

Peter Carruthers

Explaining the Empiricist Bias

Reply to Berent

Abstract: *Berent (this issue) critiques one of the three main proposals put forward by Carruthers (this issue), who suggests that cognitive scientists are biased against innateness-claims by the tacit assumptions of the mentalizing faculty. Berent proposes, instead, that the bias results from dissonance produced by a conflict between our innate dualism and our innate essentialism. The present response raises a number of difficulties for her argument.*

Keywords: dualism; emotion; empiricism; essentialism; innate ideas; theory of mind.

Carruthers (this issue) suggests that there are three major ways in which an innately channelled mentalizing (or ‘theory of mind’) faculty might bias the thinking of cognitive scientists, tacitly influencing which theories are deemed most plausible. The third suggestion in that paper is the one that will occupy us here. This is that cognitive scientists’ anti-nativist intuitions are a product of the fact that the mentalizing system contains a seemingly complete theory of belief acquisition, which makes scientific appeals to innate beliefs appear intuitively implausible. The tacit theory in question is that new beliefs are *either* a product of experience, *or* of testimony, *or* of inference from other beliefs. Each of the three disjuncts will have been directly confirmed numerous times during development, and the mentalizing system contains no fourth mode-of-acquisition principle. Moreover, in cases where beliefs emerge that have no obvious, directly observed,

Correspondence:
Email: pcarruth@umd.edu

provenance through one of the three modes of learning, one can easily imagine how they might have been learned in one of those ways nevertheless. (This is, after all, why debates about innate ideas are so hard: one can always spin a story about how the information was learned — directly or indirectly — from experience or from other people.)

Berent (this issue) rejects this suggestion, and proposes instead that people resist innateness claims in order to resolve feelings of dissonance that result from a conflict between two tacit beliefs: innate biological essentialism, on the one hand, and innate ontological dualism (which is itself a product of the mentalizing system), on the other. She suggests that people tacitly reason thus: (1) innate traits must be material (essentialism); (2) ideas are immaterial (dualism); so (3) there are no innate ideas. I see three major problems with this proposal.¹

The first is that it seems to subtly mischaracterize the nature of biological essentialism. The latter idea is that there is an internal material essence or core to each creature that *causes* the manifest species-specific and sex-specific properties of the creature. (It is *not* the thesis that the traits themselves are constituted by some inner material essence.) While many of those properties are physical ones (e.g. fur versus feathers), others are mental. Hence little kids told an adoption story about a piglet brought up by cows, and treated as a cow, nevertheless expect the grown-up pig to like acorns and dislike grass (Gelman, 2003). Since liking and disliking are mental states, it follows that there is no conflict between biological essentialism as such and innate mental properties in general. There would only be

¹ To be clear, I myself accept that people are both essentialists and tacit dualists; I just deny that it is the combination of the two that drives intuitive resistance to innate ideas/belief-like states. Moreover, I do actually think that tacit dualism has a role to play in another empiricist/nativist debate, this time over the nature of the mechanisms that underlie learning — specifically, whether all learning is *general* learning (e.g. connectionist, associationist, or Bayesian), or whether there are multiple domain-specific learning mechanisms (e.g. for mentalizing, for language, for social norms, and so on). Tacit dualism may bias people to think that we should, by default, employ the simplest theory of mental operations that we can. In contrast, if one takes seriously that the mind is a biological system, one might expect it to be composed of multiple specialized mechanisms, just as biological systems in general are, at all levels of description. But the bias, here, is just an effect of tacit dualism; essentialist beliefs need play no role. It may also be worth noting that tacit dualism is sufficient to explain people's fascination with brain-based claims about cognitive phenomena, which Berent (this issue) emphasizes in her reply to me. Essentialism plays no role here, either.

dissonance regarding innate ideas specifically if the folk thought that dualism applies only to belief-like mental states. But they don't.

This leads us to what I consider to be the second major problem with Berent's proposal. Naïve dualism involves much more than belief-like states. Until very recently almost all people, everywhere, believed in a separation between mind and body, of course (dualism is a human universal); and most believed that the mind could (and would) survive the destruction of the body. Yet no one, to the best of my knowledge, has ever thought that the surviving mind would comprise only belief-like states (ideas). On the contrary, the surviving soul is thought to retain emotions (e.g. guilt for sins committed in life; love for one's surviving children) and desires (wanting one's children to do well). And indeed, young children are just as likely to attribute emotions and desires to a dead mouse as they are to attribute epistemic states like belief (Bering and Bjorklund, 2004). So, on Berent's account, it seems that people should be just as resistant to innate emotions and innate desires as they are to innate beliefs. But as she acknowledges, they are not.

In response to these first two points Berent (personal correspondence) notes that it is an empirical matter whether kids think of properties like *liking acorns* as immaterial ones, and that in her own work she finds that people think emotions and desires are less immaterial in nature than ideas. No one has yet tested for the immateriality of *liking acorns* specifically, of course. But Bering and Bjorklund (2004) find that a large majority of kindergarteners think that the dead mouse in their vignette will still want to go home, and still loves his mum; and one might think that liking and loving are awfully close to one another in nature. If asked about liking acorns specifically, however, the kids might well respond to a presence-in-the-afterlife question by saying 'No', since many will think that there is no eating in the afterlife, and hence no opportunity to like or dislike acorns. (And indeed, Bering and Bjorklund find that children are somewhat less likely to attribute feelings of hunger and thirst to the dead mouse than love of its mother.) But this wouldn't make *liking acorns* more material than *loving mum*, of course. Unfortunately, the experiments conducted by Berent, Platt and Sandoboe (under review) on afterlife beliefs as evidence of non-physicality fail to control for this sort of point. Negative answers are treated (mistakenly, in my view) as evidence of materiality.

Furthermore (and this is now the third problem I see with her core idea), Berent doesn't seem able to explain something that she herself

acknowledges to be true, namely that people are biased in the opposite direction with respect to emotions: that is, they are actively inclined to think that some emotions *are* innate. This is arguably deeply puzzling from the perspective of her views, since emotions, just like ideas, are intuited to be immaterial (even if we concede that they are intuited to be *less* immaterial, somehow). But it has a ready explanation from my own perspective: it is because (a) the mentalizing system contains no theory of desire and emotion acquisition (let alone anything that looks like a complete theory of affective learning) and (b) it is part of our everyday experience that some emotions and desires emerge very early in infancy, and that many of them manifest in similar ways in other mammals as well. So it is quite natural to think of these as innate.

Let me now turn to say something briefly about the experimental evidence that Berent and her colleagues have collected (Berent, Platt and Sandoboe, 2019; under review; Berent, 2020). First, and most importantly, their experiments were simply not designed to test between the two competing explanations of the anti-nativist bias under discussion here. Rather, they sought evidence consistent with Berent's own proposal. Yet many of the data are equally consistent with my own view. For example, consider the finding that people are less likely to think that cognitive traits (such as thinking about magic, or keeping track of people's age) would emerge spontaneously in a desert-island-upbringing situation than would emotional and motor traits (including surprise at an unexpected event, or yawning when tired). Since there would appear to be no opportunity on a desert island to learn the information required for many of the cognitive traits, and yet since the mentalizing system biases us to think that all ideas *are* learned (in one of the three ways), we should expect people to intuit that they will not appear. But there is no such mentalizing-based presumption with respect to emotional and motor traits, and so no resistance to these appearing in a child growing up alone on a desert island.

A final point to make is that there appears to be a significant methodological problem with all of the experiments in which people are asked to think about innateness itself. This is that the notion of innateness is framed by the experimenters in biological terms rather than mental ones. So it seems to be little wonder that people are inclined to see a connection between innateness and physicality. Here are the first two sentences of the prompt used: 'Inborn traits are ones that develop in humans/infants spontaneously. Some of these traits

(e.g., having five fingers) are present in birth, but others (e.g., facial hair in men) can appear later in development.’ This suggests quite strongly that innate traits are those that are biologically fixed (or genetically caused), and the examples used for illustration will prime for physicality. But in the context of debates about innateness in cognitive science this is arguably quite the wrong notion to use. Rather, to say that a mental state or mechanism of any sort is innate is to say that it emerges in normal development *without learning* (Samuels, 2002; Ritchie, 2020). In contrast with the prompt used by Berent and colleagues, this begs no questions about the physical status of innate mental traits.

Berent might reply that her experiments are designed to probe the intuitions of ordinary people about innateness, not those of cognitive scientists (she makes this point in her commentary in this issue). But there seem to me to be two problems with this response. The first is that the ultimate target to be explained (for her as for me) is the resistance to innateness-claims *in cognitive science*. If the folk turn out to diverge from cognitive scientists in some relevant respect here, then it becomes unclear what the point is of experimenting on the folk. But the second problem is that the folk conception of innateness is a complete mess in any case, a mish-mash of distinct factors (Griffiths, 2001; Griffiths, Machery and Linquist, 2009), whereas what is at stake when cognitive scientists debate the innateness of some mental property (roughly, ‘unlearned’) is comparatively clear and straightforward.

In addition, Berent points out (personal correspondence) that some of her experiments stress the absence of learning in the description of the scenario, rather than the property of being inborn, yet the results are the same: people are more likely to think that emotional and motor traits will emerge without learning than will belief-like ones. But these are the experiments that ask participants what properties one might expect to emerge in someone brought up alone on a desert island (discussed above), or what properties one might expect to see emerging spontaneously in infants in the absence of learning (Berent, Platt and Sandoboe, 2019). And, as we have seen, these findings admit of an alternative explanation within my own framework, since the mentalizing system biases us to think that beliefs are acquired through one of three specified modes of learning, whereas it contains nothing resembling a complete theory of learning for emotions.

I acknowledge, of course, that the question before us is ultimately an empirical one. Granting that there is a bias against the innateness of ideas and beliefs, is the cause of that bias something internal to the

operations of the mentalizing faculty, as I suggest? Or does it result from a conflict between our tacit dualism and our tacit biological essentialism, as Berent maintains? I hope that experiments might be devised that can more effectively test between these two theories.²

Acknowledgments

I am grateful to Iris Berent, Chris Masciari, Shen Pan, and Aida Roige for their comments on an earlier version of this article.

References

- Berent, I. (this issue) Empiricist intuitions arise from an ontological dissonance: Reply to Carruthers, *Journal of Consciousness Studies*, **27** (7–8).
- Berent, I. (2020) *The Blind Storyteller: How We Think About Human Nature*, Oxford: Oxford University Press.
- Berent, I., Platt, M. & Sandoboe, G. (2019) People’s intuitions about innateness, *Open Mind*, **3**, pp. 101–114.
- Berent, I., Platt, M. & Sandoboe, G. (under review) How we reason about innateness: The role of dualism and essentialism, [Online], <https://doi.org/10.31234/osf.io/vy6j5>.
- Bering, J. & Bjorklund, D. (2004) The natural emergence of reasoning about the afterlife as a developmental regularity, *Developmental Psychology*, **40**, pp. 217–233.
- Carruthers, P. (this issue) How mindreading might mislead cognitive science, *Journal of Consciousness Studies*, **27** (7–8).
- Gelman, S. (2003) *The Essential Child: Origins of Essentialism in Everyday Thought*, Oxford: Oxford University Press.
- Griffiths, P. (2001) What is innateness?, *The Monist*, **85**, pp. 70–85.
- Griffiths, P., Machery, E. & Linquist, S. (2009) The vernacular concept of innateness, *Mind & Language*, **24**, pp. 605–630.
- Ritchie, J.B. (2020) What’s wrong with the minimal conception of innateness in cognitive science?, *Synthese*, [Online], <https://doi.org/10.1007/s11229-020-02543-0>.
- Samuels, R. (2002) Nativism in cognitive science, *Mind & Language*, **17**, pp. 233–265.

² Editor’s note: *JCS* will not be considering further papers in this particular debate. However, this paper will be uploaded to the *JCS* blog following publication (<https://www.imprint.co.uk/category/jcs-blog/>), where the authors concerned (and *JCS* readers generally) can continue to discuss the topic further.