

HUME'S LEGACY

A Cognitive Science Perspective

Mark Collier

Introduction

Hume's naturalistic approach to the study of human beings is a central part of his philosophical project (Stroud 1977; Garrett 1997). But the naturalist interpretation gives rise to an important challenge: how can we evaluate the status of Hume's positive account? Hume offers a series of empirical claims about why we think, feel, and act as we do. In order to determine the merits of these proposals, then, we must assess them in light of our best available theories and evidence from cognitive science.

How well does Hume's experimental philosophy fare when put to this test? Does it teach us anything of lasting importance? Some naturalist interpreters are pessimistic about its prospects. Consider the stance of Barry Stroud, for example, who refuses to endorse any of Hume's specific proposals:

[If Hume's] contributions are to be judged as part of the empirical science of man . . . then his 'results' will appear ludicrously inadequate, and there will be no reason to take him seriously.

(Stroud 1977: 223)

This type of disparaging assessment is unfortunate. There is in fact a good deal of support for many of Hume's empirical claims about the workings of the mind. Several of his hypotheses have contemporary defenders in cognitive science; others can guide future research in the field.

The goal of this chapter is to examine the main proposals of Hume's experimental philosophy from a cognitive science perspective. It assesses the current status of Hume's proposals concerning the origin of ideas (particular and general), the development of our fundamental beliefs about the world (causes, objects, and gods/spirits), the operations of the emotions, the psychological prerequisites of cooperation, and the natural foundations of moral and aesthetic judgment.

Hume offers a wide-ranging account of the principles of human nature, and as a result, this examination is broadly interdisciplinary and draws from every corner of cognitive science, including connectionist simulations of cognitive development, exemplar models of classification and memory, associative theories of contingency learning, appraisal theories of emotion, moral

and cultural psychology, experimental game theory and neuroeconomics, the cognitive science of religion, and experimental aesthetics.

Concept empiricism

Hume attempts to explain how the mind is furnished with particular ideas. His central hypothesis is that complex ideas arise automatically from associative connections between sensory primitives. The laws of association function as “principles of union or cohesion among our simple ideas,” as he puts it, and “supply the place of that inseparable connexion, by which they become united in the imagination” (T 1.1.4.6; SBN 12). Just as Newton discovered the gravitational attraction that governs the motions of celestial bodies, so Hume claims to have located the “gentle force” or mental dynamics that binds together the fluctuating contents of sensation (T 1.1.4.1; SBN 10; cf. T 1.1.4.7; SBN 13).

Connectionist models of concept acquisition allow us to develop Hume's proposal further by providing greater precision about the nature of associative learning.

Connectionism can be regarded as the outcome of returning to the original vision of the associationists, adopting their powerful idea that contiguities breed connections, and applying that idea with an unprecedented degree of sophistication.

(Bechtel and Abrahamsen 1991: 102)

Hume's classical version of associationism postulates that mental contents are naturally related in memory when they co-occur or resemble one another, but he does not attempt to explain the mechanisms that underwrite these connections. Neural networks fill this explanatory gap: connection strengths between processing units are adjusted over time in accordance with general-purpose learning rules.

Consider a simple connectionist model of associative learning: the Pattern Associator (Rumelhart, McClelland et al., 1986: 34). Its architecture consists of two layers of units connected to each other through weighted connections. Suppose the network's task is to associate the sight and smell of a rose. Each pattern of activation consists of a vector of four visual and olfactory sensory dimensions. Associative learning in pattern associators is governed by an unsupervised algorithm, known as Hebb's Rule, which gradually changes connection weights according to the inner product or similarity between input patterns. We can express this type of correlation learning in terms of a mathematical equation (Elman, Bates et al., 1996: 57).

When a pattern associator is presented with visual and auditory images of a rose, connections between units in the weight matrix are automatically adjusted to capture pair-wise regularities between sensory features. After training, each sensory pattern will automatically recall the other.

$$\Delta W_{ij} = n \cdot A_i \cdot A_j$$

$$\Delta W_{ij} = \text{memory trace}$$

$$n = \text{learning rate}$$

$$A_i = \text{activation of visual feature } i$$

$$A_j = \text{activation of olfactory feature } j$$

Figure 33.1 Hebbian Learning

This toy model illustrates how complex ideas can be learned from experience; the associative learning principles remain the same, moreover, when we scale up the model and include more sophisticated competitive algorithms such as backpropagation (Rumelhart, McClelland and the PDP Research Group 1986: 37). But is there any reason to think that human concept acquisition actually works this way? Jesse Prinz points to recent evidence from behavioral and neuroscientific studies in support of the Humean proposal that concepts are perceptually based. The behavioral tests prompt subjects to answer questions about the sensible qualities of objects (e.g. “Are lemons yellow?” “Are cranberries tart?”). Response times indicate a significant processing cost for switching between perceptual dimensions, which suggests that these objects are mentally represented in terms of sensory modalities. Neuroimaging studies provide further support for this account by revealing that sensory areas in the brain become active when we reflect on concepts (Prinz 2016: 779; cf. Prinz 2002: 127–132).

This research does not provide conclusive evidence, of course, that complex ideas are learned from experience rather than innately specified. More work is obviously required to settle the longstanding controversy over the origin of ideas. But these recent studies do show that the Humean approach to concept acquisition remains a “live theoretical option” embraced by a number of cognitive scientists (Prinz 2016; 781; cf. Prinz 2006: 111).

General ideas

Hume attempts to settle the early modern debate about whether categories, such as DOG, are mentally represented in terms of abstract ideas (Locke) or particular exemplars (Berkeley). The exemplar approach was traditionally criticized on the grounds that finite minds could not possibly represent, at one and the same time, a vast store of category members. But Hume demonstrates how the principles of association offer a rapid, parallel account of search and retrieval in memory. A category such as DOG is stored as a vast network of exemplars, according to Hume, connected by the perceptual similarity of their component features. We do not consciously entertain each of these items, since our attention span is quite limited; nevertheless, when we speak or think in general terms, these instances are activated at the level of habit, “as if they were actually present” in our awareness (EHU 12.20 n34; SBN 158n).

Contemporary exemplar theory provides a good deal of support for the plausibility of Hume’s theory of general ideas. Concepts are represented in memory, according to exemplar theorists, as separately stored instances, and recall occurs in massively parallel fashion: instances automatically resonate with similar items in memory in much the same way that tuning forks vibrate (Collier 2005a: 204). There is evidence from behavioral experiments, moreover, that concepts are stored in memory as fully determinate particulars. When people are asked to describe dogs in the Arctic, for example, they typically refer to thick fur, even though this property is not mentioned in other contexts (Prinz 2016: 780).

Cognitive scientists have also run experiments, however, that support the Lockean alternative. In a classic study, Posner and Keele (1968) asked participants to categorize a variety of training instances. They subsequently presented them with the previously unseen prototype of each category and measured their response times. The results showed that subjects were able to recognize the prototypical member as quickly as the training instances, which suggests that they must have extracted and stored an abstract summary of the category during the learning phase (Posner and Keele 1968: 354).

But exemplar models can also explain these behavioral results. Subjects compute the summed similarity of category members, according to exemplar theorists, yet they do so on the fly during search and retrieval. Exemplar theory also predicts that unseen prototypes,

therefore, will be recognized as quickly as training instances (Medin and Schaffer 1978: 214). It is extremely difficult to tease apart these rival explanations of prototype and exemplar theory, then, since they make very similar predictions about reaction times in categorization tasks (Collier 2005a: 204).

It appears that we are left, in the end, with two adequate accounts of general ideas. On one side are prototype theorists who, like Locke, believe that the mind stores abstract ideas of the shared features of category members; on the other side are exemplar theorists who, like Hume and Berkeley, maintain that the mind only stores fully determinate instances. Thus, cognitive science has not vindicated Hume's attempt to resolve the controversy over abstract ideas. But it does force us to rethink our assessment of Hume's positive account of how categories are represented in memory. Hume scholars have typically shied away from embracing his proposal; Passmore refers, for example, to the "notorious inadequacy" of Hume's theory of general ideas (Passmore 1952: 40). But this evaluation now appears unfair: contemporary versions of Hume's exemplar account lie at the cutting edge of research in cognitive psychology.

Probabilistic and causal induction

Hume offers a naturalistic account of how we make inductive inferences. He distinguishes between probabilistic and causal induction. The former type of inferences are prompted by *inconstant* conjunctions and characterized by varying degrees of uncertainty; the latter are triggered by *constant* conjunctions and constituted by feelings of necessity.

Hume maintains that our everyday probabilistic inferences can be fully explained in terms of the principles of association. When we observe either *small samples* of *As*' followed by *Bs*, *mixed frequencies* (*As* usually followed by *Bs*, but sometimes by *Cs*), or *partial resemblances* (*As* resemble each other in some respects, but not others), these event-types are weakly united in our imaginations, and novel *A*-tokens will lead us to expect *Bs* with "proportionably" low levels of inductive confidence (T 1.3.12.25; SBN 142).

There is empirical support from modern contingency learning theory for Hume's associationist explanation of probabilistic inference. Psychologists have developed an experimental paradigm to study how people learn about relations between events in the world. Subjects are presented with information about the relative frequencies with which event types co-occur, and they are asked to estimate their degree of statistical relationship. These studies converge on the fact that participants typically make implicit probability estimates that are proportional to the observed co-variation in the data (Collier 2005b: 30).

These experimental results can be explained, moreover, in terms of associative learning. A standard explanation of contingency learning in psychology is the Rescorla–Wagner model, a competitive learning rule that updates associative weights on a trial-by-trial basis. The results can also be mechanistically explained in terms of adaptive neural network models, whose learning algorithm, the delta rule, is formally equivalent to the Rescorla–Wagner rule. There is a good deal of evidence for Hume's proposal, then, that everyday probabilistic inferences can be psychologically explained at the level of associations.

What about Hume's naturalistic account of causal induction? Does it also receive support from cognitive science research? Hume maintains that when we observe constant conjunctions between events, there is a point where we no longer wait for additional confirmation: our inductive confidence jumps to a sense of certainty (T 1.3.14.20; SBN 165). This empirical proposal has not, however, been vindicated experimentally (David Shanks, personal communication). Confidence levels in causal learning tasks develop incrementally, and researchers have not observed any asymptotic jumps to maximal levels of assurance.

Hume acknowledges the limits of associationism, moreover, in explanations of causal cognition. His official position is that human beings are capable of sophisticated casual reasoning, even if many of our everyday inferences depend entirely on custom and habit. Indeed, there is a principled reason why he must hold this view: if causal inferences were based on nothing but associations, we could never discover that this was the case (Collier 2007: 179–180). Scientists of human nature would observe a *second-order* constant conjunction: their fellow men make causal attributions whenever they observe constant conjunctions, but they would not be able to figure out *why* they do so. In order to understand the psychological mechanisms responsible for these attributions, Hume recognizes, experimental philosophers must rely on explicit methodological rules for judging causes and effects.

Contemporary debates about the nature of causal reasoning often pit associative against rule-based proposals. But we can now see that framing the debate in this way gives rise to a false dilemma. Hume teaches us that both associations and rules are required in order to provide a complete account of causal inference.

Object permanence

Hume attempts to understand why we naturally believe in object permanence. Belief in the continued existence of objects is difficult to explain, however, in terms of perceptually based learning from experience. Our sensory encounter with objects, after all, is permeated by gaps. Whenever we blink or turn our heads, our acquaintance with objects become discontinuous. We can represent these broken sequences in the following way:

AAABBB__DDDEEE

The crucial question is how we learn to fill in the gaps, as it were, and infer that the occluded qualities (CCC) continue to exist unperceived.

Hume suggests at one point that belief in the persistence of objects can be explained in terms of the *coherence* of gappy series with complete sequences stored in memory. But he abandons this suggestive proposal because he failed to see how associative principles could impose a greater order than is perceived (T 1.4.2.21; SBN 197).

Research in cognitive science suggests that Hume might have been premature, however, in abandoning this coherence account of our natural belief in object permanence. Connectionist researchers have simulated the cognitive development of young children's belief in enduring objects, and these computational models provide a good deal of support for the proposal that the mind automatically fills in perceptual gaps by incorporating the sensory input into previously learned event categories (Munakata et al., 1997). The networks are trained on data consisting of sequences involving objects that pass behind occluders. In the early stages of training, the network's hidden units do not reveal any representations of temporal persistence. But as the network learns to predict the re-emergence of objects, the gaps in perception gradually cancel each other out, and the network comes to infer the continued existence of unperceived objects.

Connectionist models of cognitive development provide an existence proof for Hume's conjecture that our natural belief in a persisting world of objects can be learned through the association of partial sequences with complete standards in memory. Hume's theory of imagination was simply missing an important ingredient: a computational process, known as vector-completion, that enables the mind to recognize and fill in partially degraded input (Collier 1999: 162–164; cf. Churchland 1995: 280).

Emotions

Hume embraces a “feeling theory” of emotions, in which affections are individuated according to their distinctive phenomenological properties (Collier 2011a: 4–7; Prinz 2016: 783). This view has been subjected to severe criticism. It is often said that feeling theory cannot be reconciled with core aspects of emotions, such as their intentionality and susceptibility to rational assessment (Pitcher 1965: 327–329). But these standard objections are not as strong as they initially appear. It is true that emotions are *about* states of affairs; but they can derive this intentionality from the thoughts that cause them; similarly, emotions can be rationally assessed according to the reasonableness of their eliciting beliefs.

The armchair dismissal of feeling theory has prevented scholars from appreciating the insights of Hume's naturalistic approach to the emotions. Consider Hume's account, for example, of the causal antecedents of self-conscious emotions. Hume maintains that indirect passions such as pride and humility depend on two main variables: *valence* (pleasant or painful) and *agency* (relation to ourselves or others). This proposal receives a good deal of support from quantitative studies in social psychology in which subjects are asked to rate the causes of emotions along multiple dimensions. The experimental results confirm Hume's hypothesis that we feel pride when we perceive ourselves as responsible for pleasant outcomes, and we feel shame when we view ourselves as the source of unpleasant ones (Collier 2011a: 9–10).

There is also empirical support for Hume's conjectures about the psychological principles involved in emotional processing. Hume sketches an account where affective and cognitive pathways “assist and forward each other” (T 2.1.4.4; SBN 283). Suppose that you admire a beautiful house in your possession. This object produces a pleasant sensation and simultaneously turns attention toward yourself. It is this “double impulse” of ideas and impressions, according to Hume, that generates feelings of pride (T 2.1.5.5; SBN 287). There is ample evidence from contemporary psychology for the main outlines of Hume's proposal: affective and cognitive mechanisms are both involved in emotional processing (Zajonc 1984; Lazarus 1984).

Recent work in social psychology also corroborates Hume's speculations about the behavioral effects of emotions. Hume maintains that pride is a useful emotion that gives us “confidence and assurance” in our projects (T 3.3.2.8, T 3.3.2.14; SBN 597, 600). This claim has been vindicated by recent experiments: participants induced to feel pride demonstrate greater tenacity and perseverance in performing difficult tasks (Williams and DeSteno 2008: 1010).

Hume's feeling theory fares quite well, then, when examined from the perspective of contemporary research on the emotions. Cognitive scientists have not been directly influenced by Hume's writings, but their experimental findings “confirm and extend aspects of his account” (Prinz 2016: 785).

Moral sentimentalism

Hume attempts to understand the natural foundations of moral judgment. Why do we approve of some character traits and condemn others? Hume maintains that our evaluations are based primarily on affections and sympathy; reason is delegated a secondary role. But how well does this descriptive proposal fare when evaluated from the perspective of contemporary moral psychology?

It is helpful to distinguish two components of Hume's moral sentimentalism. The first involves *immediate* affective responses to the perception of character traits; the second involves moral judgments *mediated* by our sympathy with the welfare of others. An example of the

first type of appraisal is our affective response to the perception of uncleanness (EPM 8.13; SBN 266–267). We do not consider the downstream tendencies of such traits; rather, we automatically feel a type of moral disapproval constituted by feelings of repugnance. An example of the second type of judgment is our condemnation of almsgiving, which is based on thoughts about the harmful unintended consequences of this practice (EPM 2.18; SBN 180).

There is evidence from contemporary moral psychology that a significant portion of our everyday moral evaluations occur immediately. Consider Jonathan Haidt's influential research on implicit affective responses to vignettes such as harmless incest relations between siblings. We tend to judge the actions as morally wrong even though we cannot provide any explicit justification for these attitudes. It appears that our commitments are based on disgust reactions that take place before we have time to reason (Haidt 2012: 38). According to Haidt, the influence of the affects on our moral evaluations is corroborated by a wide range of studies in social psychology, and this experimental evidence serves to support "Hume's perverse thesis: that moral emotions and intuitions drive moral reasoning, just as surely as a dog wags its tail" (Haidt 2001: 830; cf. Haidt 2012: 49).

Another important line of evidence in favor of Hume's proposals concerning our immediate moral appraisals comes from research in experimental philosophy conducted by Shaun Nichols. Nichols asks participants to judge whether it is morally wrong for a guest at a dinner party to spit in a cup and drink it. This type of transgressive behavior violates etiquette conventions, but it does not harm others or violate their rights; one would expect, then, that it would be judged to be morally neutral. But this is not what Nichols found: a statistically significant number of participants appraised the spitting behavior as morally wrong (Nichols 2004: 22). It appears that our moral judgments in such cases depend on automatic feelings of disgust. Nichols concludes his paper by writing that "cognitive science is poised to build an empirical case that would vindicate Hume's speculation" that moral judgments depend on our natural emotional repertoire (Nichols 2004: 29).

What about the second class of moral judgments described by Hume, where evaluations are mediated by sympathy with the psychological states of others? One crucial piece of support for this proposal comes from recent work on psychopaths, who lack the ability to care about the pains or pleasures of others. Psychopaths typically fail to distinguish between morality and convention; killing innocent persons is wrong, like chewing gum in class, merely because authorities forbid it. Research on psychopathy and moral reasoning is still in early stages, but these initial studies provide preliminary support for the claim that emotional concern for the welfare of others is a necessary condition for the capacity to make reflective moral judgments (Nichols 2004: 81–82; cf. Kennett 2006).

Jesse Prinz has challenged the Humean thesis that sympathy is a causal precondition of reflective moral judgments (2011). One of Prinz's principal arguments is that the Humean proposal cannot handle cases where we ourselves are victims of crimes; it is rather strange to suppose, after all, that our moral condemnation is based on sympathy with ourselves (Prinz 2011: 219). But Hume's account can accommodate such cases. When we adopt the general point of view, we rely on cognitive pretense to bracket our beliefs about our personal interests and simulate the attitudes we would have if we were properly situated spectators (Collier 2010: 260). In this counterfactual scenario, we *would* feel disapproval toward the perpetrators, and thus when we judge a crime against us to be morally wrong, we are reporting this hypothetical disinterested feeling.

Hume's moral sentimentalism has clearly influenced the direction of contemporary moral psychology. Researchers have extended his proposal that moral evaluations are based on immediate as well as sympathy-based affective responses. Nichols is surely to correct when he remarks that "were Hume alive today, this research would be high on his reading list" (Nichols 2004: 4).

Moral diversity

In his essay, "A Dialogue," Hume argues that our core values are based on shared principles of human nature and, consequently, that moral disagreements between cultures can be rationally resolved. According to Hume, skeptics exaggerate the extent of moral diversity: there is in fact a great deal of overlapping consensus in the codes of different societies (EPM, D 27; SBN 334; cf. EPM 1.10; SBN 174). There are notable discrepancies as well. But these divergences can be explained away, he proposes, in terms of differences in factual belief or material circumstances: dueling was widely approved in France, for example, but only because they mistakenly thought that it promotes social utility (EPM, D 34; SBN 335).

The Humean approach to moral disagreement has recently been challenged on empirical grounds (Brandt 1954; Doris and Plakias 2008). Hume might have successfully explained away traditional cases of moral diversity, such as the disputes about the permissibility of infanticide. But critics claim to have isolated novel instances of moral disagreement that are *ultimate* in the sense that they remain even when we control for cognitive or socio-economic differences.

Richard Brandt was a pioneer in this approach. He conducted ethnographic studies on Hopi Reservations in the 1940s' that, he claims, reveal a fundamental moral disagreement between the Hopi and other citizens of the United States. Brandt observes in particular that the Hopi are more permissive about inflicting suffering on animals, and he could not find any factual disagreements that could explain away these attitudes (Brandt 1954: 214–215). There are reasons to doubt, however, that Brandt has discovered a genuine ultimate disagreement. It is not clear that the Hopi unanimously approve of animal cruelty, and even if they did, it is not obvious that the rest of the country exhibits greater concern for animal welfare (Collier 2013: 46).

Contemporary researchers in empirically informed moral psychology claim to improve on Brandt's approach and locate an indisputable case of ultimate moral disagreement (Doris and Stich 2005: 132; Doris and Plakias 2008: 316). These researchers do not point to ethnographic cases, such as the Nitsilik Inuits or Montenegrins; rather, they cite instances of moral disagreement much closer to home: divergent attitudes, among populations in different geographical regions of the United States, toward violent responses to personal affront. When presented with vignettes in which one person insults or cuckolds another, for example, Southerners are twice as likely as Northerners to assert that violent reprisals are "extremely justified" (Nisbett and Cohen 1996: 31–32). This disagreement appears, moreover, to be fundamental. It cannot be explained away in terms of cognitive differences, since everyone clearly understands the described facts; nor can one cite socio-economic factors, because these populations live in similar material conditions (Doris and Stich 2005: 135–136; Doris and Plakias 2008: 319–320).

The Humean approach to moral diversity can, however, accommodate this ethnographic data. First, subjects in the different regions do appear to have conflicting factual beliefs; Southerners who ascribe to codes of honor are more likely to believe, for example, that one's reputation is at stake in the vignettes. Second, these populations presumably have different conceptions of the utility of these codes; Southerners typically regard them, unlike Northerners, as serving a useful deterrence function. Finally, the available evidence indicates that honor codes are in fact dangerous when removed from their pastoral roots; the main conclusion of Nisbett and Cohen (1996)'s research, after all, is that the culture of honor is the primary factor responsible for higher rates of homicide in the South (Collier 2013: 48–49).

The Humean approach to moral diversity withstands challenges, then, from contemporary critics who draw on the emerging field of cultural psychology. Skeptics have yet to locate a single clear case of moral diversity that resists explanation in terms of the adaptation of universal values to particular contexts.

Justice

Hume offers a natural history of justice. His central thesis is that our ancestors gradually learned that they would each be better off if they mutually restrained from appropriating one another's possessions and reciprocated in economic exchanges of surplus goods and services. He recognizes that these justice conventions are difficult to explain in large-scale societies, however, where agents must trust that strangers will cooperate. These implicit agreements become even more puzzling when one assumes, as Hume does, common knowledge about the human tendency to discount the future. Why should we expect that strategically rational agents would act on their interests, given that human beings are shortsighted creatures prone toward impulsivity? (T 3.2.7.2–3; SBN 535)

One of Hume's central insights is that this dilemma, like collective action problems more generally, cannot be resolved by appeal to rational considerations alone; emotional dispositions are necessary to explain how we manage to cooperate with each other in large-scale societies. Consider Hume's solution to what we might call the Chastity Assurance Game. How can we trust our partners to remain faithful? Given the human propensity to discount the future, we would have little assurance that our partners will do so, even when they acknowledge that fidelity promotes their long-term interests. One can only cooperate with partners, therefore, who are emotionally disposed to feel "repugnance" at the very idea of infidelity (T 3.2.12.5; SBN 572). These emotional sanctions occur during deliberation and thus would reduce the immediate temptation to cheat, even in contexts where one can reasonably expect to get away with defections (Frank 1988: 82). This suggests a similar account of the psychological foundations of justice conventions: we can trust others to reciprocate in temporally extended games only if they are emotionally disposed to feel repugnance to the very idea of injustice.

What evidence is there for this conjecture? Hume reasonably worries that his proposals might be dismissed as "chimerical speculation" (T 3.2.12.6; SBN 572). Hume's natural history of justice refers, after all, to events that occurred before any written records. Does this mean that the psychological prerequisites of cooperation are beyond our empirical reach? Fortunately, researchers in game theory and neuroeconomics have developed experimental paradigms that allow us to put these speculative claims to the test. Neural activity of participants is measured during games involving social exchange and collective action. These studies indicate that areas of the brain dedicated to processing negative hedonic rewards become active when players make uncooperative moves in economic games; this reward circuitry enables participants to cooperate in these games, moreover, by inhibiting their impulsive desire for immediate gratification (Collier 2011b: 142).

Previous commentators have emphasized the game-theoretic notions in Hume's account of convention (Hardin 2007). But this is not what is distinctive about Hume's approach. Hume demonstrates that considerations of strategic rationality alone are insufficient to account for social cooperation. Our ancestors could not have established justice conventions unless they were also creatures with a heart.

Religion

In *Natural History of Religion*, Hume attempts to understand the psychological foundations of the widespread belief in gods and spirits. His explanation comes in the form of "motivated irrationality" (Kail 2007: 199). We naturally believe in the existence of supernatural agents because doing so reduces our fear and anxiety in the face of uncertainty. We acquire an illusion of control over our future welfare when we imagine that our fates are decided by anthropomorphic beings that can be propitiated by prayer and sacrifice.

Hume's account of the natural foundations of religious belief receives a good deal of support from recent work on our psychological responses to the perception of uncertainty. This research confirms that human beings are indeed vulnerable to biases such as wishful thinking and illusions of control in unpredictable situations; we also exhibit a strong tendency toward superstition and anthropomorphism in these circumstances. Hume's natural history is based on general psychological principles, then, which have been corroborated by experimental studies (Collier 2014a: 670; cf. De Cruz 2015: 658).

Contemporary researchers in the cognitive science of religion deny, however, that motivated reasoning accounts for the culturally recurrent belief in supernatural agents; they attempt to explain ontological commitment to gods and spirits, rather, in terms of purely cognitive principles, such as a "hyperactive agency detection device" (Barrett 2004: 32) or the cultural transmission of counterintuitive concepts (Boyer 1996: 95). These researchers agree with Hume that folk religious beliefs are based on universal principles of human nature rather than philosophical reasoning or argumentation, but they deny that that these commitments are motivated by passions and anxiety reduction.

The cognitive science of religion is a relatively new field, and it remains to be seen which of these alternative accounts will emerge as the best explanation of the natural foundations of religion. But it is clear that the Humean proposal, which assigns a central role to emotions, deserves a seat at the table.

Aesthetics

In his influential essay, "Of the Standard of Taste," Hume maintains that aesthetic values are based on principles of human nature. This empirical assumption buttresses his proposal that qualified (i.e. fully informed and unbiased) critics would arrive at a consensus on the merit of artwork. Skeptics would, of course, deny these claims; they maintain that aesthetic values are culturally conditioned, and as a result, that disagreements are rationally irresolvable disputes about taste (E 230). If Hume is going to refute the skeptic, then, he must discharge the assumption that aesthetic values are founded on universal aspects of the mind. But what empirical evidence could be marshaled for this controversial claim?

Recent work in experimental aesthetics provides some preliminary support for Hume's proposal. In these studies, subjects are presented with formal patterns and are asked to rank them in terms of their aesthetic appeal. These test figures consist of simple qualities and relations, so it is safe to assume that everyone is fully informed about their properties. Participants in these studies, then, would count as qualified critics. The crucial question is whether their aesthetic judgments would converge.

Consider the intriguing experiments run by Richard Latto and his colleagues. These researchers present subjects with two types of paintings: (a) early Mondrian paintings composed of horizontal and vertical spatial relations, and (b) "pseudo-Mondrian" paintings whose lines were rotated such that they became oblique. The experimental results show that participants systematically prefer the genuine over the pseudo-Mondrians (Latto et al., 2006: 983).

Why is this? The best explanation of these results, according to members of this research group, lies with the "oblique effect" in neuropsychology: orientation detection cells in the visual cortex have difficulty processing oblique contours (Latto et al., 2006: 986). There appears to be a neural basis, then, for these aesthetic preferences. The dynamic balance of the Mondrian paintings, it seems, is naturally attractive to our eyes.

Although the field of experimental aesthetics is still in its infancy, its initial results support Hume's proposal that aesthetic values are based on principles of human psychology (Collier

2014b). One might object that these studies are limited to simple formal qualities and relations; but these are the only types of stimuli that can be used in controlled experiments. As far as we can tell from experimental data, then, Hume appears to have been on the right track when he proposed that “some particular forms or qualities, from the original structure of the internal fabric, are calculated to please, and others to displease” (E 233).

Conclusion: Humean minds

When we examine Hume’s science of human nature from the perspective of cognitive science, we can see that it has much more going for it than previously recognized. Let us conclude by mentioning one implication of this revisionary evaluation for a recent debate in the Philosophy of Mind.

Fodor and Lepore attempt to refute Paul Churchland’s connectionist theory of mind on the grounds that it is Humean in character (Fodor and Lepore 1996: 160–161). Their argument runs along the following lines.

- P1: Connectionism is Humean.
- P2: Hume was “born pre-refuted.”
- C: Connectionism has been refuted.

Paul Churchland’s official reply to this challenge is to deny the first premise (Churchland 1996: 278–283). There are independent reasons to doubt whether Churchland successfully distances connectionism from Hume (Prinz 2006: 107–110). But we can now see that the motivation for this reply has been undercut. Connectionists can reply to Fodor and Lepore by denying their second premise. Many of Hume’s specific proposals about the workings of the mind have emerged as leading contenders in cognitive science. Neo-Humeans are in a position, moreover, to turn the tables on their opponents. One is left wondering whether Fodor’s Neo-Cartesian approach would fare as well when put to the empirical test.

References

- Barrett, J. L. (2004) *Why Would Anyone Believe in God?* Walnut Creek, CA: Altamira Press.
- Bechtel, W. and A. Abrahamsen (1991) *Connectionism and the Mind: An Introduction to Parallel Processing in Networks*, Cambridge, MA: Wiley-Blackwell.
- Brandt, R. (1954) *Hopi Ethics*, Chicago, IL: University of Chicago Press.
- Boyer, P. (1996) “What Makes Anthropomorphism Natural: Intuitive Ontology and Cultural Representations,” *Journal of the Royal Anthropological Institute* 2: 83–97.
- Churchland, P. (1995) *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*, Cambridge, MA: MIT Press.
- . (1996) “Second Reply to Fodor and Lepore,” in R. McCauley (ed.), *The Churchlands and Their Critics*, Oxford: Wiley-Blackwell, pp. 278–283.
- Collier, M. (1999) “Filling the Gaps: Hume and Connectionism on the Continued Existence of Unperceived Objects,” *Hume Studies* 25: 155–170.
- . (2005a) “Hume and Cognitive Science: The Current Status of the Controversy over Abstract Ideas,” *Phenomenology and the Cognitive Sciences* 4: 197–207.
- . (2005b) “A New Look at Hume’s Theory of Probabilistic Inference,” *Hume Studies* 31: 21–36.
- . (2007) “Why History Matters: Associations and Causal Judgment in Hume and Cognitive Science,” *Journal of Mind and Behavior* 28: 175–188.
- . (2010) “Hume’s Theory of Moral Imagination,” *History of Philosophy Quarterly* 27: 253–273.
- . (2011a) “Hume’s Science of Emotions: Feeling Theory without Tears,” *Hume Studies* 37: 3–18.

- . (2011b) "Hume's Natural History of Justice," in C. Taylor and S. Buckle (eds.), *Hume and the Enlightenment*, London: Pickering and Chatto, pp. 131–142.
- . (2013) "The Humean Approach to Moral Diversity," *Journal of Scottish Philosophy* 11: 41–52.
- . (2014a) "The Natural Foundations of Religion," *Philosophical Psychology* 27: 665–680.
- . (2014b) "Toward a Science of Criticism: Aesthetic Values, Human Nature, and the Standard of Taste," in M. Bruhn and D. Wehrs (eds.), *Cognition, Literature, and History*, New York: Routledge, pp. 229–242.
- De Cruz, H. (2015) "The Relevance of Hume's Natural History of Religion for Cognitive Science of Religion," *Res Philosophica* 92: 653–674.
- Doris, J. and A. Plakias (2008) "How to Argue About Disagreement: Evaluative Diversity and Moral Realism," in W. Sinnott-Armstrong (ed.), *Moral Psychology*, vol 2, Cambridge, MA: MIT Press, pp. 303–332.
- . And S. Stich (2005) "As a Matter of Fact: Empirical Perspectives on Ethics," in F. Jackson and M. Smith (eds.), *The Oxford Handbook of Contemporary Philosophy*, New York: Oxford University Press, pp. 114–154.
- Elman, J., E. Bates, M. Johnson, A. Karmiloff-Smith, D. Parisi and K. Plunkett (1996) *Rethinking Innateness: A Connectionist Perspective on Development*, Cambridge, MA: MIT Press.
- Fodor, J. and E. Lepore (1996) "Reply to Churchland," in R. McCauley (ed.), *The Churchlands and Their Critics*, Oxford: Wiley-Blackwell, pp. 159–162.
- Frank, R. (1988) *Passions Within Reason: The Strategic Role of the Emotions*, New York: W. W. Norton.
- Garrett, Don. 1997. *Cognition and Commitment in Hume's Philosophy*. New York: Oxford University Press.
- Haidt, J. (2001) "The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review* 108: 814–834.
- . (2012) *The Righteous Mind: Why Good People Are Divided by Politics and Religion*, New York: Pantheon.
- Hardin, D. (2007) *David Hume: Moral and Political Theorist*, New York: Oxford University Press.
- Kail, P. (2007) "Understanding Hume's Natural History of Religion," *Philosophical Quarterly* 57: 190–211.
- Kennett, J. (2006) "Do Psychopaths Really Threaten Moral Rationalism?" *Philosophical Explorations* 9: 69–82.
- Latto, R., D. Brain, and B. Kelly (2006) "An Oblique Effect in Aesthetics: Homage to Mondrian," *Perception* 29: 981–987.
- Lazarus, R. (1984) "On the Primacy of Cognition," *American Psychologist* 39: 124–129.
- Medin, D. and M. Schaffer (1978) "Context Theory of Classification Learning," *Journal of Experimental Psychology: Human Learning and Memory* 7: 241–253.
- Munakata, Y., J. McClelland, M. Johnson and R. Siegler (1997) "Rethinking Infant Knowledge: Toward an Adaptive Process Account of Successes and Failures in Object Performance Tasks," *Psychological Review* 104: 686–713.
- Nichols, S. (2004) *Sentimental Rules: On the Natural Foundations of Moral Judgment*, New York: Oxford University Press.
- Nisbett, R. and D. Cohen (1996) *Culture of Honor*, Boulder: Westview Press.
- Passmore, J. (1952) *Hume's Intentions*, Cambridge: Cambridge University Press.
- Pitcher, G. (1965) "Emotion," *Mind* 74: 326–346.
- Posner, M. and S. Keele (1968) "On the Genesis of Abstract Ideas," *Journal of Experimental Psychology* 77: 353–363.
- Prinz, J. J. (2002) *Furnishing the Mind: Concepts and Their Perceptual Basis*, Cambridge, MA: MIT Press.
- . (2006) "Empiricism and State Space Semantics," in B. Keeley (ed.), *Paul Churchland*, New York: Cambridge University Press, pp. 88–112.
- . (2011) "Against Empathy," *Southern Journal of Philosophy* 49: 214–233.
- . (2016) "Hume and Cognitive Science," in P. Russell (ed.), *The Oxford Handbook of Hume*, New York: Oxford University Press, pp. 777–791.
- Rumelhart, McClelland, and the PDP Research Group. (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Volume 1*, Cambridge, MA: MIT Press.
- Stroud, B. (1977) *Hume*, London: Routledge & Kegan Paul.
- Williams, L. and D. DeSteno (2008) "Pride and Perseverance: The Motivational Role of Pride," *Journal of Personality and Social Psychology* 94: 1007–1017.
- Zajonc, R. (1984) "On the Primacy of Affect," *American Psychologist* 39: 117–123.