incompatible in people's minds. People do not always choose for one explanation by dismissing the other. Instead, when assessing the credibility of their inherence intuition, perceivers are able to interpret historical constraints in hypothesis-serving ways. (For example, the girl–pink association emerged in marketing campaigns in the nineteenth century because pink is an inherently feminine color.)

Finally, the differences and similarities between an essentialist stance and the inherence heuristic remain weakly documented at a theoretical level and, indeed, hardly supported at the empirical one. C&S report evidence that their inherence scale is strongly correlated with a standard essentialism scale (Haslam et al. 2000), even after controlling for a host of cognitive, personality, and ideological dimensions. Notwithstanding the difficulties of appraising the exact nature of the inherence scale, such findings tell us nothing about, and even tend to undermine, the general argument that inherence intuitions precede essentialist beliefs.

# Author's Response

## Refining and expanding the proposal of an inherence heuristic in human understanding

Andrei Cimpian and Erika Salomon

*Department of Psychology, University of Illinois, Champaign, IL 61820.*
acimpian@psych.illinois.edu       salomon3@illinois.edu
http://psychology.illinois.edu/people/acimpian
http://www.erikasalomon.com

**Abstract:** The inherence heuristic is a cognitive process that supplies quick and effortless explanations for a wide variety of observations. Due in part to biases in memory retrieval, this heuristic tends to overproduce explanations that appeal to the inherent features of the entities in the observations being explained (hence the heuristic's name). In this response, we use the commentators' input to clarify, refine, and expand the inherence heuristic model. The end result is a piece that complements the target article, amplifying its theoretical contribution.

We have used the commentators' questions to expand the original statement of our inherence heuristic (IH) model.

Thus, we intend this response to serve as a substantive companion piece to our target article, with signposts along the way to indicate which commentaries prompted which clarification, refinement, or expansion. We are truly grateful for everyone's input – the diversity of perspectives reflected in the commentaries has given our theory a richness that would have been unattainable otherwise.

The organization of this response loosely mirrors that of the target article, in that we start with a discussion of the IH process and then move on to the phenomena (including essentialism) that we have claimed to stem from this process.

## R1. The inherence heuristic

This section capitalizes on the commentators' questions to elaborate our account of the process that underlies the IH.

### R1.1. The inherence heuristic as a general explanatory heuristic process

Most broadly (and boldly) stated, the goal of our account is to outline *the general process by which many everyday explanatory intuitions are generated*. A key claim of the account is that the process of generating such intuitions has much in common with the heuristic judgments people make in response to other difficult questions (e.g., Gilovich et al. 2002; Kahneman 2011). Specifically, this process overuses information that is easily accessible, ignoring other relevant considerations. Although questions about why the world is a certain way (e.g., why do girls like pink?) are incredibly complex (the sorts of questions that scientists spend entire careers researching), plausible-seeming answers come to mind with surprising ease. How is it that people are not stumped? How do they come up with any sort of answer (let alone an almost instantaneous one), given the daunting complexity of the explanatory task? Our theory offers a solution to this puzzle: We propose that people rely on a heuristic shortcut that leads them to explain by using only the information that's most accessible to them. Because this information often consists of the inherent features of the entities under consideration, the explanations generated will be correspondingly skewed toward inherence – hence the term *inherence heuristic*. Far from being just another quirk of human cognition, this heuristic is likely to have a powerful influence on people's understanding of the world. People seem motivated to wonder *why* from the youngest ages (e.g., Anderson et al. 1996; Callanan & Oakes 1992; Gopnik et al. 2004; Murphy & Medin 1985; Schulz 2012); thus, if our intuitive answers are shaped from the youngest ages by the proposed inherence heuristic, then the possibility of pervasive bias is high.

To reiterate, the sort of heuristic we are proposing is a general explanatory process that avails itself of the most readily accessible knowledge to generate its output and that, as a result, ends up overusing inherent features. Thus, the IH is not "devoted to" postulating inherent explanations (**Dunham**, para. 4); it inadvertently (over)postulates these explanations because inherent features are so accessible in so many circumstances. But there is nothing in the structure of the IH that prevents it from using salient, easily accessible extrinsic information (e.g., about ownership; see **Noles & Danovitch**) to generate an

explanation. The output may not be an inherence-based explanation in this case, but that is not inconsistent with our account: The proposed heuristic is an inherence heuristic not because it is somehow structurally incapable of handling anything other than inherence, but simply because it *tends* not to, given that other sorts of information aren't typically at its "fingertips" (to use Dunham's evocative metaphor [para. 2]). In effect, the IH and the explanatory heuristic described by Dunham in his thoughtful commentary are one and the same. We thank him for motivating this clearer, and more forceful, statement regarding the nature of the IH process.

This discussion has implications for how the label *inherence heuristic* should be understood. The inclusion of *inherence* in this label is intended simply as a reminder of the typical output of the heuristic process we are proposing; inherence is not the goal of this process, nor is it a criterion that's explicitly built into its structure. In other words, the term *inherence heuristic* describes the entire process illustrated in Figure 2 of the target article, whatever the output of the process may be. This is not just splitting hairs: Misunderstanding the name of the heuristic can lead to misconceptions about what we are proposing. For example, **Khalidi & Mugg** seem to have interpreted the name of the heuristic as signaling that the IH somehow selects inherent features for inclusion in explanations because of their inherence. This interpretation led Khalidi & Mugg to disagree with our claims, even though we suspect there is no real disagreement between us. Specifically, they questioned the existence of an inherence heuristic while simultaneously endorsing our main claim: that inherent features are used more often than extrinsic ones in explanations because the former are more accessible. Thus, their skepticism is directed toward a heuristic we didn't propose – one that selects explanatory features for their inherence per se.

As an aside, all of this reminds us of Kahneman and Frederick's (2002) remark that "the label that is chosen for a heuristic may take on a life of its own in subsequent theorizing" (p. 65). The process we are proposing is most transparently and accurately labeled with something along the lines of *the explanatory heuristic process that tends to output explanations that appeal to inherent facts because these facts are often the most accessible*. For obvious practical reasons, we need something a little more succinct, and the *inherence heuristic* seems to fit the bill. Although we did not provide sufficient justification for this term the first time around, we hope to have gotten ahead of the curve now and prevented further confusion about its intended meaning.

In response to points raised by the commentators, we go on to highlight several key aspects of the IH process. As a general explanatory heuristic, the IH is triggered by a wide range of explananda. Thus, the answer to **Kinzler & Sullivan**'s question about whether the IH could generate explanations both for specific instances/events and for broader uniformities is affirmative. The IH is triggered by both and likely leads to intuitions biased toward inherence in both cases. The extent of the bias, however, may be greater for broad patterns than for specific instances because of differences in what information is at its fingertips in these two cases. Extrinsic information is more readily available when the retrieval cue is a specific object or individual: Think of ownership, for example, or of causal

history. Such information is more naturally represented and stored at the level of instances, not broad sets or categories. It's Fido that belongs to our family, that caught a squirrel yesterday, or that was adopted from the shelter – not dogs as a category. The more extrinsic information accessible to the shotgun, the more frequently this information will end up in the explanatory intuitions generated, leading in the case of particular instances to an attenuation (but most likely not elimination) of the usual skew toward inherence relative to what is normatively warranted. This attenuation should not be taken to signal that the IH is solely, or even preferentially, invoked for patterns (as **Gelman & Meyer** and **Prasada** state). On our account, the IH process is invoked, and produces explanations, for all sorts of explananda, both general and specific; what differs between these cases is simply the relative extent to which inherent features dominate its output.

While on the topic of the IH as a heuristic process, we agree with **Wood** that much of the cognitive underbelly of the IH is inaccessible to conscious introspection. The IH is an intuitive (or System 1) heuristic, and lack of conscious access to its workings is a typical feature of such a process (for a nuanced discussion, see Evans & Stanovich 2013). However, we don't agree with Wood's stronger claim that the IH operates as a Fodorian peripheral module (Fodor 1983). Although it may be similar to a module in its opaqueness to introspection, the IH does not display many of the other key characteristics of such a module. For instance, there is no reasonable sense in which the IH is domain specific; rather, it operates on an incredibly broad range of inputs. The IH is not informationally encapsulated either; in fact, because it has free access to the information in semantic and episodic memory, the IH cannot be farther from encapsulation.

Although we are denying that the IH itself is a domain-specific explanatory module, we acknowledge that such modules may exist and may provide ready-made explanations for select chunks of our experiences with the world. For example, humans' ability to make sense of the interactions between physical objects seems to be facilitated by a trove of early-developing, perhaps even innate, causal-explanatory principles (e.g., Baillargeon 2004; Spelke & Kinzler 2007). More recent evidence suggests that a similar explanatory module may be available for making sense of the interactions between psychological agents (e.g., Kovács et al. 2010; Onishi & Baillargeon 2005). Thus, our claim that the IH can account for *many* everyday explanatory intuitions should not be misunderstood as stronger than it is. The IH is clearly not the only explanatory game in town, and we suspect that its influence will be felt most strongly in the (many) circumstances where these domain-specific explanatory principles are silent.

### R1.2. What's a heuristic, anyway?

Because there was some disagreement in commentators' assumptions about what a heuristic is (**Bookstein**; **Braisby**; **Dubljević & Racine**; **Gaucher & Jost**), we should clarify where we stand on this issue. Following a widely accepted view in the recent literature on heuristics and biases, we understand a heuristic to have at its core a form of *question substitution* (e.g., Kahneman 2011; Kahneman & Frederick 2002). When trying to come up with solutions to complex problems or questions, people

inadvertently rely on easily accessible information and/or simple computations more often than would be normatively warranted. This liberal reliance on cognitive "low-hanging fruit" enables people to find quick answers to questions that are enormously complicated, which is arguably a better outcome than just being perplexed. But there's a catch: The answers generated via this process no longer address the original questions but related, and cognitively simpler, ones. Thus, from an observer's perspective, it seems as though people often substitute easy questions for harder ones. For instance, when asked a question about the probability that Tom W. is a computer scientist, people often answer as though they had substituted a related, easier question about Tom W.'s similarity to the stereotypical computer scientist (Kahneman & Tversky 1973). To clarify, though, *question substitution* is just a descriptive tool – another metaphor. No question-swapping is actually (i.e., causally) involved in the process by which people generate their heuristic answers; the key feature of this process is its reliance on easily accessible information. However, this reliance leads to responses that make it seem as though people are answering a different question than the one they set out to answer, which justifies the metaphor of question substitution.

Responses that can be characterized as question substitutions have been identified in a wide array of complex judgments: People often judge stimulus familiarity as though they were rating stimulus positivity/attractiveness instead (e.g., Monin 2003); they seem to decide on appropriate punishments as though they were simply reporting how much outrage they felt at the crimes in question (e.g., Sunstein 2005); and so on. The IH model proposes that our everyday explanatory judgments often involve a similar question substitution: The complex question "What explains this observation?" is often answered (due to the high accessibility of inherent facts) as though it were the easier question "What inherent facts explain this observation?"

A corollary of this question-substitution view of the IH is that its output is – at best – an approximation of the truth. Thus, we are unmoved by examples where the IH seems to lead to mistaken judgments. (Both **Braisby** and **Dubljević & Racine** bring up such examples.) These would count as counterevidence to our claims only if we assumed that the IH is perfectly accurate. Since many of our own examples (in the target article) point to the suspect intuitions generated by the heuristic, we hope it is clear we do not endorse such an assumption. However, we are agnostic about *exactly* how often the heuristic will output the correct answer (that is, beyond the weak claim that it isn't always right). Note that we can afford to remain agnostic on this point because the truth of our hypothesis is independent of the frequency with which the IH is right: Our proposal is of a specific cognitive process by which everyday explanations are generated – a process that can be described as a sort of question substitution. This mechanistic proposal should be evaluated on its own merits, without getting mired in ideological debates about heuristic accuracy (e.g., Gigerenzer 1996; Kahneman & Tversky 1996).

### R1.3. What exactly are inherent features?

Our account makes a basic distinction between inherent and extrinsic facts. To elaborate on the definition provided

in our target article, inherent facts are those that can be said to characterize a thing "in virtue of the way that thing itself, and nothing else, is" (Lewis 1983, p. 197). The shape of an orange is an inherent fact about it; the fact that it is found next to some apples is not, and neither is the fact that it was washed in a sink – these are extrinsic facts about it. Another intuitive way to understand this distinction is that inherent facts about an object are those facts that must be true about a perfect duplicate of that object. A perfect duplicate of an orange would necessarily have the same shape as the original (as well as the same color, DNA, chemical composition, etc.); it would not, however, have the same location or history.

In part because of our cursory description of this distinction in the target article, several of our commentators understood us to mean something else by *inherent*; almost all of these alternative meanings for the term were more restrictive than our intended meaning. **Strevens** thought we meant *inherent* to be synonymous with *internal*. (Note, however, that a feature such as the shape of an orange is inherent without being internal.) Similarly, **Noles & Danovitch** thought that inherent properties are internal and nonobvious, whereas **Hampton** thought they are necessarily "deep." **Prasada** defined as inherent only those features that have a principled or causal connection with their kinds (e.g., Prasada et al. 2013). At the other end of the spectrum, **Khalidi & Mugg** understood precisely the distinction we were trying to make but then went on to use *inherent* as if it were tantamount to *perceptually salient*. (However, a feature such as the chemical composition of an orange is inherent without being perceptually salient.) In light of the foregoing expanded definition, it should be clear that none of these alternative definitions adequately capture the distinction we are making.

We should also clarify that, as **Khalidi & Mugg** helpfully pointed out, our account operates with a psychological notion of inherence rather than a metaphysical one. That is, our hypothesis does not assume that inherent properties actually exist – that they are real aspects of the world (a metaphysical claim). This is just as well, since the metaphysical notion of inherence has proven difficult to pin down (e.g., Weatherson & Marshall 2013): Many properties might at first appear inherent, but it is in fact unclear whether they are really (that is, metaphysically) so. For example, is an object's shape an inherent feature of that object, or is it an extrinsic feature because it depends on the curvature of the space around the object? However, our account can sidestep these metaphysical difficulties. *Intuitively*, the shape of an object appears to most people to be an inherent feature of that object, and that is all we need here.

Thus, our account assumes only that people intuitively conceive of some features as inherent and of others as extrinsic (a psychological claim). Note, however, that even this psychological assumption is needed in only a limited sense: The IH relies on inherent features not because of their inherence per se but rather because they tend to be low-hanging fruit for retrieval processes. Therefore, if inherent features are used just because they come to mind more readily than other facts, then the heuristic process doesn't actually need to be able to pick out inherent features. In other words, people's intuitive ability to distinguish between inherent and extrinsic facts is not directly involved in the process by which the heuristic generates its output.

So, then, what work does the inherent/extrinsic distinction do in our account? Most obviously, it is a descriptive tool that enables us (as researchers) to characterize the output of the hypothesized explanatory heuristic and to determine whether, as claimed, it typically relies on inherent facts. This distinction may also play a mechanistic role, but only *after* the IH has generated its output. Specifically, the ability to identify inherent features as such may enable reasoners to take the output of the IH and draw additional inferences from it, depending on whether it appeals to inherent or extrinsic facts. For instance, crucial inferences about the stability, inalterability, and naturalness of the phenomena explained are considerably more likely to follow from explanations that (are understood to) appeal to inherent facts.

But can people in fact draw the inherent/extrinsic distinction consistently? Is there sufficient agreement as to what features are inherent and which are extrinsic? Although this question awaits further investigation, the evidence so far leads us to believe that the variability in people's intuitions about inherence is relatively minimal, contrary to **Khalidi & Mugg**'s claim. Khalidi & Mugg raise the possibility of widespread disagreement about these matters via several examples. For instance, do people view features such as orange juice's being tangy or healthy as inherent (as we claimed), or do they view them as extrinsic because they depend on a relation between OJ and humans? The empirical data we know of point to the conclusion that these sorts of features are in fact consistently judged to be inherent: People think that lemons are sour, that carrots are crunchy, that snow is white, and so on, simply by virtue of being the kinds of things they are (Prasada & Dillingham 2006) and not because of the relations they bear to humans (specifically, to our perceptual apparatus). We don't doubt that ambiguous cases exist, but they should not detract from the conclusion that making intuitive inherent/extrinsic judgments is typically straightforward.

### R1.4. Is the inherence heuristic an inherent feature of the human mind?

Some of our commentators raised the question of whether the IH is itself an inherent feature of cognition (**Baron**; **Bookstein**; **Hampton**; **Khalidi & Mugg**; **O'Connor & Joffe**). As far as we are concerned, the answer is "yes and no." Some elements of the IH process are supplied by the endogenous structure of the human mind, whereas others are filled in by the external world. Among the likely candidates for the inherent components of the IH are the human propensity to ask *why*, the fallibility of our memory (which makes it so that stored information is not uniformly accessible), and our bounded rationality (which, among other things, leads to a satisficing reliance on easily accessible information) (e.g., Simon 1982). Nevertheless, the operation of the IH also depends in crucial ways on factors extrinsic to the mind. For instance, what information is most accessible to the IH process is in part a function of the multilayered context in which people are embedded (e.g., broad cultural beliefs, specific prior experiences; see sects. R1.5 and R1.6). How people handle the output of the IH (adopt vs. question/revise), and thus the heuristic's ultimate impact on their belief systems, may likewise be influenced by the sociocultural context (see sect. R1.7).

In sum, as with most other psychological processes, the functioning the IH depends both on features of the mind and on features of the world.

### R1.5. Which aspects of the world are noticed and encoded? Which are explained?

A number of commentators (**Braisby**; **Kinzler & Sullivan**; **Rhodes**) raised questions about the steps that precede the IH process per se: What aspects of the world do people notice and encode? And, of these, which do they seek to explain? These questions are clearly important for a comprehensive account of how the IH shapes human understanding across different domains of experience. Nevertheless, full discussion of these matters goes far beyond the scope of this response and, indeed, beyond the limits of current scientific knowledge. Here, we provide only a few speculative thoughts.

Regarding what people *attend to*, we suspect that the final answer will point to at least two sources of influence. First, human attention is guided by skeletal, domain-specific biases, arguably the products of natural selection, that highlight certain aspects of experience over others (e.g., Baillargeon et al. 2010; Chomsky 1959; Gelman 1990; Izard et al. 2009; Morton & Johnson 1991). Second, the social environment provides another powerful guide for attention, leading to preferential encoding of some observations over other, equally available, ones. In fact, several of our commentators have done groundbreaking work on exactly this topic (e.g., Bigler & Liben 2006; 2007; Gelman et al. 2010; Rhodes et al. 2012). Such sociocultural cues combine with the skeletal biases just mentioned to pare down the vast amount of information that is in principle available to humans. The end product of this filtering process is the information actually *encoded*, which is the raw material for the IH – what the heuristic is invoked to explain and what it explains with.

Importantly, however, not every bit of information that is encoded triggers the search for an explanation; the fact that humans are motivated to explain and understand does not entail that they will try to explain everything. Undoubtedly, some information is stored in semantic memory without an explanation. (Unlike **Braisby**, we don't see this claim as particularly controversial or problematic for our broader account.) The question then becomes, what *does* trigger the search for an explanation? Which facts do people seek to explain? Unfortunately, this issue has received less empirical attention than the one regarding what information people tend to encode. Some evidence suggests that children are particularly motivated to explain broad facts about the world (Cimpian & Petro 2014), whereas the findings in other studies have suggested that unexpected observations also prompt more explanation seeking (e.g., Legare et al. 2010; Weiner 1985). This cannot be the whole story, though, so much remains to be investigated here.

To conclude, the explanatory heuristic process that we are proposing is triggered by *why* questions about aspects of the world that people have, at some point, noticed. Thus, our proposal will ultimately need to dovetail with research on (a) what prompts people to ask *why*, and (b) what things people tend to notice, in order to comprehensively articulate how the IH affects human understanding across a wide variety of contexts and domains.

### R1.6. What's in the shotgun?

The first stage in our model is the *mental shotgun*: the fast, shallow memory search for relevant facts performed as soon as the IH process is set into motion. As with intuitive heuristics more generally, the operating principle of this stage is to take the path of least cognitive resistance. Here, this operating principle leads the IH to retrieve only the subset of relevant information that is most easily activated from memory. What sort of information might this be? Are there any commonalities in the content of this most-accessible information that hold across the observations being explained, the situations in which the explanation is being generated, and so on? A key insight of the IH model is that there may indeed be some systematicity in the content retrieved by the shotgun. Specifically, it is possible that – on average – the most easily activated information will consist of inherent facts about the entities in the observation being explained (see sect. 2.2 of the target article for the full argument). This imbalance in retrieval then propagates through the other stages of heuristic processing, ultimately causing an inherence skew in the intuitive explanations that the IH outputs. Notice the qualifier, though: on average. The shotgun cares only about accessibility, so if extrinsic facts happen to be most accessible for a particular observation or in a particular context, then they will be retrieved just as readily as inherent facts.

Armed with this description of the shotgun, we go on to address the questions that were raised about this stage of the IH. First, in response to comments by **Dunham** and by **Gelman & Meyer**, we reiterate that the shotgun does not select facts for their inherence. Rather, it selects facts for their accessibility, which leads only indirectly to the retrieval of a preponderance of inherent facts. Thus, the shotgun is perfectly capable of retrieving (and often does retrieve) extrinsic information; it does not need to be "limited to only inherent features" (**Noles & Danovitch**, para. 5). Second, the content retrieved by the shotgun will undoubtedly be influenced by the local and broader (e.g., cultural) context in which the IH is deployed. If extrinsic forces (e.g., situations, upbringing) figure prominently in the discourse of one's cultural community, for example, then they will be more accessible to the shotgun. The same goes for many inherent factors, too, especially less concrete ones. Self-esteem, giftedness, bodily humors, etc., are (or were) all culturally prominent inherent features that are easily retrievable by people belonging to the relevant communities in their search for explanations. In effect, the content retrieved by the shotgun will be influenced by anything that influences the accessibility of explanation-relevant information in memory, either temporarily or chronically. As such, cultural input is certain to influence the IH process. We thank several of our commentators (**Baron**; **Kinzler & Sullivan**; **O'Connor & Joffe**; **Olivola & Machery**; **Yzerbyt & Demoulin**) for prompting us to elaborate on this important point.

### R1.7. How does the storyteller work?

The next stage in the heuristic process is the *storyteller*: the search for a quick way to assemble (at least some of) the facts pulled up by the shotgun into a plausible explanatory story. The operating principle of the storyteller is the same as that of the shotgun – namely, take the path of least

resistance. Thus, the storyteller is likely to capitalize on whatever explanatory framework comes to mind most readily that can organize the information pulled up in the preceding shotgun search. This stage is also likely to terminate its search for an explanation as soon as a first plausible explanation is assembled.

A simple way to think about the relation between the shotgun and the storyteller is that the shotgun supplies the *content* of the explanations generated by the IH, whereas the storyteller supplies the *structure* of these explanations. This point brings to light the relative centrality of these two stages to the IH proposal. Because we are fundamentally making an argument about the typical *content* of everyday explanations, it is the shotgun stage that's the linchpin of this proposal. The shotgun is where the bias toward inherence originates before propagating through the rest of the IH process: If the content retrieved by the shotgun is inherence-skewed, the storyteller's options will be constrained accordingly. That is, the content retrieved will be more compatible with, and will facilitate activation of, certain explanatory structures over others. For instance, if the shotgun retrieves a number of inherent features (e.g., the refreshing smell of OJ or its tartness), the storyteller will be hard-pressed to formulate an extrinsic–historical explanation out of such features and may instead appeal to, say, their causal powers. The constraining influence of content on structure is not deterministic, but it may nevertheless be strong enough for inherence in the shotgun to translate, more often than not, into inherence at the level of the final explanations generated by the IH. (An interesting possibility, which we will mention only briefly here, is that the storyteller might be more than a neutral participant in the IH process – that this stage might actually *add* to the skew toward inherence rather than simply propagating it. For example, it is possible that inherence-based explanations are cognitively simpler to build than extrinsic ones, which might often require putting together more elements [e.g., the multiple links in the chain that connects a historical event with a current societal pattern]. If so, then the storyteller's emphasis on finding a quick solution might lead it to favor the easy-to-assemble inherent explanations, which might in turn amplify the heuristic's tendency to output such explanations.)

Because the storyteller is somewhat less central to the thrust of our proposal, we were less specific in the target article about its functioning, except to emphasize that the value it places on speed makes it relatively promiscuous in the explanations it generates and deems plausible. Importantly, however, additional details about the operation of the storyteller can be borrowed from the research exploring the structural aspects of explanations (for reviews, see Anderson et al. 1996; Keil 2006; Lombrozo 2006; 2012; Ross 1977). Whatever this research will ultimately conclude about how explanations are typically structured, about the factors that affect their plausibility, and so on, will probably apply to the functioning of the storyteller as well.

Here, we augment our description of the storyteller with a few results from this literature in order to address specific questions brought up by our commentators. First, **Marmodoro, Murphy, & Baker** (**Marmodoro et al.**) wondered where the storyteller's stories come from; seemingly, our account requires that "the causal story preexist in the mind of the reasoner" (para. 2). For example, why

would the storyteller focus on the tanginess of OJ as an explanation for morning consumption if it didn't already know at some level that tanginess might cause waking? Marmodoro et al. are right in assuming that the storyteller uses previous beliefs in assembling its explanations. However, these beliefs are often quite abstract – for example, sometimes they take the form of general mappings between certain types of causes and certain types of effects. Most relevant to the OJ case, one such abstract mapping may be that sharp perceptual stimuli (e.g., loud noises, bright lights) cause physiological arousal (e.g., being startled, being awakened). By subsuming a particular explanandum (e.g., OJ for breakfast) under a more general pattern such as this one, the storyteller can generate an explanation without having a prestored answer. People view the act of subsuming an observation under a general pattern to be explanatory (e.g., Lombrozo 2006; 2012), so there is no reason the storyteller should not be able to avail itself of this strategy in formulating its output.

Second, **Prasada** provided a plausible example of how the storyteller's operation might be influenced by the content retrieved by the shotgun. Specifically, Prasada suggested that the storyteller's output might vary depending on the representation of the inherent features at its disposal. For instance, inherent features that are represented as having a *principled* connection with their kinds (e.g., tanginess is an aspect of what it means for something to be OJ) might be more likely to license explanations with a more normative flavor (e.g., it's ideal/right that OJ is for breakfast), as well as downstream intuitions about inevitability. We endorse Prasada's prediction, and more generally we see great value in research that would further specify how the content pulled up by the shotgun shapes the storyteller's output.

Third, **Rakoczy & Cacchione** state that the storyteller's "complex inferential machinery" is unlikely to be available early in development, in part because "it appears to rely heavily on linguistic capacities" (para. 1). Although we agree that the storyteller is indeed a complex piece of inferential equipment, the jury is still out on whether it (or something like it) is available early in development. Much of the developmental evidence in our target article (e.g., Cimpian & Markman 2009; 2011) relied on language only because it is relatively easy to assess children's explanations verbally. The obvious drawback of this methodology, though, is that it prevents investigation of these issues in children who haven't yet acquired language. Nevertheless, the absence of evidence in younger children should not be taken as evidence that the relevant abilities are absent. There is nothing in the structure of storyteller, or the IH process more generally, that depends on linguistic abilities, so we predict that this heuristic process might influence reasoning even before these abilities are in place. In making this prediction, we are also encouraged by the evidence that infants' general-purpose explanatory abilities are quite sophisticated (e.g., Baillargeon 1994). Although this evidence does not speak directly to our claims, it is at least consistent with the early presence of the sort of inferential machinery that the IH relies on.

Finally, we should reiterate that, just as the content retrieved by the shotgun is influenced by context, so is the operation of the storyteller. For instance, to the extent that different cultures privilege different explanatory frameworks (e.g., Keil 2006), the differential accessibility of these frameworks would likely be reflected in the output of the storyteller. Whether this context-driven variability in the operation of the IH process is sufficient to account for the full range of cross-cultural variability in explanatory practices remains to be established. Unlike **Baron**, however, we see no need to appeal to multiple separate explanatory mechanisms to account for such variability if a single, context-sensitive mechanism can do the job just as well. (This issue is separate from that concerning whether there exist separate innate, domain-specific explanatory mechanisms, whose important role we have already acknowledged. These mechanisms are too limited in application to account for the range of observations that fall under the scope of the IH, and thus we assume that they are not the explanatory mechanisms Baron proposes as alternatives to the IH.)

### R1.8. How can the typical output of the inherence heuristic be blocked or revised?

Given the prejudicial consequences of the typical (inherence-based) output of the IH, especially when applied to social groups, a number of commentators have emphasized the importance of understanding how to block or revise this output (**Bigler & Clark**; **Kinzler & Sullivan**; **Meleady & Crisp**; **Rhodes**). Before discussing their specific suggestions, we provide some preliminary remarks. The concern here is with overriding what we have claimed is the *typical* output of the IH – namely, its inherence-based output. Although the IH (as an explanatory heuristic process) can output extrinsic explanations as well, these are not the focus of this section. A reasoner can override the inherence-based output of the IH by *blocking* it or by *revising* it. The difference between these two terms maps onto a difference in the stage at which inherence is defeated: *Blocking* occurs when inherence is avoided before the IH has run its course. Blocking might occur, for example, in circumstances where extrinsic facts happen to be easily accessible and are thus retrieved by the shotgun. Blocking is not effortful (contra **Dunham**) but rather as effortless-seeming as the rest of the IH process – in fact, blocking occurs *during* the IH process. In contrast, *revisions* occur when inherence-based explanations are overturned after they are generated, and thus *after* the IH process has run its course. As Rhodes points out, revising already-generated explanations can indeed be effortful, as well as difficult to achieve successfully. We therefore second her call to focus empirical efforts on means of blocking the inherence-based explanatory "stories" (particularly with respect to social groups, where they are wrong quite often) before they are even told.

In what follows, we briefly highlight three promising suggestions for how to block or revise the inherence-based output of the IH that are provided by three different sets of commentators. First, **Bigler & Clark** suggest that perhaps the most effective means of overriding inherence is also the most direct: providing children with the relevant extrinsic facts, especially in ways that make these facts salient and accessible (e.g., via interactive learning). Consistent with the effectiveness of this inherence-reduction strategy, Bigler and Wright (2014) review a number of studies in which teaching children about discrimination and prejudice (as extrinsic societal forces) led to corresponding reductions in children's tendency to attribute

societal phenomena to inherent causes. Interventions such as these may help both by leading children to revise existing inherence-based beliefs and by blocking further such beliefs from being generated.

Second, **Kinzler & Sullivan** raise the possibility that diverse cultural and linguistic experiences may likewise reduce reliance on inherence-based intuitions. Kinzler & Sullivan's proposal seems to work on two levels: Diverse experiences may increase the diversity of the information accessible to the shotgun, increasing the probability that inherence-based explanations are blocked before they are generated. In addition, having to process varied experiences may, over time, also enhance processing resources such as cognitive control (e.g., Bialystok et al. 2004; Kovács & Mehler 2009), which might increase children's ability to scrutinize the output of the IH rather than simply accepting it as is. In turn, such scrutiny may increase the likelihood of revisions to inherence-based intuitions.

Third, similarly to **Kinzler & Sullivan**, **Meleady & Crisp** suggest that chronic exposure to violations of one's inherence-based notions (e.g., in diverse environments where reasoners can interact positively with members of outgroups) may predispose one to question the value of these heuristic intuitions and, as a consequence, adopt a more analytic cognitive style. This style may also be facilitated by improvements in cognitive resources brought about by having to repeatedly resolve the conflict between one's prior beliefs and the evidence contradicting them.

As **Meleady & Crisp** rightly pointed out, our original statement of the IH model left unspecified the mechanisms by which the typical inherence-based output of this heuristic could be overcome. The foregoing suggestions expand this dimension of the account, laying the groundwork for new empirical research and for interventions to curb the unwanted side effects of inherence-based thinking.

### R1.9. Where do inherence-based explanations come from?

Two commentaries question the psychological reality of the heuristic process we have proposed, arguing that inherence-based explanations emerge via other processes. **Gelman & Meyer** suggest that these explanations stem from a *realist assumption* about patterns in the world, whereas for **Yzerbyt & Demoulin** their source can be found in a basic *need for coherence and predictability*. We will address each of these possibilities in turn but begin with a few general thoughts about the psychological reality of the IH process.

For us, one of the most appealing features of the IH proposal is its non-mysterious nature – its ability to account for an impressive range of cognitive behavior with a simple process whose components are all well understood and independently documented and whose psychological reality is beyond doubt. In a nutshell, we are proposing that *why* questions trigger a search for relevant information in memory; that this memory search pulls up the information that is most easily accessible, which is typically skewed toward inherent rather than extrinsic (e.g., historical, contextual) facts; and that, due to a tendency to satisfice and to lax supervision by working-memory-dependent (System 2) processes, this retrieval bias toward inherence ends up translating into a bias in the final product of the

cognitive system (i.e., explanations). The empirical foundation for each of these components is solid enough that it actually seems difficult to imagine how the process of generating explanations could proceed *without* them: It seems difficult to imagine how the search for explanations could proceed without a memory search, how this memory search could proceed without sensitivity to well-documented asymmetries in the accessibility of different types of information, and so on. Consequently, because any arguments against the psychological reality of the IH process would have to spell out why these basic elements are peripheral rather than central to the process of generating explanations, we suspect that such arguments face an uphill battle.

**Gelman & Meyer** suggest that the inherence-based explanations that we hypothesized to stem from the IH actually stem from a more basic assumption that uniformities in the environment are "real, stable, and nonaccidental" (para. 5). On this view, people are predisposed to see patterns as stable and natural, which leads them to search their memories for inherent facts that can rationalize this predisposition. Thus, inherence-based explanations may not be the result of the process we outlined but rather just a by-product of people's realist assumption about regularities in the world.

To begin, we note an important similarity between the two accounts. Memory retrieval is a part of **Gelman & Meyer**'s alternative account, just as it is of ours: In most circumstances, people generate inherence-based explanations by retrieving inherent facts from memory. The key difference from our proposal is that, for Gelman & Meyer, the search for relevant information in memory is not triggered *directly* by the prompt to explain (as it is in the IH process). Rather, there is an intermediary step: the activation of an assumption about the nature of reality, which then prompts a targeted search for inherent facts consistent with this assumption. However, if inherent facts are more likely to be retrieved anyway because of their greater accessibility, then what need is there for an extra assumption or bias that guides our explanatory intuitions toward these facts? That is, it seems unnecessary to invoke a new psychological entity, especially one whose origin and nature are unclear, when the prevalence of inherence-based explanations can be accounted for by a set of non-mysterious cognitive processes, such as those in the IH, whose reality has already been documented. The typical (inherence-based) output of the IH could potentially be *described* as a sort of assumption, but it would nevertheless be the IH process that is psychologically real.

Are there cases, however, where inherence-based explanations could stem *only* from a realist assumption? If such cases existed, they could provide reason to endorse this alternative proposal over ours. In that spirit, **Gelman & Meyer** raise the point that inherence-based explanations are generated even for *novel or unfamiliar facts*. Because the realist assumption they propose is formulated at an abstract level, it can easily apply to novel patterns. In contrast, since the IH relies on previous knowledge, it might appear that it cannot handle novel patterns because the shotgun would not be able to retrieve much of anything. Or would it? Let us consider the evidence that Gelman & Meyer have in mind here, which suggests that children generate inherence-based explanations for novel patterns such as that snakes have holes in their teeth or that girls are good

at a game called *gorp* (Cimpian & Markman 2009; 2011). Although these patterns are indeed unfamiliar to the 4-year-olds in these studies, many of the *constituents* of the patterns are actually familiar (e.g., snakes, teeth, girls, games), so there is quite a bit of information in memory for the shotgun to retrieve and pass on to the storyteller. In other words, even if the explanandum itself is novel, not all its constituents are novel, and thus there is almost always something for the shotgun to retrieve as raw material for the IH.[1] The shotgun would probably be thwarted only by *complete* novelty (e.g., Jabberwocky-style questions such as "why are borogoves mimsy?"), but of course a realist assumption would not be of much use here either, since it is impossible to explain such facts. In sum, the IH account has little difficulty accounting for the presence of inherence-based explanations for unfamiliar patterns, and thus the added value of a realist assumption is still in question.

It is also notable that the IH proposal covers a broader range of phenomena than **Gelman & Meyer**'s claim about a realist assumption. For example, our proposal subsumes the evidence for a correspondence bias – that is, the tendency to attribute instances of behavior to inherent traits more often than is warranted (e.g., Gilbert & Malone 1995; Jones & Harris 1967). Although the accessibility difference between inherent and extrinsic facts may be less dramatic here than in the case of broad patterns, inherent features of the person (or of people in general) are still likely to be easily retrieved by the shotgun, leading to their overuse – all other things being equal (that is, assuming the context [e.g., culture] does not make extrinsic facts more readily available than they would otherwise be). Gelman & Meyer's realist assumption is explicitly formulated as applying to patterns and regularities, and thus cannot provide an account of inherence-based explanations for instances. In sum, the IH process seems both more parsimonious than Gelman & Meyer's proposal (because it does not need to appeal to special assumptions of unclear origin) and better able to account for the full range of explanations that appear to be biased toward inherence.

**Yzerbyt & Demoulin** raise another possibility, suggesting that inherence-based explanations stem from a basic need to see the world as a "coherent and predictable" (para. 4) place that can be brought under one's control. People seek explanations in terms of inherent features because these explanations fill this need particularly well. Although Yzerbyt & Demoulin intended their proposal as a rival to our own, there are ways in which it could be complementary, rather than contradictory, to it. For example, if Yzerbyt & Demoulin are correct in assuming that inherence-based explanations provide a sense of control over one's environment (but see Jones 1979), then fulfillment of this basic need could provide a functional reason for the presence of an inherence heuristic in human cognition. We have already argued that the IH process is present in part because it fulfills the drive to *understand* the world. Yzerbyt & Demoulin's point suggests that it may also fulfill the (related) drive to *control* the world.[2]

At a more mechanistic level, this motivation to see the world as a stable, predictable place may also modulate the operation of the IH process. For example, high levels of this motivation, whether chronic or situation-specific, might enhance the accessibility of inherent facts in memory (e.g., Anderson et al. 1996), leading to a stronger inherence skew in the content of the shotgun; similarly,

high levels of this motivation may lead reasoners to endorse the inherence-based output of the IH immediately and without question. Thus, in addition to suggesting a possible reason *why* people rely on the IH, the motivation to see the world as predictable and controllable may be directly involved in the *how* of the heuristic – in the specifics of its operation.

However, **Yzerbyt & Demoulin**'s intention was to propose an *alternative process* for generating inherence-based explanations, not to simply offer suggestions regarding the process we proposed. Their alternative process follows a course similar to **Gelman & Meyer**'s: Explanations appealing to inherent features emerge from a memory search that (a) occurs only after the activation of the hypothesized need to see the world as stable and predictable, and (b) targets inherent features specifically, so as to generate explanations that fill this need. The first question we raise here is the one we raised about Gelman & Meyer's proposal: Is the motivation invoked by Yzerbyt & Demoulin necessary (at a mechanistic level) to produce inherence-based explanations? Most likely, it is not. The process outlined in our proposal, which relies on basic, well-understood cognitive components, is sufficient to generate inherence-based explanations *without any motivated influences*. Although this heuristic process *can* be influenced by reasoners' needs and goals, its output will be biased toward inherence even in the absence of such an intervention. What, then, is to be gained by invoking the need for an orderly, controllable world?

Even though this psychological motive may not be necessary to arrive at inherence-based explanations, perhaps it is *sufficient*. In that case, we would have to adjudicate between two alternative, individually sufficient paths to inherence-based explanations: the IH process and **Yzerbyt & Demoulin**'s motivated process. What reasons are there to favor one over the other? Yzerbyt & Demoulin provide two justifications for their preferred mechanism. The first is the occurrence of vague inherence-based explanatory intuitions (e.g., "there is *something* about girls that explains why they wear pink"). These intuitions are argued to be the initial explanatory products of people's need for a predictable world, before inherent features are found in memory (or discovered via later learning) that can be used to elaborate these initial intuitions into something less vague. But this point cannot differentiate between the two accounts: As explained in our target article, such vague intuitions are a potential output of the IH process as well. Specifically, they occur when the content retrieved by the shotgun cannot be assembled into a sensible explanation, and as a result reasoners are left with a vague impression that some of the (predominantly inherent) facts activated from memory will ultimately explain the observation being considered. Phenomenologically, this failure of the storyteller to come up with a concrete explanation will often translate into the intuitions Yzerbyt & Demoulin have in mind (e.g., "there is *something* about girls…"). Thus, the occurrence of these intuitions is not a reason to favor Yzerbyt & Demoulin's motivated model. The second justification Yzerbyt & Demoulin provided for this model was that, due to confirmation bias, inherence-based intuitions are resistant to change. However, it is unclear to us why the existence of a confirmation bias in human cognition should provide unique support for

Yzerbyt & Demoulin's motivated model. Once an inherence-based intuition has been generated, even a vague one, confirmation bias will make it resistant to change regardless of the process by which this intuition came about (e.g., Nickerson 1998). Consequently, Yzerbyt & Demoulin's proposed process cannot derive special support from the general inertia that characterizes human beliefs.

So far, we have argued that **Yzerbyt & Demoulin**'s reasons for favoring their motivated process over ours may not be convincing. Are there reasons to favor our proposal instead? To begin, we highlight the promise of the IH model via a brief historical analogy. When the correspondence bias was first documented in the late 1960s, one of the accounts offered was that this bias stems from the fundamental need to predict and control the world. Perceiving inherent dispositions in others' behaviors was thought to satisfy this need by organizing the complex stream of behavioral data into invariant "chunks" (i.e., dispositions) that could then be used to predict and influence future behaviors (see Heider 1958). Even in the early stages of research on the correspondence bias, though, few researchers believed that motivational drives could provide a sufficient account of this bias (e.g., Jones 1979; Ross 1977). Jones (1979), for example, expressed serious reservations about the need-for-control argument: "How can this ubiquitous proneness to attributional error really facilitate control? Why should we feel more in control when we think we understand a personal disposition than when we think we understand the situational context of behavior?" (p. 116). In part as a result of these doubts, research into the cognitive mechanisms underlying this bias continued unabated until ultimately settling on a model that appeals to fast heuristic processes operating on easily accessible information with loose supervision/correction by working-memory-dependent processes (e.g., Gilbert 2002; Gilbert et al. 1988; 2003; Trope & Gaunt 2000). Motivational factors such as the need for predictability and control are peripheral in contemporary accounts of the correspondence bias. The point of the analogy is this: If heuristic cognitive mechanisms were sufficient to account for the inherence bias in people's explanations for behavior, it seems plausible that such mechanisms will also be sufficient to account for the more general inherence bias under investigation here.

Consistent with this possibility, inherence-based explanations seem to behave exactly how one would expect the output of a heuristic (as opposed to a purely motivational) process to behave. First, these explanations are most prevalent in circumstances where inherent facts are most accessible and thus most likely to be overused for generating quick explanations (e.g., Cimpian & Erickson 2012; Cimpian & Markman 2009; 2011). Second, inherence-based explanations are judged to be more plausible when people's cognitive resources are taxed by a secondary task (Salomon & Cimpian, in preparation). Again, this is precisely what we would predict: Heuristic responses generally increase under cognitive load (e.g., Epley & Gilovich 2006; Gilbert et al. 1988). Third, inherence-based explanations are more strongly endorsed by people with lower scores on a fluid-intelligence test (Salomon & Cimpian 2014).[3] The same is true of the output of other intuitive heuristics (e.g., Stanovich & West 2000; see also Stanovich & West 2008). Fourth, inherence-based explanations are more strongly endorsed by children (Cimpian & Steinberg, in

press), a developmental difference that parallels that identified for other heuristics (e.g., Kokis et al. 2002; Toplak et al. 2014). Fifth, inherence-based explanations are evaluated more positively by people who tend to adopt less reflective or effortful thinking styles (Salomon & Cimpian 2014). The negative relationship between reflective thinking/cognitive styles and reliance on heuristic intuitions is also well established (e.g., Epley & Gilovich 2006; Stanovich & West 2000; Toplak et al. 2014). Together, these data speak to the plausibility of our heuristic model as a source of people's inherence-based intuitions; they also suggest that these intuitions are unlikely to be just the by-products of a motivation to see the world as predictable and controllable, as argued by **Yzerbyt & Demoulin**. (These data may also pose a challenge to **Gelman & Meyer**'s proposal of a realist assumption. It is unclear how a simple assumption about the nature of reality would be able to account for this sort of systematic variability in the prevalence of inherence-based explanations.)

To conclude, neither of the alternative sources of inherence-based intuitions proposed by our commentators (a realist assumption about regularities, and a need for predictability and control) appears to provide a viable alternative to the IH account.

## R2. The inherence heuristic and system justification

In this section, we discuss the potential role of the IH in promoting a tendency to defend the societal status quo. To reiterate, the typical output of this heuristic process consists of explanations that appeal to inherent facts. Regardless of what observations this output is used to explain, it will tend to make these observations seem natural and sensible (rather than contingent or arbitrary). If, for example, someone explains why OJ is for breakfast by appealing to its sleep-chasing tangy taste, then it is only a small further step to assume that this pairing of OJ with breakfast makes perfect sense. It should be apparent that this way of thinking has direct relevance to the phenomenon of system justification. When used to gain an understanding of one's society, the typical (inherence-based) output of the IH will tend to make whatever sociopolitical configuration is in place when the IH is triggered seem natural and sensible, and thus deserving of one's support. Thus, the IH can "fire system-justifying bullets" (**Gaucher & Jost**, para. 3) for completely non-motivated, non-ideological reasons – simply because inherent facts (or "facts") are often easily accessible in memory and are thus low-hanging fruit for a process that's looking for quick-and-dirty explanations. In other words, a motivation to defend the status quo is *not necessary* for status-quo-defending attitudes to emerge as a result of the IH process. This is not to say that motivational states are never involved in the operation of the IH – they can be, as explained in our target article, and they may thus contribute to the generation of system-justifying outputs. But the important point here is that defense motives are not needed for such outputs to emerge.

Our argument so far is that the IH process is sufficient to give rise to system justification – even in the absence of motivated influences. Given the ubiquity of explanation in

everyday life, we also suspect that the IH is a prolific source of system-justifying cognitions. We do not, however, endorse the more extreme claim that the IH is necessarily involved in *any* instance of system justification. Psychological states that reinforce the status quo may emerge via other processes as well. For example, exposure to the repeated pairing of certain groups with positively valenced markers of high status (e.g., wealth, power) could result, via simple associative learning, in similarly valenced emotions toward these high-status groups, regardless of one's own group membership (see **Newheiser & Olson**). Thus, the IH is a frequent entry point into system justification, but it is not the only one.

We now go on to address some of the points raised in the commentaries. **Gaucher & Jost** wondered "why – in the absence of social and motivational considerations – the mental shotgun fires system-justifying bullets" (para. 8). The answer is simple: because inherent, and thereby system-justifying, bullets (e.g., stereotypes about the inherent traits of different groups) are at its fingertips (e.g., Devine 1989). In effect, system-justifying explanations are a prime example of how easy heuristic judgments can really miss the mark, with serious consequences. We should clarify, however, that this heuristic framework for understanding system justification does not negate the role of defense motives in this phenomenon, a role that Gaucher & Jost and **Uhlmann, Zhu, Brescoll, & Newman** (**Uhlmann et al.**) illustrate with many convincing examples. Rather, our argument acknowledges the importance of these motives and fits them into a broader account that also spells out some of the possible cognitive underpinnings of people's tendency to defend the status quo. This account has additional advantages besides the deeper mechanistic understanding it may provide. For example, by highlighting the frequency and ease with which inherence-based explanations are generated, our proposal provides new insight into why defense motives co-opt these explanations with such remarkable regularity. Another advantage of our heuristic account is that it unifies a number of apparently disparate areas of research. To us, the possibility that far-flung phenomena such as the correspondence bias and system justification are even partially underlain by a simple explanatory heuristic is intellectually exciting and worthy of serious consideration.[4]

In their commentary, **Uhlmann et al.** endorse the existence of deep links between the IH and system justification but caution against equating one phenomenon with the other. We agree. As clarified earlier, we see the IH as one of multiple pathways to system justification. Other pathways include low-level, implicit associations of the sort described by **Newheiser & Olson**, as well as motivational influences that don't operate through the IH process: Although the motivation to defend the status quo can indeed skew the output of the heuristic, it almost certainly leads to system-justifying judgments by means unrelated to the IH as well (e.g., beliefs about karma or fate). The only potential point of disagreement with Uhlmann et al. arises with respect to their claim that system justification is *necessarily* motivated. Because the IH can produce inherence-based explanations for societal patterns without any motivated prompts, and because the content of these explanations can in and of itself justify existing societal arrangements, Uhlmann et al.'s claim that motives are definitional of system justification seems too strong.

**Newheiser & Olson** add a developmental layer to this discussion by reviewing some of their evidence of system justification in childhood (in particular, South African minority children's striking outgroup preferences) and by raising interesting questions about how this evidence should be interpreted: Does one need to invoke inherence-based explanations to account for children's implicit preference for high-status groups, or would low-level associations suffice? Newheiser & Olson argue that associations may be all that is needed to explain their data, and we agree. However, this is not to say that low-level associations and the ensuing implicit preferences for high-status groups exhaust the full extent of system justification in childhood. As **Bigler & Clark** remind us, children not only detect regularities in their social world but also endeavor to explain these regularities. If the IH is part of the process by which these regularities are explained (and Bigler & Clark agree that it most likely is), then we would expect to see stronger, more explicit forms of system justification in childhood as well. That is, we would expect children to display not only implicit preferences for high-status groups (of the sort described by Newheiser & Olson) but also explicit cognitions that frame existing societal structures as fair because they are the natural by-products of the relevant groups' inherent features. Hints in the existing literature suggest that such cognitions may indeed be present: Consider, for instance, children's tendency to legitimize the dearth of women and people of color in positions of authority by appealing to these groups' inherent features (e.g., lower intelligence; Bigler et al. 2008), or the fact that 4-year-olds explain even novel facts about social groups (e.g., that girls are good at *gorp*) in terms of inherent traits (Cimpian & Erickson 2012; Cimpian & Markman 2011). In addition to these clues, direct tests of our account's developmental predictions are currently under way, with promising results (Hussak & Cimpian 2013). Because some of these tests involve novel social categories, we are also able to rule out the possibility that children's system-legitimizing explanations occur only as a means of rationalizing an already-acquired implicit preference for the high-status groups (as argued by Newheiser & Olson).

To summarize, we argued that the IH is a sufficient – and powerful – source of system-justifying attitudes, with roots that stretch deep into early childhood.

## R3. The inherence heuristic and is–ought inferences

The typical (inherence-based) output of the IH leads to a tendency to defend the status quo in part because it makes the status quo seem sensible and appropriate – *the way things should be*. If the tangy taste of OJ is invigorating, then it stands to reason that OJ *should* be consumed for breakfast (as opposed to, say, dinner). Inferences about *shoulds* and *oughts* are even easier to make when the inherence-based output of the IH already relies on value-laden explanatory notions (e.g., the taste of OJ makes it ideal for breakfast) as opposed to causal notions (see, e.g., sect. 2.2.3 of the target article). For these reasons, we argued that the IH may be one of the psychological mechanisms underlying the frequent transitions from *is* to *ought* that Hume (1740/2000) observed in his

contemporaries' reasoning – transitions that seem no less common today. In this section, we clarify a few aspects of our argument and then discuss a proposal for an alternative source of is–ought inferences (brought up by **Bartsch & Estes**).

First, it is important to note that our hypothesis concerns human psychology, not metaethical principles: The IH proposal explains why people are prone to take what is the case as a signal of what should be the case. However, this proposal does not commit us to any particular position on the issue of whether moral truths can in fact be legitimately derived from statements of fact. Although we suspect that many is–ought inferences are indefensible (which is why we occasionally used the term *is–ought errors* in the target article), we do not want to be taken as saying that moral judgments are necessarily divorced from, and altogether a different sort of entity than, empirical judgments (for defenses of naturalistic moral realism, see Boyd 1988, Railton 1986, and many others). We thank **Dubljević & Racine** for the prompt to clarify our position on this issue.

We also want to clarify a related matter that came up in **Dubljević & Racine**'s commentary: The IH is a psychological process that informs everyday explanatory (and, by extension, moral) reasoning. We are more hesitant to read the operation of the IH into the deliberative reasoning processes that, say, philosophers might rely on in constructing their arguments about what makes something morally virtuous or reprehensible. Thus, we would caution against using examples such as those mentioned by Dubljević & Racine (e.g., Kant's deontological arguments) as illustrations of heuristic reasoning. This is not to say that vestiges of intuitive heuristics (including the IH) are never present in scientific or philosophical arguments. We suspect that they are, but differentiating between the products of heuristic and deliberative thought becomes more difficult in these circumstances.

To reiterate, we are arguing that is–ought inferences may stem from the operation of the IH. **Bartsch & Estes** propose an alternative source for these inferences, particularly as they occur in children's reasoning. Specifically, Bartsch & Estes argue that children may assume that the world is exactly as it should be because it conforms to human goals and intentions: "If the way things are is the way people *want* them to be, then maybe it's all good" (para. 6). We are in complete agreement with Bartsch & Estes regarding the centrality of reasoning about mental states in children's cognitive life. With respect to explanations per se, however, we argue that such mentalistic reasoning is not as ubiquitous as Bartsch & Estes claim. We first detail how this argument follows from the IH proposal. We then review some of the evidence on this point, which seems to favor our account over Bartsch & Estes's.

According to the IH proposal, the content of people's everyday explanations tends to be supplied by shallow memory searches that often start with the main constituents of the explanandum as retrieval cues. For example, if children are trying to explain why OJ is for breakfast, they may start by retrieving any easily accessible facts about OJ and breakfast. Although these most-accessible facts may occasionally involve mental states (e.g., Mom likes OJ), in many circumstances the shotgun will retrieve only facts about the relevant objects' inherent features (e.g., OJ is sour), which are the sort of facts that dominate our semantic representations (e.g., McRae et al. 1997;

2005). As a result of this retrieval bias, heuristic explanations for facts about the world may actually appeal to human decisions less, rather than more, often than would be normatively warranted – the opposite of **Bartsch & Estes**'s prediction.

Is there evidence that can adjudicate between these claims? To begin, although **Bartsch & Estes** invoke childhood *artificialism* (a hypothesized tendency to see human agency at the origin of most phenomena and events) to support their argument, we note that Piaget's (1929/1967) conclusions on this topic were contradicted by subsequent work (e.g., Gelman & Kremer 1991). In fact, our reading of the relevant developmental literature points to the conclusion that children typically *fail* to see the human agency behind many features of their world that in reality are the result of human intentions and decisions. (We are not alone in this interpretation of the literature. For instance, the realist assumption proposed by **Gelman & Meyer** has a similar flavor.) As an illustration, children often fail to understand that certain aspects of the world, such as modes of dress or word–referent mappings, are due to mutable social conventions (e.g., Brook 1970; Gabennesch 1990; Kalish 1998; Lockhart et al. 1977). Instead, children seem to understand many of these conventions as permanent fixtures of the world that cannot be changed even if social consensus favored such a change (that is, even if people wanted a change). Similar intuitions were also documented in a recent series of studies investigating the developmental course of the IH (Cimpian & Steinberg, in press). As would be expected if the IH were present early in development, preschool-age children often denied that coins are round, fire trucks are red, and so forth, because people wanted them to be so, endorsing instead explanations that appealed to the inherent natures of the relevant objects. In summary, the totality of the developmental evidence suggests to us that an appreciation of human intentions is not as central to children's explanations of the world as Bartsch & Estes argue. Rather, children tend to assume that the origin of many environmental uniformities can be found *within* these uniformities themselves (specifically, in the features of their constituents), as predicted by our IH account.

## R4. The inherence heuristic and psychological essentialism

In our target article, we proposed that the IH, with its tendency to output explanations that appeal to inherent facts, could serve as a precursor to psychological essentialism. This was perhaps the riskiest move in our paper, in no small measure because essentialism is a moving target. As was clear from our commentaries, there is no general agreement on what counts as essentialism; and if we don't know what essentialism is, it is of course impossible for any theory (including ours) to successfully explain how it emerges. In what follows, we provide a brief analysis of the construct of essentialism as we see it reflected in the commentaries. Based on this analysis, we conclude that this construct has become so diluted and broad that it is unlikely to be picking out a *single* psychological phenomenon. We then clarify which of the phenomena the *essentialism* label currently refers to we propose to explain with the IH. Finally, we answer the questions that came up

regarding the developmental process by which the IH leads to the emergence of these specific essentialist phenomena.

### R4.1. What is essentialism?

We share **Haslam**'s impression that the term *essentialism* has been applied very liberally in recent years, to the point where it has become hard to interpret and perhaps no longer useful for further theorizing. As an illustration of this issue, next we list the disparate intuitions that our commentators labeled as essentialist. To preview, the only feature these intuitions all seem to share is that they involve an appreciation for a reality beyond the obvious – for the fact that there is more to the world than meets the eye. But do these intuitions necessarily stem from the same psychological process? This is the crucial question if we are to continue studying essentialism as a unitary phenomenon. Unfortunately, it seems to us that the answer to it is probably "no." In fact, a variety of different cognitive processes lead people to rely on nonobvious factors, including many whose outputs have not (so far, at least) been included under the essentialism rubric (e.g., reasoning about others' [nonobvious] mental states). Thus, if the only feature shared by all phenomena labeled as *essentialism* is an appreciation for the nonobvious, and if this feature can arise from a diversity of sources, it seems hopeless to attempt a unified theory of essentialism. Instead, more progress might be made if we first grouped essentialist intuitions (i.e., intuitions that appeal to nonobvious factors) by their suspected etiology/origin and then investigated each of the resulting groups independently. This sort of classification, however, goes beyond what we can accomplish in the present response. Here, we limit ourselves to making the point about the diverse etiologies of essentialist intuitions and then carving out a portion of these intuitions for the IH to explain.

In this spirit, here are some of the intuitions that were mentioned in the commentaries as instances of essentialism:

*Intuition 1.* A thing's category membership is given by a nonobvious physical entity present in it. For instance, Fido is a dog, and possesses dog-typical properties, by virtue of some of its DNA. This seems to be the intuition that is most widely classified as essentialist (e.g., **Dennehy**; **Haslam**; **Hood**; **Noles & Danovitch**; **Olivola & Machery**).

*Intuition 2.* A thing's unique identity is given by a nonobvious physical entity present in it (**Hood**). For instance, Fido is the individual it is, with its unique properties, by virtue of some of its DNA.

*Intuition 3.* A thing's unique identity is given by a nonobvious "essential element" that picks out the thing across space and time – perhaps something akin to a mental index that tracks it (**Hood**, para. 4). For instance, the ship of Theseus maintains its identity over a complete change in its material composition because it retains this essential identity-preserving element. The essence in this intuition is nonphysical (because it is completely independent of the object's material composition), which contrasts with Intuition 2 (where the essence inheres in the object).

*Intuition 4.* A thing's value (monetary, sentimental, etc.) is given in part by a nonobvious entity obtained via the thing's relations to other entities (e.g., oneself, a famous person, a historical event; see **Hood** and **Noles & Danovitch**). For instance, a pen that we own, a pen that belonged to Beyoncé at some point, and a pen that was used to sign the U.S. Constitution are all endowed with some nonobvious entity (an essence) that gives them a "special status" (Noles & Danovitch, para. 3) and makes them more valuable than pens that are molecule-for-molecule identical. This is the essentialist intuition often invoked in discussions of ownership and authenticity (e.g., Gelman 2013; Hood & Bloom 2008). Essences account for the *value* of an object here rather than for its category membership (Intuition 1), physical features (Intuitions 1 and 2), or unique identity (Intuitions 2 and 3). These value-conferring essences are acquired or lost over the course of the object's history, depending on the presence of the relevant relations (e.g., whether the object is in one's possession). This fluidity contrasts with Intuitions 1–3, where the essences seemed more than just temporary attributes that an object can gain or lose with the circumstances. Also of note, the essences involved here need not be physical. When a person assumes ownership of an object, for example, the object appears to acquire whatever essential entity that gives it its special status even without any physical contact with its new owner (and thus presumably without the possibility of any physical changes occurring to it).

*Intuition 5.* There is a deeper, nonobvious reality behind the surface data (**Bookstein**). This seems to be the cognitive impulse behind much scientific activity, perhaps even down to the level of the statistical methods, as Bookstein points out. For example, the logic behind something as simple as calculating the mean of a number of observations seems to presuppose the existence of a true value of which these observations are indicative. This essentialist intuition places almost no constraint on what the deeper reality (or the essence) might be – just that it exists and that it is the source of the observed data.

These are just the examples of essentialism that came up in our commentaries – more can be found in the broader literature, but the present set is sufficient to make our point: Although using the term *essentialism* to refer to all of these intuitions gives the impression that they are all outputs of the same psychological process, it is hard to see how that could be the case. Given what is known about our species' neurocognitive architecture, it seems implausible, for example, that a process that tracks object identity across space and time (Intuition 3) would be one and the same as a process that computes the subjective reward value of "emotional objects" (**Hood**, para. 4; Intuition 4) or a process that provides causal explanations for a category's features (Intuition 1); these judgments all involve different types of computations that are likely to be instantiated in different brain substrates. Thus, rather than forming a coherent *essentialism* category, judgments that assume the presence of nonobvious entities appear to arise via a range of different mechanisms. Some of these judgments (such as Intuition 5 about the presence of a deeper reality) might stem from the basic human motivation to understand and explain what is observed (e.g., Callanan & Oakes 1992; Murphy & Medin 1985; Schulz

2012). Others (such as Intuition 3 about identity maintenance across material changes) may grow out of attentional tracking processes of the sort that have been reported in the visual cognition literature (e.g., Pylyshyn 1989; Scholl 2001). Yet others (such as Intuition 4 about value-conferring essences) may develop in part as elaborations of early-developing mechanisms for tagging psychological agents with value based on their actions (e.g., helping vs. hindering) (e.g., Hamlin et al. 2007; Kuhlmeier et al. 2003). And these are probably only a small subset of the cognitive processes whose outputs appeal to nonobvious entities and could thus fall under the scope of essentialism.

The upshot of this argument is that no single cognitive mechanism can account for the full complement of intuitions that are currently called *essentialist*. This conclusion applies to the IH mechanism as well. Thus, our claim that the IH serves as a foundation for psychological essentialism must be understood in conjunction with the definition of essentialism we used in the target article, which roughly equated essentialism with Intuition 1 (about the existence of physical essences that explain category membership and the features associated with it).[5] As a general explanatory process whose output overuses inherent facts, the IH possesses the computational resources – and in fact seems uniquely suited – to lay the groundwork for the development of Intuition 1. The IH cannot, however, supply the computations that are at the core of many other essentialist intuitions (e.g., identity tracking) and is therefore unlikely to be their precursor.

### R4.2. How might the inherence heuristic lead to the development of (Intuition 1) essentialism?

We proposed that Intuition 1 (which applies to natural and social kinds) may emerge as an elaboration of the IH's output over the first few years of life. Briefly, this hypothesized transition might occur as follows: Initially, children invoke the IH to explain the uniformities observed across the members of a category. In many cases, however, children may not be able to assemble the inherent facts pulled up by the shotgun into coherent explanations; as a result, children may be left with a vague impression that some of these inherent facts will ultimately be sufficient to explain the uniformities in question. At this point, children haven't yet developed Intuition 1 because they are not conceiving of the characteristics of the relevant categories as emerging from a single, internal source. However, the unresolved, to-be-determined output of the IH may in fact account for other essentialist intuitions, particularly ones that impose fewer constraints on the location and nature of the explanatory "essences." For example, this early output of the IH might be exactly what **Strevens** glosses as a belief in the "causal efficacy of category membership" (para 7).[6] In other words, the heuristic impression that *some* inherent facts about the members of the category will eventually explain their characteristics could easily be mistaken for, and thus described as, a belief in the causal efficacy of category membership per se. This is only a conjecture at this point, though, so it will be important to test empirically whether we are correct: whether what Strevens describes as an explanatory appeal to the brute fact of category membership (e.g., the mere fact of being dogs causes dogs to bark) is actually an appeal to inherent facts about the category (e.g., some features of dogs cause them to bark).

So how does Intuition 1 emerge out of the early intuitions supplied by the IH? The unresolved nature of many of the heuristic's outputs over the first few years of life may prompt children to look for plausible means of arriving at more specific (and perhaps more satisfying) explanations. We argued that children could use at least two sources of early causal knowledge to elaborate the vague output of the IH, both of which would also bring this output closer in content to Intuition 1. In particular, children could rely on their preexisting causal beliefs about insides (e.g., Newman et al. 2008; Setoh et al. 2013) and internal energy (e.g., Gottfried & Gelman 2005; Hatano & Inagaki 1994; Morris et al. 2000). Both of these are broad causal principles that help children reason about the basic functioning of living things (e.g., their movement, their growth). However, when used as a means of refining the output of the IH with respect to why the members of a particular category have the features they do, these causal forces will be translated into category-specific versions of their general forms. For example, even though children might at first conceive of insides as the completely category-general causal force behind animal movement (with no distinctions made between the insides of dogs, cats, birds, etc.), when children later use this belief in their attempt to determine which inherent fact(s) *about dogs* might explain their characteristic features, it will necessarily take on a category-specific form (that is, children will need to appeal to *dog* insides and their unique causal powers). Our claim is that, by way of such refinement processes, children gradually arrive at intuitions that can be described as Intuition 1.

Several commentators had questions about this elaboration process. For example, **Marmodoro et al.** wondered, among other things, what motivates the transition to essences (is it something that children find "rationally compelling" [para. 5], or is it learned?), what is gained by an appeal to essences, and why children should assume that a single essential property explains the rest of a category's properties. On our account, the transition to (Intuition 1) essentialism is at least partially driven by a motivation to resolve the ambiguity caused by the storyteller's inchoate output and thus to find a more satisfying explanation for the uniformities observed within a category. Elaborating this ambiguous output with the help of preexisting notions about insides and/or energy may in fact be a rationally compelling (or at least defensible) means of identifying a more specific, and thereby more satisfying, explanation. Thus, we suspect that essentialist beliefs are in large part constructed rather than learned, especially since young children encounter very little talk about essences or essencelike causes (Gelman et al. 1998; but see Gelman et al. 2010; Rhodes et al. 2012). (Of course, later learning could alter the content of these early essentialist beliefs, such as when children start learning about genetics.) As the preceding discussion suggests, what children gain by transitioning over to Intuition 1 is (what appears to them to be) a more satisfying explanation than the one that the IH was able to provide. Finally, note that the emphasis on a single essential feature is simply a by-product of this elaboration process, not a goal of it. Because notions about insides and/or energy are available to guide children's search for a more specific explanation, and because these

notions rely on a single causally efficacious entity, children *end up* with beliefs about a single essential feature causing many others.

**Kinzler & Sullivan** asked whether the elaboration process we are proposing involves true conceptual change or simply the enrichment of existing conceptual content. Given the cognitive leaps involved in (a) bringing prior beliefs about insides and/or energy to bear on the output of the IH, and in (b) creatively twisting these category-general causal forces into category-specific explanatory roles, we might classify this process as an instance of conceptual change. Point (b) is also relevant to a comment by **Bastian**, who questioned whether internal energy, which is a general and fluid causal force that does not distinguish between kinds, could be involved in the process by which kind-specific essences emerge. We agree with Bastian about the category-general nature of internal energy and other presumed vitalistic processes. In fact, this is precisely the reason why we emphasized that children's beliefs about internal energy would have to be *twisted* or *distorted* to fit the mold of an essencelike inherent feature. This may be a big conceptual leap for children, but the considerable causal power usually assigned to this energy makes it a very appealing means of refining the output of the IH into a plausible story.

Finally, several commentators pointed to potential social influences on the content and endorsement of essentialist beliefs (**O'Connor & Joffe**; **Olivola & Machery**; **Rhodes**). As detailed throughout this response, the operation of the IH is sensitive to such sociocultural factors. As a result, our account is in principle capable of accounting for their influence on the heuristic's offshoots.

As a final note, we shamelessly highlight **Haslam**'s assessment of the present proposal as a "useful and potentially generative idea" that "deserves serious consideration by psychologists who study essentialist thinking" (para. 8). Naturally, we share Haslam's optimism. And, since the bulk of empirical work is ahead of us, we hope that the answers provided here have clarified our theoretical position to a point where others may be enticed to join us in testing its predictions.

### NOTES

**1.** There is no guarantee, however, that the storyteller will be able to shape the content at its disposal into an explanation. Whenever it fails to do so, the output of the IH will be the vaguer, "to-be-determined" sort of intuition – namely, that some of the facts retrieved from memory are relevant to an explanation in a to-be-determined way.

**2.** At this point, the most we can say is that the IH fulfills these needs by providing a *subjective sense* of understanding and control. Whether the IH truly enhances people's understanding and control of the environment depends on further questions about its normative accuracy.

**3. Yzerbyt & Demoulin** asked whether the IH scale that we described in our target article might measure people's adherence to social norms rather than a reliance on inherence-based explanations. Although this scale correlates with scales measuring people's adherence to social norms (as our account would predict as well), we have evidence that it also taps directly into people's explanations of the world. For example, in a construct validation study described by Salomon and Cimpian (2014), we asked people to justify why they agree or disagree with our scale items. As we would expect, agreement with our scale items was overwhelmingly justified with inherent reasons (e.g., traffic lights use red to signal stop because "red is a warning sign in

nature"). Conversely, disagreement with the scale items was justified with extrinsic reasons (e.g., "we have learned that red means stop so I feel this is why red means stop").

**4.** As a note of historical interest, we are not the first to posit deeper commonalities between the correspondence bias and system justification. In the last paragraph of their classic article on the fundamental attribution error (which we refer to as *correspondence bias* here), Ross et al. (1977) speculate that "this distortion in social judgment [i.e., the bias to attribute behavior to corresponding inherent traits] could provide a particularly insidious brake upon social mobility, whereby the disadvantaged and powerless overestimate the capabilities of the powerful who, in turn, inappropriately deem members of their own caste well-suited to their particular leadership tasks" (p. 494).

**5.** The IH might also be at the source of intuitions similar to Intuition 2 (about the physical essences of unique individuals), especially since the IH is also triggered to explain specific features or behaviors. We leave that argument for another occasion, however.

**6.** We should also reiterate that *inherent* is not synonymous with *internal*, and thus that **Strevens**'s arguments against what he calls *insides essentialism* miss their target when directed at our proposal.

## References

[The letters "a" and "r" before author's initials stand for target article and response references, respectively]

Ahn, W., Kalish, C., Gelman, S. A., Medin, D., Luhmann, C., Atran, S., Coley, J. D. & Shafto, P. (2001) Why essences are essential in the psychology of concepts. *Cognition* 82:59–69. doi:10.1016/S0010-0277(01)00145-7. [aAC, TCD]

Allport, G. W. (1954) *The nature of prejudice*. Addison-Wesley. [MR]

Anderson, C. A., Krull, D. S. & Weiner, B. (1996) Explanations: Processes and consequences. In: *Social psychology: Handbook of basic principles*, ed. E. T. Higgins & A. W. Kruglanski, pp. 271–96. Guilford Press. [arAC]

Aristotle (1890) *The ethics of Aristotle: The Nichomachean ethics*, trans. D. P. Chase. W. Scott. [VD]

Astington, J. W. (2001) The paradox of intention: Assessing children's metarepresentational understanding. In: *Intentions and intentionality: Foundations of social cognition*, ed. B. F. Malle, L. J. Moses & D. A. Baldwin, pp. 85–103. MIT Press. [KB]

Astuti, R., Carey, S. & Solomon, G. (2004) Constraints on conceptual development: A case study of the acquisition of folkbiological and folksociological knowledge in Madagascar. *Monographs of the Society for Research in Child Development* 69(3, whole no. 277). [ASB]

Atran, S. (1993) *Cognitive foundations of natural history: Towards an anthropology of science*. Cambridge University Press. [aAC]

Atran, S. (1998) Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences* 21:547–609. [aAC, ASB, SAG]

Baetu, I. & Baker, A. G. (2009) Human judgments of positive and negative causal chains. *Journal of Experimental Psychology: Animal Behavior Processes* 35 (2):153–68. [AM]

Baillargeon, R. (1994) Physical reasoning in young infants: Seeking explanations for impossible events. *British Journal of Developmental Psychology* 12:9–33. [rAC]

Baillargeon, R. (2004) Infants' physical world. *Current Directions in Psychological Science* 13:89–94. [rAC]

Baillargeon, R., Scott, R. M. & He, Z. (2010) False-belief understanding in infants. *Trends in Cognitive Sciences* 14:110–18. [rAC]

Baker, A. G., Murphy, R. A. & Vallée-Tourangeau, F. (1996) Associative and normative models of causal induction: Reacting to versus understanding cause. In: *The psychology of learning and motivation, vol. 34*, ed. D. R. Shanks, K. J. Holyoak & D. L. Medin, pp. 1–45. Academic Press. [AM]

Bargh, J. A. (1999) The cognitive monster: The case against controllability of automatic stereotype effects. In: *Dual-process theories in social psychology*, ed. S. Chaiken & Y. Trope, pp. 361–82. Guilford Press. [aAC]

Bargh, J. A., Bond, R. N., Lombardi, W. J. & Tota, M. E. (1986) The additive nature of chronic and temporary sources of construct accessibility. *Journal of Personality and Social Psychology* 50:869–78. [aAC]

Baron, A. & Banaji, M. (2009) Evidence of system justification in young children. *Social and Personality Psychology Compass* 3:918–26. [aAC]

Baron, J. (2005) *Rationality and intelligence*. Cambridge University Press. [aAC]

Bartsch, K. & Wellman, H. M. (1995) *Children talk about the mind*. Oxford University Press.   [KB]

Bastian, B. & Haslam, N. (2006) Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology* 42:228–35. doi:10.1016/j.jesp.2005.03.003.   [aAC, TCD, RM]

Bastian, B. & Haslam, N. (2007) Psychological essentialism and attention allocation: Preferences for stereotype-consistent versus stereotype-inconsistent information. *Journal of Social Psychology* 147:531–41. doi:10.3200/SOCP.147.5.531-542.   [BB, TCD]

Bastian, B. & Haslam, N. (2008) Immigration from the perspective of hosts and immigrants: Roles of psychological essentialism and social identity. *Asian Journal of Social Psychology* 11:127–40. doi:10.1111/j.1467-839X.2008.00250.x.   [BB, TCD]

Bastian, B., Bain, P., Buhrmester, M. D., Gómez, A., Vázquez, A., Knight, C. & Swann, W. B. (2013) Moral Vitalism: Seeing good and evil as real, agentic forces. Unpublished manuscript.   [BB]

Bastian, B., Jetten, J. & Fasoli, F. (2011) Cleansing the soul by hurting the flesh: The guilt-reducing effect of pain. *Psychological Science* 22:334–35.   [BB]

Beauchamp, T. & Childress, J. (2009) *Principles of biomedical ethics*, 6th ed. Oxford University Press.   [VD]

Beggan, J. (1992) On the social nature of nonsocial perception: The mere ownership effect. *Journal of Personality and Social Psychology* 62:229–37.   [NSN]

Ben-Zeev, A. & Dennehy, T. C. (2014) When boys wear pink: Gendered color cues evoke risk taking. *Psychology of Men and Masculinity*. Advance online publication. doi:10.1037/a0034683.   [TCD]

Bertulis, A. & Bulatov, A. (2001) Distortions of length perception in human vision. *Biomedicine* 1:3–26.   [GW]

Bialystok, E. (1988) Levels of bilingualism and levels of metalinguistic awareness. *Developmental Psychology* 4:560–67.   [KDK]

Bialystok, E., Craik, F. I. M., Klein, R. & Viswanathan, M. (2004) Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging* 19:290–303.   [rAC]

Bigler, R. S., Arthur, A. E., Hughes, J. M. & Patterson, M. M. (2008) The politics of race and gender: Children's perceptions of discrimination and the U.S. presidency. *Analyses of Social Issues and Public Policy* 8:1–30.   [RSB, rAC]

Bigler, R. S. & Liben, L. S. (2006) A developmental intergroup theory of social stereotypes and prejudice. *Advances in Child Development and Behavior* 34:39–89.   [RSB, rAC]

Bigler, R. S. & Liben, L. S. (2007) Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science* 16:162–71.   [RSB, rAC]

Bigler, R. S. & Wright, Y. F. (2014) Reading, writing, arithmetic, and racism? Risks and benefits to teaching children about intergroup biases. *Child Development Perspectives* 8:18–23.   [RSB, rAC]

Birnbaum, D., Deeb, I., Segall, G., Ben-Eliyahu, A. & Diesendruck, G. (2010) The development of social essentialism: The case of Israeli children's inferences about Jews and Arabs. *Child Development* 81:757–77.   [CO]

Blanchar, J. & Eidelman, S. (2013) Perceived system longevity increases system justification and the legitimacy of inequality. *European Journal of Social Psychology* 43:238–45.   [DG]

Bloch, M., Solomon, G. & Carey, S. (2001) Zafimaniry: An understanding of what is passed on from parents to children: A cross-cultural investigation. *Journal of Culture and Cognition* 1(1):43–68.   [ASB]

Bloom, P. (1996) Intention, history, and artifact concepts. *Cognition* 60:1–29.   [aAC]

Bloom, P. (2000) *How children learn the meanings of words*. MIT Press.   [aAC]

Bloom, P. (2004) *Descartes' baby: How the science of child development explains what makes us human*. Basic Books.   [aAC]

Bloom, P. (2010) *How pleasure works: Why we like what we like*. W. W. Norton.   [BH]

Bookstein, F. L. (2014) *Measuring and reasoning: Numerical inference in the sciences*. Cambridge University Press.   [FLB]

Boroditsky, L., Schmidt, L. A. & Phillips, W. (2003) Sex, syntax, and semantics. In: *Language in mind: Advances in the study of language and thought*, ed. D. Gentner & S. Goldin-Meadow, pp. 61–79. MIT Press.   [aAC]

Boyd, R. (1999) Homeostasis, species, and higher taxa. In: *Species: New interdisciplinary essays*, ed. R. A. Wilson, pp. 141–85. MIT Press.   [JAH]

Boyd, R. N. (1988) How to be a moral realist. In: *Essays on moral realism*, ed. G. Sayre-McCord, pp. 181–228. Cornell University Press.   [rAC]

Brescoll, V. & LaFrance, M. (2004) The correlates and consequences of newspaper reports of research on sex differences. *Psychological Science* 15:515–20.   [CO].

Brescoll, V. L. & Uhlmann, E. L. (2013) [Unpublished raw data.] Yale University.   [ELU]

Brescoll, V. L., Uhlmann, E. L. & Newman, G. N. (2013) The effects of system-justifying motives on endorsement of essentialist explanations for gender differences. *Journal of Personality and Social Psychology* 105:891–908.   [DG, ELU]

Brook, J. S. (1970) A test of Piaget's theory of "nominal realism." *Journal of Genetic Psychology* 116:165–75.   [arAC]

Bublitz, D. & Prasada, S. (2013) Normative dimensions of generic knowledge. Poster presented at the 35th Annual Cognitive Science Conference, Berlin, Germany, July 31–August 3.   [SP]

Bullot, N. J. & Rebar, R. (2013) The artful mind meets art history: Toward a psycho-historical framework for the science of art appreciation. *Behavioral and Brain Sciences* 36:123–80.   [NSN]

Cacchione, T., Hrubesch, C., Call, J. & Rakoczy, H. (submitted) Great apes are psychological essentialists.   [HR]

Cacchione, T., Schaub, S. & Rakoczy, H. (2013) Fourteen-month-old infants infer the continuous identity of objects on the basis of non-visible causal properties. *Developmental Psychology* 49(7):1325–29. doi:10.1037/a0029746.   [HR]

Cacioppo, J. T., Petty, R. E. & Kao, C. F. (1984) The efficient assessment of need for cognition. *Journal of Personality Assessment* 48:306–307.   [aAC]

Callahan, D. (1996) Can nature serve as a moral guide? *Hastings Center Report* 26 (6):21–22.   [VD]

Callanan, M. A. & Oakes, L. M. (1992) Preschoolers' questions and parents' explanations: Causal thinking in everyday activity. *Cognitive Development* 7:213–33.   [rAC]

Carey, S. (1996) Cognitive domains as modes of thought. In: *Modes of thought: Explorations in culture and cognition*, ed. D. R. Olson & N. Torrance, pp. 187–215. Cambridge University Press.   [aAC]

Carey, S. (2009) *The origin of concepts*. Oxford University Press.   [SAG]

Cashmore, A. R. (2010) The Lucretian swerve: The biological basis of human behavior and the criminal justice system. *Proceedings of the National Academy of Sciences USA* 107:4499–504.   [BB]

Cattell, R. B. (1986) *Intelligence: Its structure, growth, and action*. Elsevier.   [aAC]

Chen, S. & Chaiken, S. (1999) The heuristic–systematic model in its broader context. In: *Dual-process theories in social and cognitive psychology*, ed. S. Chaiken & Y. Trope, pp. 73–96. Guilford Press.   [aAC]

Chomsky, N. (1959) A review of B. F. Skinner's *Verbal Behavior*. *Language* 35:26–58.   [rAC]

Cimpian, A. (2010) The impact of generic language about ability on children's achievement motivation. *Developmental Psychology* 46:1333–40.   [aAC]

Cimpian, A. (2013) Generic statements, causal attributions, and children's naive theories. In: *Navigating the social world: What infants, children, and other species can teach us*, ed. M. R. Banaji & S. A. Gelman, pp. 269–74. Oxford University Press.   [aAC]

Cimpian, A., Arce, H. C., Markman, E. M. & Dweck, C. S. (2007) Subtle linguistic cues affect children's motivation. *Psychological Science* 18:314–16.   [aAC]

Cimpian, A. & Cadena, C. (2010) Why are dunkels sticky? Preschoolers infer functionality and intentional creation for artifact properties learned from generic language. *Cognition* 117:62–68.   [aAC]

Cimpian, A. & Erickson, L. C. (2012) The effect of generic statements on children's causal attributions: Questions of mechanism. *Developmental Psychology* 48:159–70.   [arAC]

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.   [arAC, SAG]

Cimpian, A. & Markman, E. M. (2011) The generic/nongeneric distinction influences how children interpret new information about social others. *Child Development* 82:471–92.   [arAC]

Cimpian, A., Mu, Y. & Erickson, L. C. (2012) Who is good at this game? Linking an activity to a social category undermines children's achievement. *Psychological Science* 23:533–41.   [aAC]

Cimpian, A. & Petro, G. (2014) Building theory-based concepts: Four-year-olds preferentially seek explanations for features of kinds. *Cognition* 131:300–10.   [rAC]

Cimpian, A. & Scott, R. M. (2012) Children expect generic knowledge to be widely shared. *Cognition* 123:419–33.   [aAC]

Cimpian, A. & Steinberg, O. D. (in press) The inherence heuristic across development: Systematic differences between children's and adults' explanations for everyday facts. *Cognitive Psychology*. doi:10.1016/j.cogpsych.2014.09.001   [arAC]

Clark, C. M. & Bigler, R. S. (2013) The emergence and trajectory of heterosexist attitudes towards relationships during early childhood. Poster presented at the First International Conference on LGBT [lesbian, gay, bisexual, and transgender] Psychology and Related Fields, Lisbon, Portugal.   [RSB]

Clark, E. V. (1988) On the logic of contrast. *Journal of Child Language* 15:317–35.   [aAC]

Coleman, J. & Hong, Y. (2008) Beyond nature and nurture: The influence of lay gender theories on self-stereotyping. *Self and Identity* 7:34–53.   [CO]

Cox, D. M. (2006) *Principles of statistical inference*. Cambridge University Press.   [FLB]

Crisp, R. J. & Meleady, R. (2012) Adapting to a multicultural future. *Science* 336:853–55. doi:10.1126/science.1219009.   [RM]

Crisp, R. J. & Turner, R. N. (2011) Cognitive adaptation to the experience of social diversity. *Psychological Bulletin* 137:242–66. doi:10.1037/a0021840.  [RM]

Csibra, G. & Gergely, G. (2011) Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society Series B: Biological Sciences* 366:1149–57.  [ASB]

Cutright, K. M., Wu, E. C., Banfield, J. C., Kay, A. C. & Fitzsimons, G. J. (2011) When your world must be defended: Choosing products to justify the system. *Journal of Consumer Research* 38:62–77.  [ELU]

Dar-Nimrod, I. & Heine, S. J. (2011) Genetic essentialism: On the deceptive determinism of DNA. *Psychological Bulletin* 137:800–18.  [aAC, CO]

Dawkins, R. (2010) *The greatest show on earth: The evidence for evolution.* Free Press.  [BH]

De Neys, W. (2006) Dual processing in reasoning: Two systems but one reasoner. *Psychological Science* 17:428–33.  [aAC]

Dennett, D. C. (1987) *The intentional stance.* MIT Press.  [KB]

Devine, P. G. (1989) Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology* 56:5–18.  [arAC]

Dewey, J. (1922) *Human nature and conduct: An introduction to social psychology.* Holt.  [VD]

Diesendruck, G., Goldfein-Elbaz, R., Rhodes, M., Gelman, S. A. & Neumark, N. (2013) Cross-cultural differences in children's beliefs about the objectivity of social categories. *Child Development* 84:1906–17.  [MR]

Diesendruck, G. & Haber, L. (2009) God's categories: The effect of religiosity on children's teleological and essentialist beliefs about categories. *Cognition* 110:100–14. doi:10.1016/j.cognition.2008.11.001.  [TCD, CO]

Diesendruck, G. & HaLevi, H. (2006) The role of language, appearance, and culture in children's social category-based induction. *Child Development* 77:539–53.  [MR]

Diesendruck, G. & Markson, L. (2001) Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology* 37:630–41.  [aAC]

Diesendruck, G. & Markson, L. (2011) Children's assumption of the conventionality of culture. *Child Development Perspectives* 5:189–95.  [aAC]

DiVitto, B. & McArthur, L. Z. (1978) Developmental differences in the use of distinctiveness, consensus, and consistency information for making causal attributions. *Developmental Psychology* 14:474–82.  [aAC]

Dunham, Y., Baron, A. S. & Carey, S. (2011) Consequences of "minimal" group affiliations in children. *Child Development* 82:793–811.  [MR]

Dunham, Y., Newheiser, A., Hoosain, L., Merrill, A. & Olson, K. R. (2014) From a different vantage: Intergroup attitudes among children from low- and inter-mediate-status racial groups. *Social Cognition* 32:1–21.  [A-KN]

Dunning, D., Leuenberger, A. & Sherman, D. A. (1995) A new look at motivated inference: Are self-serving theories of success a product of motivational forces? *Journal of Personality and Social Psychology* 69:58–68.  [ELU]

Dupré, J. (1981) Natural kinds and biological taxa. *Philosophical Review* 90:66–90.  [aAC]

Dweck, C. S. (1999) *Self-theories: Their role in motivation, personality, and development.* Psychology Press.  [aAC]

Dweck, C. S. (2006) *Mindset: The new psychology of success.* Random House.  [aAC]

Eagly, A. H. & Steffen, V. J. (1984) Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology* 46:735–54.  [DG]

Eco, U. (1999) *Kant and the platypus: Essays on language and cognition.* Secker & Warburg.  [NB]

Eidelman, S., Crandall, C. S. & Pattershall, J. (2009) The existence bias. *Journal of Personality and Social Psychology* 97:765–75.  [aAC]

Epley, N. & Gilovich, T. (2006) The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychologic8al Science* 17:311–18.  [rAC]

Evans, J. St. B. T. (2003) In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences* 7:454–59.  [aAC]

Evans, J. St. B. T. (2006) The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin and Review* 13:378–95.  [aAC]

Evans, J. St. B. T. (2008) Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology* 59:255–78.  [aAC]

Evans, J. St. B. T. (2009) How many dual process theories do we need: One, two or many? In: *In two minds: Dual processes and beyond,* ed. J. St. B. T. Evans & K. Frankish, pp. 1–32. Oxford University Press.  [aAC]

Evans, J. St. B. T. & Stanovich, K. E. (2013) Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science* 8:223–41.  [rAC]

Fasig, L. G. (2000) Toddlers' understanding of ownership: Implications for self-concept development. *Social Development* 9:370–82.  [NSN]

Fisher, P. (2002) Reinventing vitalism. *Homeopathy* 91:61–62.  [BB]

Fiske, A. P., Kitayama, S., Markus, H. R. & Nisbett, R. E. (1998) The cultural matrix of social psychology. In: *The handbook of social psychology, vol. 2,* 4th edition, ed. D. T. Gilbert, S. T. Fiske & G. Lindzey, pp. 915–81. McGraw-Hill.  [CO]

Fiske, S. T. & Taylor, S. E. (1984) *Social cognition.* Addison-Wesley.  [aAC]

Fiske, S. T. & Taylor, S. E. (1991) *Social cognition,* 2nd edition. McGraw-Hill.  [RM]

Fiske, S. T. & Taylor, S. E. (2008) *Social cognition: From brains to culture.* McGraw-Hill.  [VYY]

Fodor, J. A. (1983) *The modularity of mind: An essay on faculty psychology.* MIT Press.  [GW, rAC]

Frazer, J. G. (1890/1959) *The golden bough: A study in magic and religion,* abridged edition, ed. T. H. Gaster. Macmillan. (Original work published in 1890.)  [BB]

Frazer, J. G. (1890/1958) *The golden bough: A study in magic and religion.* Macmillan. (Original work published in 1890.)  [aAC]

Frazier, B. N. & Gelman, S. A. (2009) Developmental changes in judgments of authentic objects. *Cognitive Development* 24:284–92.  [NSN]

Frazier, B. N., Gelman, S. A. & Wellman, H. M. (2009) Preschoolers' search for explanatory information within adult–child conversation. *Child Development* 80:1592–611.  [SAG]

Frederick, S. (2002) Automated choice heuristics. In: *Heuristics and biases: The psychology of intuitive judgment,* ed. T. Gilovich, D. Griffin & D. Kahneman, pp. 548–58. Cambridge University Press.  [aAC]

Friedman, O., Van de Vondervoort, J. W., Defeyter, M. A. & Neary, K. R. (2013) First possession, history, and young children's ownership judgments. *Child Development* 84:1519–25.  [SAG]

Friedrich, J. (2005) Naturalistic fallacy errors in lay interpretations of psychological science: Data and reflections on the Rind, Tromovitch, and Bauserman (1998) controversy. *Basic and Applied Social Psychology* 27:59–70.  [aAC]

Friedrich, J. (2010) Naturalistic fallacy errors and behavioral science news: The effects of editorial content and cautions on readers' moral inferences and perceptions of contributors. *Basic and Applied Social Psychology* 32:369–83.  [aAC]

Friedrich, J., Kierniesky, N. & Cardon, L. (1989) Drawing moral inferences from descriptive science: The impact of attitudes on naturalistic fallacy errors. *Personality and Social Psychology Bulletin* 15:414–25.  [aAC]

Gabennesch, H. (1990) The perception of social conventionality by children and adults. *Child Development* 61:2047–59.  [arAC]

Garcia, J. & Koelling, R. A. (1966) Relation of cue to consequence in avoidance learning. *Psychonomic Science* 4:123–24.  [AM]

Gaucher, D., Jost, J. T. & Laurin, K. (2013) [System threat and essentialism.] Unpublished raw data.  [DG]

Gawronski, B. (2004) Theory-based bias correction in dispositional inference: The fundamental attribution error is dead, long live the correspondence bias. *European Review of Social Psychology* 15:183–217.  [aAC]

Gawronski, B., Bodenhausen, G. & Becker, A. (2007) I like it, because I like myself: Associative self-anchoring and post-decisional change of implicit evaluations. *Journal of Experimental Social Psychology* 43:221–32.  [NSN]

Gelman, R. (1990) First principles organize attention to and learning about relevant data: Number and the animate–inanimate distinction as examples. *Cognitive Science* 14:79–106.  [arAC]

Gelman, S. A. (2003) *The essential child: Origins of essentialism in everyday thought.* Oxford University Press.  [aAC, SAG]

Gelman, S. A. (2013) Artifacts and essentialism. *Review of Philosophy and Psychology* 4:449–63.  [rAC]

Gelman, S. A. & Bloom, P. (2000) Young children are sensitive to how an object was created when deciding what to name it. *Cognition* 76:91–103.  [aAC]

Gelman, S. A., Coley, J. D., Rosengren, K. S., Hartman, E. & Pappas, A. (1998) Beyond labeling: The role of parental input in the acquisition of richly struc-tured categories. *Monographs of the Society for Research in Child Development* 63(1, whole no. 253).  [arAC]

Gelman, S. A., Collman, P. & Maccoby, E. (1986) Inferring properties from cate-gories versus inferring categories from properties: The case of gender. *Child Development* 57:396–404.  [MR]

Gelman, S. A. & Diesendruck, G. (1999) A reconsideration of concepts: On the compatibility of psychological essentialism and context sensitivity. In: *Concep-tual development: Piaget's legacy,* ed. E. K. Scholnick, K. Nelson, S. A. Gelman & P. H. Miller, pp. 79–102. Erlbaum.  [TCD]

Gelman, S. A., Frazier, B. N., Noles, N. S., Manczak, E. & Stillwell, S. M. (2014) How much would children pay for Harry Potter's glasses? Developing an ap-preciation for the value of authentic objects. *Journal of Cognition and Devel-opment.* Available at: http://www.tandfonline.com/doi/full/10.1080/15248372.2013.815623#.U-kDiYBdVrw.  [NSN]

Gelman, S. A. & Gottfried, G. M. (1996) Children's causal explanations of animate and inanimate motion. *Child Development* 67:1970–87.  [aAC]

Gelman, S. A. & Kalish, C. W. (1993) Categories and causality. In: *Emerging themes in cognitive development, vol. 2,* ed. R. Pasnak & M. L. Howe, pp. 3–32. Springer.  [aAC]

Gelman, S. A. & Kremer, K. (1991) Understanding natural cause: Children's ex-planations of how objects and their properties originate. *Child Development* 62:396–414.  [rAC]

Gelman, S. A., Manczak, E. M. & Noles, N. S. (2012) The nonobvious basis of ownership: Preschool children trace the history and value of owned Objects. *Child Development* 83:1732–47.  [SAG, NSN]

Gelman, S. A. & Markman, E. M. (1986) Categories and induction in young children. *Cognition* 23:183–209. [aAC]

Gelman, S. A., Meyer, M. A. & Noles, N. S. (2013) History and essence in human cognition. *Behavioral and Brain Sciences* 36:142–43. [NSN]

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 for Research in Child Development* 69(1, Whole No. 275). [aAC]

Gelman, S. A., Ware, E. A. & Kleinberg, F. (2010) Effects of generic language on category content and structure. *Cognitive Psychology* 61:273–301. [arAC, SAG]

Gelman, S. A. & Waxman, S. R. (2009) Response to Sloutsky: Taking development seriously: Theories cannot emerge from associations alone. *Trends in Cognitive Sciences* 13:332–33. [SAG]

Gelman, S. A. & Wellman, H. M. (1991) Insides and essences: Early understandings of the non-obvious. *Cognition* 38:213–44. [aAC]

Gigerenzer, G. (1996) On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). *Psychological Review* 103:592–96. [rAC]

Gigerenzer, G. (2008) Moral intuition=fast and frugal heuristics? In: *Moral psychology, vol. 2: The cognitive science of morality – Intuition and diversity*, ed. W. Sinnott-Armstrong, pp. 1–26. MIT Press. [VD]

Gigerenzer, G., Todd, P. M. & the ABC Research Group (1999) *Simple heuristics that make us smart*. Oxford University Press. [NB]

Gilbert, D. T. (2002) Inferential correction. In: *Heuristics and biases: The psychology of intuitive judgment*, ed. T. Gilovich, D. Griffin & D. Kahneman, pp. 1–11. Cambridge University Press. [rAC]

Gilbert, D. T. & Malone, P. S. (1995) The correspondence bias. *Psychological Bulletin* 117:21–38. [arAC]

Gilbert, D. T., Pelham, B. W. & Krull, D. S. (1988) On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology* 54:733–40. [arAC, YD]

Gilbert, D. T., Pelham, B. W. & Krull, D. S. (2003) The psychology of good ideas. *Psychological Inquiry* 14:258–60. [rAC]

Giles, J., Legare, C. H. & Samson, J. (2008) Psychological essentialism and cultural variation: Children's beliefs about aggression in the United States and South Africa. *Infant and Child Development* 17:137–50. [CYO]

Gilovich, T. & Griffin, D. (2002) Introduction – Heuristics and biases: Then and now. In: *Heuristics and biases: The psychology of intuitive judgment*, ed. T. Gilovich, D. Griffin & D. Kahneman, pp. 1–18. Cambridge University Press. [aAC]

Gilovich, T., Griffin, D. & Kahneman, D. (2002) *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press. [aAC]

Gil-White, F. (2001) Are ethnic groups biological "species" to the human brain? *Current Anthropology* 42:515–54. [aAC]

Goodman, N. (1983) *Fact, fiction, and forecast*, 4th edition. Harvard University Press. (Original work published in 1955.) [SP]

Gopnik, A. (1998) Explanation as orgasm. *Minds and Machines* 8:101–18. [aAC]

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T. & Danks, D. (2004) A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review* 111:3–32. [arAC]

Gopnik, A. & Schulz, L. (2004) Mechanisms of theory formation in young children. *Trends in Cognitive Sciences* 8:371–77. [SAG]

Gopnik, A. & Wellman, H. M. (1994) The theory theory. In: *Mapping the mind: Domain specificity in culture and cognition*, ed. L. A. Hirschfeld & S. A. Gelman, pp. 257–93. Cambridge University Press. [KB]

Gottfried, G. M. & Gelman, S. A. (2005) Developing domain-specific causal-explanatory frameworks: The role of insides and immanence. *Cognitive Development* 20:137–58. [arAC]

Graham, S. A., Stock, H. & Henderson, A. M. (2006) Nineteen-month-olds' understanding of the conventionality of object labels versus desires. *Infancy* 9:341–50. [aAC]

Greenwald, A. G., McGhee, D. E. & Schwartz, J. L. K. (1998) Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology* 74:1464–80. [aAC]

Gweon, H. & Schulz, L. E. (2011) 16-month-olds rationally infer causes of failed actions. *Science* 332:1524. [aAC]

Hall, D. G. (1998) Continuity and persistence of objects. *Cognitive Psychology* 37:28–59. [BH]

Hamlin, J. K., Wynn, K. & Bloom, P. (2007) Social evaluation by preverbal infants. *Nature* 450:557–59. [rAC]

Haslam, N. & Ernst, D. (2002) Essentialist beliefs about mental disorders. *Journal of Social and Clinical Psychology* 21:628–44. doi:10.1521/jscp.21.6.628.22793. [TCD]

Haslam, N. & Levy, S. R. (2006) Essentialist beliefs about homosexuality: Structure and implications for prejudice. *Personality and Social Psychology Bulletin* 32:471–85. doi:10.1177/0146167205276516. [TCD, NH]

Haslam, N., Rothschild, L. & Ernst, D. (2000) Essentialist beliefs about social categories. *British Journal of Social Psychology* 39:113–27. doi:10.1348/014466600164363. [arAC, TCD, NH, CO, VYY]

Haslam, N., Rothschild, L. & Ernst, D. (2002) Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology* 41:87–100. doi:10.1348/014466602165072. [aAC, NH, RM]

Haslam, N. & Whelan, J. (2008) Human natures: Psychological essentialism in thinking about differences between people. *Social and Personality Psychology Compass* 2:1297–312. [aAC]

Hatano, G. & Inagaki, K. (1994) Young children's naive theory of biology. *Cognition* 50:171–88. [arAC]

Hay, D. F. (2006) Yours and mine: Toddlers' talk about possessions with familiar peers. *British Journal of Developmental Psychology* 24:39–52. [NSN]

Heider, F. (1958) *The psychology of interpersonal relations*. Wiley. [rAC]

Helwig, C. C., Tisak, M. S. & Turiel, E. (1990) Children's social reasoning in context: Reply to Gabennesch. *Child Development* 61:2068–78. [aAC]

Helwig, C. C. & Turiel, E. (2011) Children's social and moral reasoning. In: *Wiley-Blackwell handbook of childhood social development*, 2nd edition, ed. C. Hart & P. Smith, pp. 567–83. Wiley-Blackwell. [aAC]

Henrich, J. & Broesch, J. (2011) On the nature of cultural transmission networks: Evidence from Fijian villages for adaptive learning biases. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366(1567):1139–48. doi:10.1098/rstb.2010.0323. [ASB]

Henrich, J., Heine, S. & Norenzayan, A. (2010) The weirdest people in the world? *Behavioral and Brain Sciences* 33:61–135. doi:10.1017/S0140525X0999152X. [ASB]

Henrich, J. & McElreath, R. (2003) The evolution of cultural evolution. *Evolutionary Anthropology* 12(3):123–35. doi:10.1002/evan.10110. [ASB]

Herrmann, E., Call, J., Lloreda, M., Hare, B. & Tomasello, M. (2007) Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science* 317:1360–66. [ASB]

Heussen, D. & Hampton, J. A. (2008) Ways of explaining properties. In: *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, ed. B. C. Love, K. McRae & V. M. Sloutsky, pp. 143–48. Cognitive Science Society. [JAH]

Hewstone, M. (1990) The "ultimate attribution error"? A review of the literature on intergroup causal attribution. *European Journal of Social Psychology* 20:311–35. [VYY]

Higgins, E. T. (1996) Knowledge activation: Accessibility, applicability, and salience. In: *Social psychology: Handbook of basic principles*, ed. E. T. Higgins & A. W. Kruglanski, pp. 133–68. Guilford Press. [aAC]

Higgins, E. T. & Bryant, S. L. (1982) Consensus information and fundamental attribution error: The role of development and in-group versus out-group knowledge. *Journal of Personality and Social Psychology* 43:889–900. [aAC]

Higgins, E. T. & King, G. (1981) Accessibility of social constructs: Information processing consequences of individual and contextual variability. In: *Personality, cognition, and social interaction*, ed. N. Cantor & J. F. Kihlstrom, pp. 69–121. Erlbaum. [aAC]

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

Higgins, E. T., Rholes, W. S. & Jones, C. R. (1977) Category accessibility and impression formation. *Journal of Experimental Social Psychology* 13:141–54. [aAC]

Hodson, G. & Busseri, M. A. (2012) Bright minds and dark attitudes: Lower cognitive ability predicts greater prejudice through right-wing ideology and low intergroup contact. *Psychological Science* 23:187–95. [NH]

Hoffman, C. & Hurst, N. (1990) Gender stereotypes: Perception or rationalization? *Journal of Personality and Social Psychology* 58:197–208. [aAC, DG]

Hood, B. (2009) *SuperSense: Why we believe the unbelievable*. HarperOne. [BH]

Hood, B. M. & Bloom, P. (2008) Children prefer certain individuals over perfect duplicates. *Cognition* 106:455–62. [arAC, BH]

Hooper, E. M. (1890) Hints on home dress-making. *The Ladies' Home Journal* 7(12):23. [aAC]

Horton, R. (1967) African traditional thought and Western science: Part II. The "closed" and "open" predicaments. *Africa: Journal of the International African Institute* 37:155–87. [aAC]

Hume, D. (1739/1975) *Treatise of human nature*. Clarendon Press. (Original work published in 1739.) [VD]

Hume, D. (1740/2000) *A treatise of human nature*, ed. D. F. Norton & M. J. Norton. Oxford University Press. (Original work published 1740.) [arAC]

Hurlbert, A. & Ling, Y. (2007) Biological components of sex differences in color preference. *Current Biology* 17:R623–25. [aAC]

Hussak, L. & Cimpian, A. (2013) Why do people think they live in a fair society? A new perspective on the cognitive origins of system justification. Paper presented at the Annual Convention of the Cognitive Development Society, Memphis, TN, October 18–19, 2013. [rAC]

Inagaki, K. & Hatano, G. (2004) Vitalistic causality in young children's naive biology. *Trends in Cognitive Science* 8:356–62. [BB]

Izard, V., Sann, C., Spelke, E. S. & Streri, A. (2009) Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences USA* 106:10382–85. [rAC]

Jaeggi, S. M., Buschkuehl, M., Jonides, J. & Shah, P. (2011) Short and long term benefits of cognitive training. *Proceedings of the National Academy of Sciences USA* 108:10081–86. doi:10.1073/pnas.1103228108. [RM]

Jayaratne, T. E., Gelman, S., Feldbaum, M., Sheldon, J., Petty, E. & Kardia, S. (2009) The perennial debate: Nature, nurture, or choice? Black and White Americans' explanations for individual differences. *Review of General Psychology* 13:24–33. [CO]

Joffe, H. & Staerklé, C. (2007) The centrality of the self-control ethos in Western aspersions regarding outgroups: A social representational approach to stereotype content. *Culture and Psychology* 13:395–418. [CO]

Jones, E. E. (1979) The rocky road from acts to dispositions. *American Psychologist* 34:107–17. [rAC]

Jones, E. E. & Davis, K. E. (1965) From acts to dispositions: The attribution process in social psychology. In: *Advances in experimental social psychology, vol. 2*, ed. L. Berkowitz, pp. 219–66. Academic Press. [aAC]

Jones, E. E. & Harris, V. A. (1967) The attribution of attitudes. *Journal of Experimental Social Psychology* 3:1–24. [arAC]

Jost, J. T. (2001) Outgroup favoritism and the theory of system justification: A paradigm for investigating the effects of socioeconomic success on stereotype content. In: *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition*, ed. G. Moskowitz, pp. 89–102. Psychology Press. [aAC]

Jost, J. T. & Banaji, M. R. (1994) The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology* 33:1–27. [aAC, DG, A-KN, ELU]

Jost, J. T. Banaji, M. R. & Nosek, B. A. (2004) A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology* 25:881–919. [aAC, ELU]

Jost, J. T. & Burgess, D. (2000) Attitudinal ambivalence and the conflict between group and system justification motives in low status groups. *Personality and Social Psychology Bulletin* 26:293–305. [aAC]

Jost, J. T., Gaucher, D., & Stern, C. (2015). "The world isn't fair": A system justification perspective on social stratification and inequality. In: *APA Handbook of Personality and Social Psychology. Volume 2*. ed. J. F. Dovidio & J. Simpson, pp. 317–40. American Psychological Association. [DG]

Jost, J. T., Glaser, J., Kruglanski, A. W. & Sulloway, F. (2003) Political conservatism as motivated social cognition. *Psychological Bulletin* 129:339–75. [aAC]

Jost, J. T. & Hamilton, D. L. (2005) Stereotypes in our culture. In: *On the nature of prejudice: Fifty years after Allport*, ed. J. F. Dovidio, P. Glick & L. Rudman, pp. 208–24. Blackwell. [DG]

Jost, J. T. & Hunyady, O. (2005) Antecedents and consequences of system-justifying ideologies. *Current Directions in Psychological Science* 14:260–65. [aAC]

Jost, J. T., Kivetz, Y., Rubini, M., Guermandi, G. & Mosso, C. (2005) System-justifying functions of complementary regional and ethnic stereotypes: Cross-national evidence. *Social Justice Research* 18:305–33. [ELU]

Kahneman, D. (2002) Maps of bounded rationality: A perspective on intuitive judgment and choice. *Nobel Prize Lecture* (accessed March 2011). Available at: http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2002/kahnemann-lecture.pdf. [GW]

Kahneman, D. (2011) *Thinking, fast and slow*. Farrar, Straus, and Giroux. [arAC, KB, FLB]

Kahneman, D. & Frederick, S. (2002) Representativeness revisited: Attribute substitution in intuitive judgment. In: *Heuristics and biases: The psychology of intuitive judgment*, ed. T. Gilovich, D. Griffin & D. Kahneman, pp. 49–81. Cambridge University Press. [arAC, NB, VD]

Kahneman, D., Knetsch, J. L. & Thaler, R. H. (1990) Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy* 98:1325–48. [NSN]

Kahneman, D. & Tversky, A. (1973) On the psychology of prediction. *Psychological Review* 80:237–51. [arAC]

Kahneman, D. & Tversky, A. (1996) On the reality of cognitive illusions. *Psychological Review* 103:582–91. [rAC]

Kail, R. (1991) Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin* 109:490–501. [aAC]

Kalish, C. W. (1998) Reasons and causes: Children's understanding of conformity to social rules and physical laws. *Child Development* 69:706–20. [arAC]

Kalish, C. W. & Lawson, C. A. (2008) Development of social category representations: Early appreciation of roles and deontic relations. *Child Development* 79:577–93. [aAC]

Kamins, M. L. & Dweck, C. S. (1999) Person versus process praise and criticism: Implications for contingent self-worth and coping. *Developmental Psychology* 35:835–47. [aAC]

Kane, M. J. & Engle, R. W. (2002) The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin and Review* 9:637–71. [aAC]

Kant, I. (1785/1993) *Grounding for the metaphysics of morals*, 3rd edition, trans. J. W. Ellington. Hackett. (Original work published in 1785.) [VD]

Kay, A. C. & Friesen, J. (2011) On social stability and social change: Understanding when system justification does and does not occur. *Current Directions in Psychological Science* 20:360–64. [aAC, DG]

Kay, A. C., Gaucher, D., Peach, J. M., Laurin, K., Friesen, J., Zanna, M. P. & Spencer, S. J. (2009) Inequality, discrimination, and the power of the status quo: Direct evidence for a motivation to see the way things are as the way they should be. *Journal of Personality and Social Psychology* 97:421–34. [aAC, A-KN]

Kay, A. C., Jost, J. T. & Young, S. (2005) Victim derogation and victim enhancement as alternate routes to system justification. *Psychological Science* 16:240–46. [ELU]

Keil, F. C. (1989) *Concepts, kinds, and cognitive development*. MIT Press. [HR, MS]

Keil, F. C. (2006) Explanation and understanding. *Annual Review of Psychology* 57:227–54. [rAC]

Keller, J. (2005) In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology* 88:686–702. doi:10.1037/0022-3514.88.4.686. [aAC, TCD, DG, RM, CO]

Kelley, H. H. (1967) Attribution theory in social psychology. *Nebraska Symposium on Motivation* 15:192–238. [aAC]

Kelley, H. H. (1973) The processes of causal attribution. *American Psychologist* 28:107–28. [aAC]

Kelley, H. H. & Michela, J. L. (1980) Attribution theory and research. *Annual Review of Psychology* 31:457–501. [aAC]

Keysar, B. & Bly, B. M. (1995) Intuitions of the transparency of idioms: Can one keep a secret by spilling the beans? *Journal of Memory and Language* 34:89–109. [aAC]

Keysar, B. & Bly, B. M. (1999) Swimming against the current: Do idioms reflect conceptual structure? *Journal of Pragmatics* 31:1559–78. [aAC]

Kinzler, K., Shutts, K. & Correll, J. (2010) Priorities in social categories. *European Journal of Social Psychology* 40:581–92. [MR]

Kokis, J. V., Macpherson, R., Toplak, M. E., West, R. F. & Stanovich, K. E. (2002) Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology* 83:26–52. [arAC]

Kornblith, H. (1995) *Inductive inference and its natural ground: An essay in naturalistic epistemology*. MIT Press. [aAC]

Kovács, A. M. & Mehler, J. (2009) Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences USA* 106:6556–60. [rAC]

Kovács, A. M., Teglas, E. & Endress, A. D. (2010) The social sense: Susceptibility to others' beliefs in human infants and adults. *Science* 330:1830–34. [rAC]

Kruglanski, A. W. (1996) Motivated social cognition: Principles of the interface. In: *Social psychology: Handbook of basic principles*, ed. E. T. Higgins & A. W. Kruglanski, pp. 493–520. Guilford Press. [aAC]

Kruglanski, A. W., Webster, D. W. & Klem, A. (1993) Motivated resistance and openness to persuasion in the presence or absence of prior information. *Journal of Personality and Social Psychology* 65:861–76. [aAC]

Krull, D. S. (1993) Does the grist change the mill? The effect of the perceiver's inferential goal on the process of social inference. *Personality and Social Psychology Bulletin* 19:340–48. [aAC]

Krull, D. S. & Dill, J. C. (1996) On thinking first and responding fast: Flexibility in social inference processes. *Personality and Social Psychology Bulletin* 22 (9):949–59. [YD]

Krull, D. S. & Erickson, D. J. (1995) Judging situations: On the effortful process of taking dispositional information Into account. *Social Cognition* 13(4):417–38. doi:10.1521/soco.1995.13.4.417. [YD]

Krull, D. S., Loy, M. H. M., Lin, J., Wang, C. F., Chen, S. & Zhao, X. (1999) The fundamental fundamental attribution error: Correspondence bias in individualist and collectivist cultures. *Personality and Social Psychology Bulletin* 25:1208–19. [aAC]

Kuhlmeier, V., Wynn, K. & Bloom, P. (2003) Attribution of dispositional states by 12-month-olds. *Psychological Science* 14:402–408. [rAC]

Kunda, Z. (1990) The case for motivated reasoning. *Psychological Bulletin* 108:480–98. [ELU]

Langton, R. & Lewis, D. (1998) Defining "intrinsic." *Philosophy and Phenomenological Research* 58(2):333–45. [MAK]

Laszlo, P. (2007) *Citrus: A history*. University of Chicago Press. [aAC]

Lau, G. P., Kay, A. C. & Spencer, S. J. (2008) Loving those who justify inequality: The effects of system threat on attraction to women who embody benevolent sexist ideals. *Psychological Science* 19:20–21. [ELU]

Laurin, K., Shepard, S. & Kay, A. C. (2010) Restricted emigration, system inescapability, and the defense of the status quo: System-justifying consequences of restricted exit opportunities. *Psychological Science* 21:1075–82. [DG, ELU]

Le Pelley, M. E., Reimers, S. J., Calvini, G., Spears, R., Beesley, T. & Murphy, R. A. (2010) Stereotype formation: Biased by association. *Journal of Experimental Psychology: General* 139:138–61. [AM]

Legare, C. H., Gelman, S. A. & Wellman, H. M. (2010) Inconsistency with prior knowledge triggers children's causal explanatory reasoning. *Child Development* 81:929–44. [rAC]

Lerner, M. J. (1980) *The belief in a just world: A fundamental delusion.* Plenum [aAC, DG]

Leslie, S. J. (2013) Essence and natural kinds: When science meets preschooler intuition. In: *Oxford studies in epistemology, vol. 4*, ed. T. Gendler & J. Hawthorne, pp. 108–66. Oxford University Press.   [aAC]

Levine, L. E. (1983) Mine: Self-definition in 2-year-old boys. *Developmental Psychology* 19:544–49.   [NSN]

Levy, S. R., Stroessner, S. J. & Dweck, C. S. (1998) Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology* 74:1421–36. doi:10.1037/0022-3514.74.6.1421.   [aAC, RM]

Lewis, D. (1983) Extrinsic properties. *Philosophical Studies* 44:197–200.   [MAK, rAC]

Leyens, J.-P., Yzerbyt, V. Y. & Corneille, O. (1996) The role of concept applicability in the emergence of the overattribution bias. *Journal of Personality and Social Psychology* 70:229–39.   [VYY]

Lillard, A. S. & Flavell, J. H. (1990) Young children's preference for mental state versus behavioral descriptions of human action. *Child Development* 61:731–41.   [KB]

Lindeman, M. & Saher, M. (2007) Vitalism, purpose and superstition. *British Journal of Psychology* 98:33–44.   [BB]

Lipton, P. (2004) *Inference to the best explanation*, 2nd edition. Routledge.   [aAC]

Liviatan, I. & Jost, J. T. (2011) Motivated social cognition in the service of the status quo. Special issue: System justification theory. *Social Cognition* 29:231–37. [aAC]

LoBue, V. & DeLoache, J. S. (2011) Pretty in pink: The early development of gender-stereotyped colour preferences. *British Journal of Developmental Psychology* 29:656–67.   [aAC]

Lockhart, K. L., Abrahams, B. & Osherson, D. N. (1977) Children's understanding of uniformity in the environment. *Child Development* 48:1521–31.   [arAC]

Lohman, D. F. (1989) Human intelligence: An introduction to advances in theory and research. *Review of Educational Research* 59:333–73.   [aAC]

Lombrozo, T. (2006) The structure and function of explanations. *Trends in Cognitive Sciences* 10:464–70.   [rAC]

Lombrozo, T. (2012) Explanation and abductive inference. In: *Oxford handbook of thinking and reasoning*, ed. K. J. Holyoak & R. G. Morrison, pp. 260–76. Oxford University Press.   [rAC]

Luo, Y., Kaufman, L. & Baillargeon, R. (2009) Young infants' reasoning about physical events involving inert and self-propelled objects. *Cognitive Psychology* 58:441–86.   [aAC]

Machery, E., Mallon, R., Nichols, S. & Stich, S. P. (2004) Semantics, cross-cultural style. *Cognition* 92:B1–B12.   [CYO]

Machery, E., Olivola, C. Y., Cheon, H., Kurniawan, I. T., Mauro, C., Struchiner, N. & Susianto, H. (submitted) Is folk essentialism a fundamental feature of human cognition?   [CYO]

Machery, E., Olivola, C. Y. & de Blanc, M. (2009) Linguistic and metalinguistic intuitions in the philosophy of language. *Analysis* 69:689–94.   [CYO]

Mackie, J. L. (1977) *Ethics: Inventing right and wrong*. Penguin Books.   [GW]

Mackintosh, N. J. (1975) A theory of attention: Variations in associability of stimuli with reinforcement. *Psychological Review* 82(4):276–98.   [AM]

Maglo, K. N. (2011) The case against biological realism about race: From Darwin to the post-genomic era. *Perspectives on Science* 19:361–90.   [aAC]

Mahalingam, R. (2003a) Essentialism, culture, and beliefs about gender among the Aravanis of Tamil Nadu, India. *Sex Roles* 49:489–96.   [CYO]

Mahalingam, R. (2003b) Essentialism, culture, and power representations of social class. *Journal of Social Issues* 59:733–49.   [DG, CYO]

Mahalingam, R. (2007) Essentialism, power, and the representation of social categories: A folk sociology perspective. *Human Development* 50:300–19.   [CO]

Mahalingam, R. & Rodriguez, J. (2003) Essentialism, power and cultural psychology of gender. *Journal of Cognition and Culture* 3:157–74.   [CO]

Mahalingam, R. & Rodriguez, J. (2006) Culture, brain transplants and implicit theories of identity. *Journal of Cognition and Culture* 6:453–62.   [CO]

Malinvaud, E. (1966) *Statistical methods of econometrics*. North-Holland.   [FLB]

Markman, E. M. (1989) *Categorization and naming in children: Problems of induction*. MIT Press.   [aAC]

Martin, K. A. (2009) Normalizing heterosexuality: Mother's assumptions, talk, and strategies with young children. *American Sociological Review* 74:190–207. [RSB]

Matthews, M., Levin, S. & Sidanius, J. (2009) A longitudinal test of the model of political conservatism as motivated social cognition. *Political Psychology* 30:921–36.   [aAC]

Maurer, D., Pathman, T. & Mondloch, C. J. (2006) The shape of boubas: Sound–shape correspondences in toddlers and adults. *Developmental Science* 9:316–22.   [aAC]

Mauss, M. (1902/1972) *A general theory of magic*, trans. R. Brain. W. W. Norton. (Original work published in 1902.)   [BB]

Mayr, E. (1963) *Animal species and evolution*. Belknap Press.   [BH]

McArthur, L. A. (1972) The how and what of why: Some determinants and consequences of causal attribution. *Journal of Personality and Social Psychology* 22:171–93.   [aAC]

McRae, K., Cree, G. S., Seidenberg, M. S. & McNorgan, C. (2005) Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods* 37:547–59.   [rAC]

McRae, K., De Sa, V. R. & Seidenberg, M. S. (1997) On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General* 126:99–130.   [arAC]

McRae, K. & Jones, M. N. (2013) Semantic memory. In: *The Oxford handbook of cognitive psychology*, ed. D. Reisberg. Oxford University Press.   [aAC]

Medin, D. L. & Ortony, A. (1989) Psychological essentialism. In: *Similarity and analogical reasoning*, ed. S. Vosniadou & A. Ortony, pp. 179–95. Cambridge University Press.   [aAC, TCD, BH]

Mendes, N., Rakoczy, H. & Call, J. (2008) Ape metaphysics: Object individuation without language. *Cognition* 106(2):730–49.   [HR]

Mendes, N., Rakoczy, H. & Call, J. (2011) Primates do not spontaneously use shape properties for object individuation: A competence or a performance problem? *Animal Cognition* 14:407–14.   [HR]

Meyer, M., Leslie, S.-J., Gelman, S. A. & Stilwell, S. M. (2013) Essentialist beliefs about bodily transplants in the United States and India. *Cognitive Science* 37:668–710.   [aAC]

Miyamoto, Y. & Kitayama, S. (2002) Cultural variation in correspondence bias: The critical role of attitude diagnosticity of socially constrained behavior. *Journal of Personality and Social Psychology* 83:1239–48.   [ASB]

Molden, D. C., Plaks, J. E. & Dweck, C. S. (2006) "Meaningful" social inferences: Effects of implicit theories on inferential processes. *Journal of Experimental Social Psychology* 42(6):738–52. doi:10.1016/j.jesp.2005.11.005.   [YD]

Monin, B. (2003) The warm glow heuristic: When liking leads to familiarity. *Journal of Personality and Social Psychology* 85:1035–48.   [rAC]

Moore, G. E. (1903/1971) *Principia ethica*. Cambridge University Press. (Original work published in 1903.)   [VD]

Morris, M. & Peng, K. (1994) Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology* 67:949–71.   [aAC, ASB]

Morris, S. C., Taplin, J. E. & Gelman, S. A. (2000) Vitalism in naive biological thinking. *Developmental Psychology* 36:582–95.   [arAC, BB]

Morton, J. & Johnson, M. H. (1991) CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review* 98:164–81.   [rAC]

Morton, T. A., Postmes, T., Haslam, S. A. & Hornsey, M. J. (2009) Is there anything essential about essentialism? *Journal of Personality and Social Psychology* 96:653–64.   [NH]

Moya, C., Boyd, R. & Henrich, J. (under review) Reasoning about cultural and genetic transmission: Developmental and cross-cultural evidence from Peru, Fiji, and the U.S. on how people make inferences about trait and identity transmission.   [ASB]

Mueller, C. M. & Dweck, C. S. (1998) Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology* 75:33–52.   [aAC]

Muraven, M., Baumeister, R. & Tice, D. M. (1999) Longitudinal improvement of self-regulation through practice: Building self-control strength through repeated exercise. *Journal of Social Psychology* 139:446–57. doi:10.1016/j.jesp.2009.12.011.   [RM]

Murdock, G. P., Wilson, S. F. & Frederick, V. (1978) World distribution of theories of illness. *Ethnology* 17:449–70.   [aAC]

Murphy, G. L. (2004) *The big book of concepts*. MIT Press.   [aAC]

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

Murphy, R. A., Mondragón, E. & Murphy, V. A. (2009) Contingency, temporal order and generalization: Building blocks of causal cognition. *International Journal of Comparative Psychology* 22:61–74.   [AM]

Napier, J. L., Mandisodza, A. N., Andersen, S. M. & Jost, J. T. (2006) System justification in responding to the poor and displaced in the aftermath of Hurricane Katrina. *Analyses of Social Issues and Public Policy* 6:57–73.   [aAC, DG]

Needham, P. (2000) What is water? *Analysis* 60:13–21.   [aAC]

Needham, P. (2011) Microessentialism: What is the argument? *Noûs* 45:1–21. [aAC]

Nemeroff, C. & Rozin, P. (1994) The contagion concept in adult thinking in the United States: Transmission of germs and of interpersonal influence. *Ethos* 22:158–86.   [BB]

Newheiser, A., Dunham, Y., Merrill, A., Hoosain, L. & Olson, K. R. (2014) Preference for high status predicts implicit outgroup bias among children from low-status groups. *Developmental Psychology*. 50(4):1081–90.   [A-KN]

Newheiser, A. & Olson, K. R. (2012) White and Black American children's implicit intergroup bias. *Journal of Experimental Social Psychology* 48:264–70. [A-KN]

Newman, G. E., Diesendruck, G. & Bloom, P. (2011) Celebrity contagion and the value of objects. *Journal of Consumer Research* 38:215–28.   [NSN]

Newman, G. E., Herrmann, P., Wynn, K. & Keil, F. C. (2008) Biases towards internal features in infants' reasoning about objects. *Cognition* 107 (2):420–32.   [arAC, HR]

Newman, G. E. & Keil, F. C. (2008) Where is the essence? Developmental shifts in children's beliefs about internal features. *Child Development* 79(5):1344–56. [aAC]

Nickerson, R. S. (1998) Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* 2:175–220. [arAC]

No, S., Hong, Y., Liao, H., Lee, K., Wood, D. & Chao, M. (2008) Lay theory of race affects and moderates Asian Americans' responses toward American culture. *Journal of Personality and Social Psychology* 95:991–1004. [CO]

Norenzayan, A., Choi, I. & Nisbett, R. E. (2002) Cultural similarities and differences in social inference: Evidence from behavioral predictions and lay theories of behavior. *Personality and Social Psychology Bulletin* 28(1):109–20. [ASB]

Nosek, B. A., Banaji, M. R. & Greenwald, A. G. (2002) Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice* 6:101–15. [A-KN]

O'Connor, C., Rees, G. & Joffe, H. (2012) Neuroscience in the public sphere. *Neuron* 74:220–26. [CO]

O'Connor, P. T. & Kellerman, S. (2009) *Origins of the specious: Myths and misconceptions of the English language.* Random House. [JAH]

Oderberg, D. (2007) *Real essentialism.* Routledge. [AM]

Öhman, A. & Mineka, S. (2001) Fears, phobias, and preparedness: Toward an evolved module of fear learning. *Psychological Review* 108:483–522. [AM]

Olson, K. R., Shutts, K., Kinzler, K. D. & Weisman, K. G. (2012) Children associate racial groups with wealth: Evidence from South Africa. *Child Development* 83:1884–99. [A-KN]

Onishi, K. H. & Baillargeon, R. (2005) Do 15-month-old infants understand false beliefs? *Science* 308:255–58. [KB, rAC]

Orvis, B. R., Cunningham, J. D. & Kelley, H. H. (1975) A closer examination of causal inference: The roles of consensus, distinctiveness, and consistency information. *Journal of Personality and Social Psychology* 32:605–16. [aAC]

Osherson, D. N. & Markman, E. M. (1975) Language and the ability to evaluate contradictions and tautologies. *Cognition* 3:213–26. [aAC]

Paoletti, J. B. (2012) *Pink and blue: Telling the boys from the girls in America.* Indiana University Press. [aAC]

Peirce, C. S. (1934) A neglected argument for the existence of God [1908]. In: *Collected papers of Charles Sanders Peirce, vol. 6: Scientific metaphysics*, ed. C. Hartshorne & P. Weiss, pp. 311–39. Belknap Press of Harvard University Press. [FLB]

Pereira, M. E., Estramiana, J. L. A. & Gallo, I. S. (2010) Essentialism and the expression of social stereotypes: A comparative study of Spain, Brazil and England. *Spanish Journal of Psychology* 13:808–17. [CYO]

Pettigrew, T. F. & Tropp, L. R. (2000) Does intergroup contact reduce prejudice? Recent meta-analytic findings. In: *Reducing prejudice and discrimination: The Claremont Symposium on Applied Social Psychology*, ed. S. Oskamp, pp. 93–114. Erlbaum. [RSB]

Pfeffer, K., Cole, B. & Dada, K. (1998) Attributions for youth crime among British and Nigerian primary school children. *Journal of Social Psychology* 138:251–53. [CYO]

Phillips, W. & Santos, L. R. (2007) Evidence for kind representations in the absence of language: Experiments with rhesus monkeys (*Macaca mulatta*). *Cognition* 102:455–63. [HR]

Phillips, W., Shankar, M. & Santos, L. R. (2010) Essentialism in the absence of language? Evidence from rhesus monkeys (*Macaca mulatta*). *Developmental Science* 13(4):F1–F7. [HR]

Piaget, J. (1929) *The child's conception of the world.* Routledge & Kegan Paul. [KB]

Piaget, J. (1929/1967) *The child's conception of the world*, trans. J. Tomlinson & A. Tomlinson. Littlefield, Adams & Co. (Original work published 1929.) [arAC]

Pinker, S. (1994) *The language instinct: How the mind creates language.* William Morrow. [aAC, SAG]

Prasada, S. & Dillingham, E. (2006) Principled and statistical connections in common sense conception. *Cognition* 99:73–112. [arAC, SP]

Prasada, S. & Dillingham, E. (2009) Representation of principled connections: A window onto the formal aspect of common sense conception. *Cognitive Science* 33:401–48. [aAC, SP]

Prasada, S., Hennefield, L. & Otap, D. (2012) Conceptual and linguistic representations of kinds and classes. *Cognitive Science* 36:1224–50. [SP]

Prasada, S., Khemlani, S., Leslie, S.-J. & Glucksberg, S. (2013) Conceptual distinctions amongst generics. *Cognition* 126:405–22. [SP, rAC]

Premack, D. & Premack, A. J. (1996) Origins of human social competence. In: *The cognitive neurosciences*, ed. M. S. Gazzaniga, pp. 205–18. MIT Press. [aAC]

Prentice, D. A. & Miller, D. T. (2007) Psychological essentialism of human categories. *Current Directions in Psychological Science* 16:202–6. [aAC]

Pylyshyn, Z. (1989) The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition* 32:65–97. [rAC]

Pyszczynski, T. & Greenberg, J. (1987) Toward an integration of cognitive and motivational perspectives on social inference: A biased-hypothesis testing model. *Advances in Experimental Social Psychology* 20:297–340. [aAC]

Quine, W. V. O. (1960) *Word and object.* MIT Press. [JAH]

Racine, E. (2008) Which naturalism for bioethics? A defense of moderate (pragmatic) naturalism. *Bioethics* 22(2):92–100. [VD]

Railton, P. (1986) Moral realism. *Philosophical Review* 95:163–207. [rAC]

Ramachandran, V. S. & Hubbard, E. M. (2001) Synaesthesia – A window into perception, thought and language. *Journal of Consciousness Studies* 8:3–34. [aAC]

Rangel, U. & Keller, J. (2011) Essentialism goes social: Belief in social determinism as a component of psychological essentialism. *Journal of Personality and Social Psychology* 100:1056–78. [NH]

Raven, J. C. (1960) *Guide to the standard progressive matrices: Sets A, B, C, D and E.* H. K. Lewis & Co. [rAC]

Rehder, B. & Burnett, R. C. (2005) Feature inference and the causal structure of categories. *Cognitive Psychology* 50:264–314. doi:10.1016/j.cogpsych.2004.09.002. [TCD]

Repacholi, B. M. & Gopnik, A. (1997) Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology* 33:12–21. doi:10.1037/0012-1649.33.1.12. [KB]

Rhodes, M. (2012) Naïve theories of social groups. *Child Development* 83:1900–16. [MR]

Rhodes, M. (2014) Children's explanations as a window into their intuitive theories of the social world. *Cognitive Science*. doi: 10.1111/cogs.12129. [Epub ahead of print] [MR]

Rhodes, M. & Chalik, L. (2013) Social categories as markers of intrinsic interpersonal obligations. *Psychological Science* 6:999–1006. [MR]

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* 59:244–74. [aAC, CO, MR]

Rhodes, M. & Leslie, S. J. (in preparation) Essentialism underlies the development of prejudiced attitudes. [MR]

Rhodes, M., Leslie, S. J. & Tworek, C. M. (2012) Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences USA* 109:13526–31. [arAC, SAG, MR]

Richerson, P. & Boyd, R. (2005) *Not by genes alone: How culture transformed human evolution.* University of Chicago Press. [ASB]

Roets, A. & Van Hiel, A. (2011) The role of need for closure in essentialist entitativity beliefs and prejudice: An epistemic needs approach to racial categorization. *British Journal of Social Psychology* 50:52–73. [NH]

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

Rosenblum, T. & Pinker, S. A. (1983) Word magic revisited: Monolingual and bilingual children's understanding of the word-object relationship. *Child Development* 54:773–80. [aAC]

Ross, H., Tesla, C., Kenyon, B. & Lollis, S. (1990) Maternal intervention in toddler peer conflict: The socialization of principles of justice. *Developmental Psychology* 26:994–1003. [NSN]

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

Ross, L. D., Amabile, T. M. & Steinmetz, J. L. (1977) Social roles, social control, and biases in social-perception processes. *Journal of Personality and Social Psychology* 35:485–94. [arAC]

Rozenblit, L. R. & Keil, F. C. (2002) The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science* 26:521–62. [JAH]

Rozin, P., Markwith, M. & Ross, B. (1990) The sympathetic magical law of similarity, nominal realism and neglect of negatives in response to negative labels. *Psychological Science* 1:383–84. [aAC]

Rozin, P., Millman, L. & Nemeroff, C. (1986) Operation of the laws of sympathetic magic in disgust and other domains. *Journal of Personality and Social Psychology* 50:703–12. [aAC]

Rozin, P. & Nemeroff, C. (1990) The laws of sympathetic magic: A psychological analysis of similarity and contagion. In: *Cultural psychology: Essays on comparative human development*, ed. J. W. Stigler, R. A. Shweder & G. Herdt, pp. 205–32. Cambridge University Press. [NSN]

Ruble, D. N., Feldman, N. S., Higgins, E. T. & Karlovac, M. (1979) Locus of causality and the use of information in the development of causal attributions. *Journal of Personality* 47:595–614. [aAC]

Russell, B. (1967) *The autobiography of Bertrand Russell, vol. 1.* Allen and Unwin. [FLB]

Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996) Statistical learning by 8-month-old infants. *Science* 274:1926–28. [aAC]

Salomon, E. & Cimpian, A. (in preparation) *Experimental evidence for an inherence heuristic.* [rAC]

Salomon, E., & Cimpian, A. (2014). The inherence heuristic as a source of essentialist thought. *Personality and Social Psychology Bulletin* 40:1297–315. [arAC]

Samuelson, W. & Zeckhauser, R. (1988) Status quo bias in decision making. *Journal of Risk and Uncertainty* 1:7–59. [aAC]

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. [aAC]

Saylor, M. M., Ganea, P. A. & Vázquez, M. D. (2010) What's mine is mine: Twelve-month-olds use possessive pronouns to identify referents. *Developmental Science* 14:859–64. [NSN]

Scholl, B. J. (2001) Objects and attention: The state of the art. *Cognition* 80:1–46. [rAC]

Schulz, L. E. (2012) The origins of inquiry: Inductive inference and exploration in early childhood. *Trends in Cognitive Sciences* 16:382–89. [arAC]

Schwarz, N. & Clore, G. L. (1983) Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology* 45:513–23. [aAC]

Scribner, S. & Cole, M. (1981) *The psychology of literacy*. Harvard University Press. [aAC]

Seiver, E., Gopnik, A. & Goodman, N. D. (2013) Did she jump because she was the big sister or because the trampoline was safe? Causal inference and the development of social attribution. *Child Development* 8:443–54. [aAC]

Setoh, P., Wu, D., Baillargeon, R. & Gelman, R. (2013) Young infants have biological expectations about animals. *Proceedings of the National Academy of Sciences USA* 110:15937–42. [rAC]

Shah, A. K. & Oppenheimer, D. M. (2008) Heuristics made easy: An effort-reduction framework. *Psychological Bulletin* 134(2):207–22. [VD]

Shea, D. (2010) Pat Robertson: Haiti "cursed" by "pact to the devil." *Huffington Post*, March 18. Available at: http://www.huffingtonpost.com/2010/01/13/pat-robertson-haiti-curse_n_422099.html. [BB]

Shweder, R. A., Mahapatra, M. & Miller, J. G. (1987) Culture and moral development. In: *The emergence of morality in young children*, ed. J. Kagan & S. Lamb, pp. 1–82. University of Chicago Press. [aAC]

Sidgwick, H. (1874/1981) *Methods of ethics*, 7th edition. Hackett. (Original work published in 1874.) [VD]

Simon, H. A. (1957) *Models of man*. Wiley. [NB]

Simon, H. A. (1982) *Models of bounded rationality, vol. 3: Empirically grounded economic reason*. MIT Press. [rAC]

Simons, D. J. & Keil, F. C. (1995) An abstract to concrete shift in the development of biological thought: The *insides* story. *Cognition* 56:129–63. [aAC]

Sloman, S. A. (1996) The empirical case for two systems of reasoning. *Psychological Bulletin* 119:3–22. [aAC]

Sloutsky, V. M. & Fisher, A. V (2008) Attentional learning and flexible induction: How mundane mechanisms give rise to smart behaviors. *Child Development* 79:639–51. [SAG]

Smetana, J. G., Schlagman, N. & Adams, P. W. (1993) Preschool children's judgments about hypothetical and actual transgressions. *Child Development* 64:202–14. [aAC]

Snyder, M. (1984) When belief creates reality. In: *Advances in experimental social psychology, vol. 18*, ed. L. Berkowitz, pp. 248–306. Academic Press. [VYY]

Snyder, M. & Swann, W. B. (1978) Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology* 36:1202–12. [aAC]

Sober, E. (1994) *From a biological point of view: Essays in evolutionary philosophy*. Cambridge University Press. [aAC]

Solak, N., Jost, J. T., Sümer, N. & Clore, G. (2012) Rage against the machine: The case for system-level emotions. *Social and Personality Psychology Compass* 6:674–90. [DG]

Spelke, E. S. & Kinzler, K. D. (2007) Core knowledge. *Developmental Science* 10:89–96. [rAC]

Stanovich, K. E. (1999) *Who is rational? Studies of individual differences in reasoning*. Erlbaum. [aAC]

Stanovich, K. E. (2011) *Rationality and the reflective mind*. Oxford University Press. [aAC]

Stanovich, K. E. & West, R. F. (1997) Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology* 89:342–57. [aAC]

Stanovich, K. E. & West, R. F. (2000) Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences* 23:645–65. [arAC, GW]

Stanovich, K. E. & West, R. F. (2008) On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology* 94:672–95. [rAC]

Stigler, S. M. (1986) *The history of statistics: The measurement of uncertainty to 1900*. Belknap Press of Harvard University Press. [FLB]

Strevens, M. (2000) The essentialist aspect of naive theories. *Cognition* 74:149–75. [aAC, MS]

Strevens, M. (2001) Only causation matters: Reply to Ahn et al. *Cognition* 82:71–76. [aAC]

Sunstein, C. R. (2005) Moral heuristics. *Behavioral and Brain Sciences* 28:531–73. [rAC, VD]

Sutherland, S. L. & Cimpian, A. (in preparation) An inherence heuristic in explanation promotes the belief that words are well suited for their referents. [aAC]

Tadmor, C. T., Chao, M. M., Hong, Y-y. & Polzer, J. T. (2013) Not just for stereotyping any more: Racial essentialism reduces domain-general creativity. *Psychological Science* 24:99–105. [NH]

Tambiah, S. J. (1968) The magical power of words. *Man* 3:175–208. [aAC]

Taylor, M., Rhodes, M. & Gelman, S. (2009) Boys will be boys; cows will be cows: Children's essentialist reasoning about gender categories and animal species. *Child Development* 80:461–81. [MR]

Thaler, R. (1980) Towards a positive theory of consumer choice. *Journal of Economic Behavior and Organization* 1:39–60. [NSN]

Todd, A. R., Molden, D. C., Ham, J. & Vonk, R. (2011) The automatic and co-occurring activation of multiple social inferences. *Journal of Experimental Social Psychology* 47(1):37–49. [YD]

Toplak, M. E., West, R. F. & Stanovich, K. E. (2014) Rational thinking and cognitive sophistication: Development, cognitive abilities, and thinking dispositions. *Developmental Psychology* 50(4):1037–48. [rAC]

Trope, Y. & Gaunt, R. (1999) A dual-process model of overconfident attributional inferences, In: *Dual-process theories in social psychology*, ed. S. Chaiken & Y. Trope, pp. 161–78. Guilford Press. [aAC]

Trope, Y. & Gaunt, R. (2000) Processing alternative explanations of behavior: Correction or integration? *Journal of Personality and Social Psychology* 79:344–54. [arAC]

Trope, Y. & Gaunt, R. (2007) Attribution and person perception. In: *The Sage handbook of social psychology*, ed. M. A. Hogg & J. Cooper, pp. 176–94. Sage. [aAC]

Tversky, A. & Kahneman, D. (1973) Availability: A heuristic for judging frequency and probability. *Cognitive Psychology* 5:207–32. [NB, rAC]

Tversky, A. & Kahneman, D. (1974) Judgment under uncertainty: Heuristics and Biases. *Science*, 185:1124–31. [AM]

Tversky, A. & Kahneman, D. (1983) Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* 90:293–315. [aAC]

Tylor, E. B. (1871/1974) *Primitive culture: Researches into the development of mythology, philosophy, religion, art, and custom*. Gordon Press. (Original work published in 1871.) [BB]

Unger, R. M. (1983) The critical legal studies movement. *Harvard Law Review* 96:561–675. [aAC]

Vallentyne, P. (1997) Intrinsic properties defined. *Philosophical Studies* 88 (2):209–19. [MAK]

Verkuyten, M. (2003) Discourses about ethnic group (de-)essentialism: Oppressive and progressive aspects. *British Journal of Social Psychology* 42:371–91. [NH]

Wason, P. C. (1960) On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology* 12:129–40. [aAC]

Waxman, S. (2010) Names will never hurt me? Naming and the development of racial and gender categories in preschool-aged children. *European Journal of Social Psychology* 40:593–610. [MR]

Weatherson, B. & Marshall, D. (2013) Intrinsic vs. extrinsic properties. In: *The Stanford encyclopedia of philosophy*, spring 2013 edition, ed. E. N. Zalta. Available at: http://plato.stanford.edu/archives/spr2013/entries/intrinsic-extrinsic. [rAC]

Weiner, B. (1985) "Spontaneous" causal thinking. *Psychological Bulletin* 97:74–84. [arAC]

Welder, A. N. & Graham, S. A. (2006) Infants' categorization of novel objects with more or less obvious features. *Cognitive Psychology* 52:57–91. [aAC]

Wellman, H. M. (1990) *The child's theory of mind*. MIT Press. [KB]

Wellman, H. M. (2010) Developing a theory of mind. In: *The Wiley-Blackwell handbook of childhood cognitive development*, ed. U. Goswami, 2nd ed., pp. 258–84. Wiley. [SAG]

Wigner, E. (1960) The unreasonable effectiveness of mathematics in the physical sciences. *Communications in Pure and Applied Mathematics* 13:1–14. [FLB]

Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D. & Tannock, R. (1999) Development of inhibitory control across the life span. *Developmental Psychology* 35:205–13. [aAC]

Williams, M. J. & Eberhardt, J. L. (2008) Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology* 84:1033–47. doi:10.1037/0022-3514.94.6.1033. [aAC, RM, CO]

Wilson, D. S., Dietrich, E. & Clark, A. B. (2003) On the inappropriate use of the naturalistic fallacy in evolutionary psychology. *Biology and Philosophy* 18:669–81. [aAC]

Wilson, E. O. (1998) *Consilience: The unity of knowledge*. Knopf. [FLB]

Wold, H. A. & Joreskog, K. G. (1982) *Systems under indirect observation: Causality, structure, prediction, 2 vols*. North-Holland. [FLB]

Wood, G. (2012) Attributing agency: Fast and frugal or all things considered? In: *Scientific approaches to the philosophy of religion*, ed. Y. Nagasawa, pp. 71–91. Palgrave Macmillan. [GW]

Woodman, G. F., Vogel, E. K. & Luck, S. J. (2012) Flexibility in visual working memory: Accurate change detection in the face of irrelevant variations in position. *Visual Cognition* 20:1–28. [SAG]

Wright, S. (1968) *Evolution and the genetics of populations, vol. 1: Genetics and Biometric Foundations*. University of Chicago Press. [FLB]

Wright, S. (1969) *Evolution and the genetics of populations, vol. 2: The theory of gene frequencies*. University of Chicago Press. [FLB]

Wright, S. (1977) *Evolution and the genetics of populations, vol. 3: Experimental results and evolutionary deductions*. University of Chicago Press. [FLB]

Wright, S. (1978) *Evolution and the genetics of populations, vol. 4: Variability within and among natural populations*. University of Chicago Press. [FLB]

Xu, F. (2007) Sortal concepts, object individuation, and language. *Trends in Cognitive Sciences* 11(9):400–406. [HR]

Yzerbyt, V. Y., Corneille, O. & Estrada, C. (2001) The interplay of subjective essentialism and entitativity in the formation of stereotypes. *Personality and Social Psychology Review* 5:141–55. [aAC, RM, VYY]

Yzerbyt, V. Y., Rocher, S. & Schadron, G. (1997) Stereotypes as explanations: A subjective essentialistic view of group perception. In: *The psychology of stereotyping and group life*, ed. R. Spears, P. Oakes, N. Ellemers & A. Haslam, pp. 20–50. Basil Blackwell. [VYY]

Zhao, J., Al-Aidroos, N. & Turk-Browne, N. B. (2013) Attention is spontaneously biased toward regularities. *Psychological Science* 24:667–77. [aAC]

Zhong, C. B. & Liljenquist, K. (2006) Washing away your sins: Threatened morality and physical cleansing. *Science* 313:1451–52. [BB]

Zhu, L., Kay, A. C. & Eibach, R. P. (2013) A test of the flexible ideology hypothesis: System justification motives interact with ideological cuing to predict political judgments. *Journal of Experimental Social Psychology* 49:755–58. [ELU]