Hume on Mental Representation and Intentionality

Jonathan David Cottrell
Wayne State University, jonathan.cottrell@wayne.edu

Follow this and additional works at: https://digitalcommons.wayne.edu/philofrp

Part of the Philosophy Commons

Recommended Citation
Cottrell JD. Hume on mental representation and intentionality. Philosophy Compass. 2018; e12505.
https://doi.org/10.1111/phc3.12505

This Article is brought to you for free and open access by the Philosophy at DigitalCommons@WayneState. It has been accepted for inclusion in Philosophy Faculty Research Publications by an authorized administrator of DigitalCommons@WayneState.
HUME ON MENTAL REPRESENTATION AND INTENTIONALITY

JONATHAN COTTRELL
WAYNE STATE UNIVERSITY

PENULTIMATE VERSION; PLEASE CITE OR QUOTE FROM PUBLISHED VERSION, AVAILABLE ONLINE AT <HTTPS://DOI.ORG/10.1111/PHC3.12505>

ABSTRACT

The past two decades have seen an explosion of literature on Hume’s views about mental representation and intentionality. This essay gives a roadmap of this literature, while arguing for two main interpretive claims. First, Hume aims to naturalize all forms of mental representation and intentionality, i.e. to explain them in terms of properties and relations that are found throughout the natural world (not just in minds) and that are not, individually, peculiar to representational or intentional things. Second, Hume holds that the passions are not representational, but do have intentionality extrinsically.

KEYWORDS

David Hume; mental representation; intentionality; naturalism; impressions; ideas; passions
Claims and arguments about mental representation—that is, representation by mental items or perceptions—are central to Hume’s philosophy. For example, in Treatise Book 1 and the first Enquiry, he aims to clarify what certain philosophically important ideas (a species of perception) are “of” or “represent,” with radical results: he argues that none of our ideas represent external objects, causal necessity, or persons as his philosophical opponents claim to understand these things; so, we cannot conceive these things to be as his opponents claim they are (T 1.2.6, 1.3.14, 1.4.6, E 7).¹ In Treatise Book 2, he claims that passions (which are impressions, another species of perception) do not represent anything (T 2.3.3.5/415). Based on this claim, he argues that reason cannot compete with the passions for the control of the will (T 2.3.3.5–17/415–17), and (in Book 3) that reason is not the source of “moral distinctions,” i.e. our grasp of the difference between virtue and vice (T 3.1.1.9–10/458). This is the foundation of his case against rationalist theories of morality, and in favor of his sentimentalist alternative.²

To understand these central claims and arguments, we must understand Hume’s views about mental representation—most fundamentally, his view about what it is for a perception to represent something.

Philosophers today often discuss mental representation together with intentionality: the property of being of, about, or directed on something.³ Many mental items have this property.

---

¹ Hume’s works are cited as follows. References to ‘T’ are to Hume (2007), followed by book, part, section and, where appropriate, paragraph numbers. References to ‘E’ are to Hume (2000), followed by section and, where appropriate, paragraph numbers. These are followed by page numbers in Hume (1978) and (1975), respectively, set off by a slash mark.


³ Crane (2001, chapter 1) is a helpful ahistorical introduction to intentionality.
(Some philosophers—but not Hume—hold that all mental items have it.) For example, Romeo’s love is directed on Juliet. His belief (in Act 5) that Juliet is dead is directed on a certain state of affairs: namely, Juliet’s being dead.4

Some philosophers use ‘representation’ and ‘intentionality’ as synonyms.5 But this should be avoided in connection with Hume. He may not intend his talk of “representation” to express a form of intentionality. Even if he does, he may not regard representation as the only form of intentionality. We should not prejudge these questions at the outset.

The past two decades have seen an explosion of literature on Hume’s views about mental representation and intentionality. Here, I aim to give a roadmap of this literature, while arguing for two main interpretive claims. First, Hume aims to naturalize all forms of mental representation and intentionality. Second, he holds that the passions are not representational, but are extrinsically intentional.

The first claim needs explanation. In recent decades, numerous philosophers have argued that intentional and representational properties, such as being directed on Juliet (a property of Romeo’s love) or representing Juliet as being dead (a property of his belief in Act 5), are not explanatorily basic features of the universe; instead, they can be explained in terms of properties and relations that are found throughout the natural6 world (not just in minds) and that are not,

---

4 At this point in the play, Juliet is not dead. So, this example brings out an important and puzzling feature of intentionality: intentional mental items can be directed on things that do not exist, or (as in this example) on states of affairs that do not obtain.

5 For example, see Jacob (2014).

6 As Hume notes, the word ‘nature’ is several-ways ambiguous (T 3.1.2.7–9/473–75). Different interpretations of it will give rise to different conceptions of naturalization. For present purposes, we can afford to ignore the ambiguity because, on any reasonable interpretation of the word ‘nature’, the properties and relations in terms of which Hume...
individually, peculiar to intentional or representational things.\textsuperscript{7} A philosopher who provides such an explanation can be said to have \textit{naturalized} intentionality and representation (or, equivalently, to have given \textit{a naturalistic theory} of them). If I am correct, then Hume aims to naturalize all the intentional and representational properties of minds and perceptions. For him, the properties and relations to which a naturalistic theory may appeal (its \textit{explanantia}) would include sensible qualities such as colors, tastes, and smells;\textsuperscript{8} the intrinsic qualities of our passions, e.g. the distinctive quality that we experience when we feel anger; and the seven kinds of “philosophical relations” that he distinguishes (T 1.1.5), i.e. resemblance, identity, spatiotemporal relations, quantitative relations, qualitative relations, contrariety, and causation. I argue below that, on any plausible interpretation of Hume, \textit{resemblance} and \textit{causation} play central roles in his naturalistic theory of mental representation and intentionality.

\textsuperscript{7} Jerry Fodor (1987, 97) gives a classic statement of this view: “I suppose that sooner or later the physicists will complete the catalogue they’ve been compiling of the ultimate and irreducible properties of things. When they do, the likes of \textit{spin}, \textit{charm}, and \textit{charge} will perhaps appear upon their list. But \textit{aboutness} surely won’t; intentionality simply doesn’t go that deep. It’s hard to see, in face of this consideration, how one can be a Realist about intentionality without also being, to some extent or other, a Reductionist. If the semantic and the intentional are real properties of things, it must be in virtue of their identity with (or maybe of their supervenience on?) properties that are themselves \textit{neither} intentional \textit{nor} semantic. If aboutness is real, it must really be something else.”

\textsuperscript{8} The inclusion of sensible qualities on this list may be surprising, because mental items are not usually believed to have such qualities. Hume, however, holds that some perceptions have—or, better, \textit{are instances of}—sensible qualities; see §3, below. Cf. Garrett (2006, 302–3).
§1 Perceptions and “Representation”

For Hume, the building-blocks of our mental lives are perceptions. He divides these into impressions, the building-blocks of experiences, and ideas, the building-blocks of thoughts (T 1.1.1.1/1, E 2.1–3/17–18). He subdivides impressions into impressions of sensation, which compose our sensory experiences, and impressions of reflection, which compose our passionate and sentimental experiences (T 1.1.2.1/7–8, 2.1.1.1–3/275–76). Perceptions (whether impressions or ideas) are either simple, i.e. partless, or complex, i.e. composed of parts (T 1.1.1.2/2). Every simple idea is copied from, i.e. causally derives from and resembles, a simple impression (T 1.1.1.7/4). At least some perceptions “represent”: for example, a simple idea “exactly represent[s]” the impression from which it is copied (ibid.); and the abstract idea of a man “represents men of all sizes and all qualities” (T 1.1.7.2/18). But what exactly is a perception? And how should we interpret Hume’s talk of “representation”?

Broadly speaking, there are three possible views of what a perception is. The Act View says that perceptions are perceivings (a kind of mental act).\(^9\) The Object View says that perceptions are things perceived (objects of perceivings).\(^10\) And the Dual-Aspect View says that a perceiving

---

\(^9\) Wright (2009, 60) ascribes the Act View to Hume. The evidence that Hume accepts this view includes his claims that “[t]o hate, to love, to think, to feel, to see; all this is nothing but to perceive” (T 1.2.6.7/67) and that “all the actions of seeing, hearing, judging, loving, hating, and thinking, fall under this denomination [i.e., fall under the term ‘perception’]” (T 3.1.1.2/456). These claims suggest that perceptions are acts of hating, loving, thinking, and so forth.

\(^10\) Norton (2000, I16–I17), Stroud (1977, Chapter 2), and Waxman (1994, 10 and 18) ascribe the Object View to Hume. For evidence that Hume accepts the Object View, see below.
and a thing perceived are two inseparable aspects of a perception (or, in Hume’s terminology, that there is only a “distinction of reason”\footnote{For Hume on “distinctions of reason,” see T 1.1.7.17–18/24–25.} between a perceiving and its object).\footnote{Ainslie (2015, 211–17) and Alanen (2006, 182–83 and 193) ascribe the Dual-Aspect View to Hume. Broughton (2006, 43) may also mean to do so: she claims that perceptions are “conscious states” (perceivings?) that have “contents” (things perceived?). The evidence that Hume accepts the Dual-Aspect view includes his distinction between “the objects, of which we were thinking” and “the action of the mind in the meditation” (T 1.3.8.16/106), and his related distinction between two resemblances in our minds: a resemblance among “perceptions” and a resemblance among “act[s] of the mind” (T 1.4.2.35n/204–5; see also T 1.2.5.21/61–62).} 12

Which view we attribute to Hume will shape how we interpret his talk of “representation.” For example, consider his claim that every simple idea “represent[s]” its corresponding impression, i.e. that from which it is copied. If he holds the Act View, this claim may mean that every simple idea \textit{re-presents} the object that its corresponding impression \textit{presented}, i.e. that simple ideas are perceivings directed on the same objects (the same perceived things) as their corresponding impressions. Similarly, if he holds the Dual-Aspect View, this claim may mean that every simple idea’s object-aspect exactly resembles that of its corresponding impression; again, this could be described as \textit{re-presentering} what the impression \textit{presented}. Either way, his claim means that a simple idea is \textit{of} or \textit{about} whatever the impression was \textit{of} or \textit{about}.

But what if Hume holds the Object View, on which ideas are things perceived? When we say that something perceived “represents” something, we often mean that the former is \textit{of} or \textit{about} the latter. For example, when we say that a painted portrait that we see “represents” a certain person, we mean that the portrait is \textit{of} or \textit{about} that person. Analogously, if Hume holds the Object View, his claim about simple ideas “representing” impressions may mean that every idea is \textit{of} or \textit{about} whatever
about its corresponding impression. Indeed, it is hard to see what else this claim could mean, if he holds the Object View.\footnote{According to some, Hume’s claim that simple ideas “represent” impressions means only that they resemble impressions. Karl Schafer (2015, 1000n17) notes that Peter Loptson proposed this to him. Loptson’s interpretation is consistent with the Object View. However, as Schafer observes, it fits badly with Hume’s claim that “Ideas always represent the objects or impressions, from which they are deriv’d, and can never without a fiction represent or be apply’d to any other” (T 1.2.3.11/37). Here, Hume equates representing an “object or impression” with being “apply’d to” it. As Schafer notes, it is hard to see what the phrase ‘be apply’d to’ could mean in this context, other than be used for thinking of or about. Presumably, an idea can be used for thinking of or about a certain thing only if the idea is, itself, a representation of or about that thing. So, Hume does not seem to use ‘represent’ to express (only) resemblance, when he claims that ideas represent impressions. Instead (or, perhaps, in addition) he uses it to express of-ness or about-ness.}

There are at least three reasons to think Hume holds the Object View. First, he holds that a mind’s perceptions (and nothing else) are “present with” or “present to” that mind (T 1.1.3.1/8, 1.2.6.7–8/67–68; see also T Abs 5/647), and that something present to a mind appears to (T 1.4.2.37/206) or is perceived by (for example, is seen or felt by) that mind (T 1.4.2.38/206–7). Taken together, these commitments imply that perceptions are things perceived.

Second, Hume treats the following claims as if they were equivalent: i) we (insofar as we are “vulgar”) “attribute a distinct continu’d existence to the very things [we] feel or see” (T 1.4.2.14/193, italics added); and ii) we “attribute a distinct and continu’d existence” to “impressions” (T 1.4.2.15/194, italics added). His doing so suggests that, in his view, impressions are felt or seen, hence perceived.
Third, Hume aims to naturalize the intentionality involved in perceiving—the directedness of perceiving on an object (the thing perceived). He asks “what we mean by this seeing, and feeling, and perceiving” (T 1.4.2.38/207), and answers:

External objects are seen, and felt, and become present to the mind; that is, they acquire such a relation to a connected heap of perceptions, as to influence them very considerably in augmenting their number by present reflections and passions, and in storing the memory with ideas. (T 1.4.2.40/207)

Here, Hume speaks of “external objects,” but later in the paragraph he calls these “sensible objects or perceptions,” i.e. impressions of sensation. (This is because he is explaining a “vulgar” belief that mistakes such impressions for “bodies” or “external objects.”) In his view, then, an impression’s being perceived consists in its being causally related to other perceptions in a distinctive way (so as “to influence them very considerably,” etc.). Causal relations are found throughout the natural world; they are not peculiar to intentional or representational mental items. So, in this passage, Hume is trying to naturalize the intentionality involved in perceiving.

The Act and Dual-Aspect Views conflict with this goal. They imply that every simple perception is, or has an aspect that is, an act of perceiving. Simple perceptions are the fundamental entities in Hume’s theory of mind: he does not try to explain their intrinsic nature in other, more basic, terms. So, the Act and Dual-Aspect Views imply that perceiving involves an explanatorily basic kind of intentionality: the directedness of a simple perception on its object (Act View) or of its act-aspect on its object-aspect (Dual-Aspect View). And so, if Hume held either view, he could

---

14 Hume explains what it is for an impression to be perceived in terms of that impression’s causal relations to mental representations: ideas formed in the memory. This explanation counts as naturalistic only if Hume also naturalizes the representationality of these ideas. In the following sections, I argue that he aims to do so.
not consistently aim to naturalize the intentionality involved in perceiving. In contrast, the Object View does not say that simple perceptions are, or have aspects that are, acts of perceiving. If Hume held this view, he could consistently aim to naturalize the intentionality involved in perceiving. Since he has this aim, we should interpret him as holding the Object View.

In the following sections, I will assume that Hume holds the Object View, and that when he says ideas “represent” impressions, he means that they are of or about those impressions—not that they re-present things that those impressions once presented. But readers who reject these assumptions can still accept much of what I say, by construing my claims about “perceptions” as claims about the objects or object-aspects of perceptions.

§2 The Copy Theory of Representation

The Object View says that a mind perceives, hence is directed on, the perceptions “present to” it. In order for the mind to be directed on anything other than these perceptions, one of them must represent that thing. But what is it for a perception to represent something?

Numerous passages suggest Hume sees an important connection between representation and copying. We have seen him claim that every simple idea represents the impression from which

---

15 Ainslie (personal correspondence) notes that the Dual-Aspect View allows Hume to explain the mind’s directedness upon external objects (that is, things with a “continu’d and distinct existence”) in terms of its directedness upon “image-contents” that lack continued and distinct existence. But this would not count as naturalizing intentionality, in my sense: directedness upon “image-contents” is found only in minds, and is peculiar to mental items with intentionality; so, a naturalistic theory cannot treat this property as explanatorily basic. As I understand T 1.4.2.40/207, Hume is trying to explain the mind’s directedness upon impressions (which he calls both “external objects” and “sensible objects or perceptions”) in terms of the causal relations in which these impressions stand. This would count as naturalizing the mind’s directedness upon its impressions.
it is copied (T 1.1.1.7/4; for similar claims, see T 1.2.3.11/27 and T 1.3.14.6/157). Elsewhere, he argues that we lack ideas of certain things—objects specifically different from impressions, causal power, and a self that continues invariably throughout our lives—because we lack impressions from which to copy them (T 1.2.6.8/67–68, 1.3.14.11/161, and 1.4.6.2/251–52). And he seems to argue that passions (which are simple impressions of reflection) do not represent because they are not copies (T 2.3.3.5/415).

We may be tempted to infer that, for Hume, $x$ represents $y$ iff (if and only if) $x$ is copied from $y$. But this would be too quick. Hume allows that some complex ideas represent things from which they are not copied. For example, we can form a complex idea of the New Jerusalem, a city paved with gold and walled with rubies (T 1.1.1.4/3).

Still, this complex idea must be composed of simple ideas copied from things like parts of the New Jerusalem; and these simple ideas must be arranged like those parts. So, once we take Hume’s views about complex perceptions into account, we might hypothesize that he accepts the following theory of representation. If $x$ is simple, then $x$ represents $y$ iff $x$ is copied from $y$. If $x$ is complex, then $x$ represents $y$ iff $x$’s simple parts are copied from things like $y$’s parts and are arranged like $y$’s parts. (If $x$ is complex and copied from $y$, then $x$’s parts are copied from and arranged like $y$’s parts; so, by this theory, $x$ represents $y$.) Let us call this the Copy Theory of Representation (CTR).

As stated here, CTR is still too crude for Hume’s purposes. We have a complex idea of the New Jerusalem (T 1.1.1.4/3). But the parts of the New Jerusalem do not exist. So, no impressions resemble these parts. And so, no ideas are copied from things that resemble these parts. Taken together with CTR, this implies that we have no complex idea of the New Jerusalem. Contradiction. To avoid this problem, CTR needs to be spelled out in modal terms: If $x$ is
It is controversial whether Hume accepts CTR. In an influential paper, Rachel Cohon and David Owen (1997) argue that he does. Don Garrett (2006) argues against this interpretation. In a series of recent publications, David Landy (2012, 2015, 2017) defends the view that Hume accepts CTR against Garrett’s objections. Karl Schafer (2015) offers a compromise. The following sections address this controversy, while examining the three main kinds of perceptions that Hume distinguishes: impressions of sensation, ideas, and impressions of reflection.

§3 Impressions of Sensation

In Hume’s view, when we have sensory experiences, one or more impressions of sensation are “present to the mind,” or perceived (T 1.1.1.1/1, Abs 5/647). He claims that only sensible qualities are perceived by the senses: “If it be perceiv’d by the eyes, it must be a colour; if by the ears, a sound; if by the palate, a taste; and so of the other senses” (T 1.1.6.1/16). For Hume, only particulars exist—there are no universals (T 1.1.7.6/19). It follows that impressions of sensation are particular quality-instances: entities such as this particular instance of redness (the one in this tomato) and this particular instance of hardness (the one in this hammer).

Hume’s views about impressions of sensation may seem to show that he does not accept CTR. He sometimes says that such impressions “represent”: for example, he writes of “[t]hat complex, then x represents y iff x’s simple parts are copied from things, the zs, such that if y existed, then y’s parts would resemble the zs … etc. For ease of exposition, in the main text, I ignore this complication.

17 In Hume’s view, every particular instance of a color is identical to a particular instance of a spatial property (T 1.1.7.18/25). So, the particular instance of redness in this tomato = the particular instance of roundness in this tomato. (Really, there is just one quality-instance here, which we might call an instance of red-round-ness.) This point does not affect my arguments in this essay; for ease of exposition, I ignore it.

18 For a similar interpretation, see Kail (2007, 28–29).
compound impression, which represents extension” (T 1.2.3.15/38). But he suspends judgment about what causes them (T 1.3.5.2/84). This commits him to suspending judgment about whether they are copies (if their causes do not resemble them, they are not copies). So, it may seem, he cannot consistently accept CTR, which implies that a perception represents only if it is a copy or composed of copies.¹⁹

But this reasoning is too quick. In the Treatise, the best evidence that Hume regards impressions of sensation as representational comes from passages that may not give his considered view, because they precede his discussion of belief in external objects or “bodies” in “Of scepticism with regard to the senses” (T 1.4.2).²⁰ This section distinguishes two forms of belief in bodies: the “vulgar” form (hereafter, ‘V’), which attributes the defining properties of bodies to

---

¹⁹ Cohon and Owen (1997) seem to think that, if Hume accepts CTR, this commits him to denying that impressions of sensation represent. For the reasons below, I think this is incorrect.

²⁰ This evidence includes T 1.2.3.15/38, quoted in the main text, and T 1.2.1.5/28, where Hume writes that our senses “give us disproportion’d images of things, and represent as minute and uncompounded what is really great and compos’d of a vast number of parts.” Garrett (2006, 304) claims that T 1.4.4.14/231 provides evidence that Hume does not hold that impressions of sensation are non-representational: if he did, we should expect him to appeal to this principle here, because it would support his intended conclusion (that impressions of touch do not represent solidity); but he does not do so. Donald Ainslie (personal correspondence) claims that the same could be said of T 1.4.5.3/232–33, where Hume argues that no impression of sensation represents a substance. I am not persuaded by Garrett’s and Ainslie’s claims about these passages. It is good argumentative strategy to rely on the weakest claims that one needs, in order to support one’s intended conclusion. In T 1.4.4.14/231 and 1.4.5.3/232–33, Hume does not need to invoke the strong principle that impressions of sensation are non-representational, in order to support his intended conclusions. So, we should not expect him to invoke this principle in these paragraphs, even if he accepts it. And so, the fact that he does not invoke this principle in these paragraphs is not good evidence that he does not accept it.
what are, in fact, impressions of sensation; and the “philosophical” form (hereafter, ‘P’), which distinguishes impressions from bodies and says that the former represent and are caused by the latter (T 1.4.2.14/193, 1.4.2.31/201–2, 1.4.2.46/211–12, 1.4.2.56/217–18). If impressions of sensation represent, in Hume’s considered view, then he must accept P. Some scholars say he does, but this is questionable. He argues that P cannot be supported by reason (T 1.4.2.47/212) and “must derive all its authority” from V (T 1.4.2.39/213). Since he has argued that V is false (T 1.4.2.44–46/210–11), this may be an ironic way of saying that P has no authority at all. In the following sections, he distinguishes two sub-varieties of P: an ancient one, which construes a body as a collection of qualities inhering in a substratum (T 1.4.3); and a modern one, which construes a body as having only primary qualities (T 1.4.4). He argues that neither form of P is intelligible. So, it is not clear that he accepts P.

---

21 Hume’s view that the “vulgar” attribute continued and distinct existence (the defining properties of bodies) to what are, in fact, impressions (T 1.4.2.14–15/193–94, 1.4.2.31/201–2, 1.4.2.36/205) is a further reason to think he accepts the Object View of perceptions: if impressions of sensation were not objects—things perceived—then how could we mistake them for continued and distinct pens, papers, hats, shoes, stones, and so forth (T 1.4.2.31/202)? (I thank Don Garrett for this point.)

22 For example, see Garrett (2006, 306).

23 As evidence that Hume accepts P, Don Garrett (2006, 304) cites this Enquiry passage, which seems to endorse P: “no man, who reflects, ever doubted, that the existences, which we consider, when we say, this house and that tree, are nothing but perceptions in the mind, and fleeting copies or representations of other existences, which remain uniform and independent” (E 12.9/152). However, it is not clear that this passage gives Hume’s considered view. It may express a non-skeptical stage in a dialectic that ultimately leads to skepticism. (The passage must, at least, involve exaggeration. Hume cannot really think that “no man, who reflects, ever doubted” whether our impressions of sensation are representations of other existences. In the Treatise, he says that we do in fact have such doubts—at least for a short time—following philosophical reflection on the senses (T 1.4.2.56–57/217–18).) And even if this passage
Moreover, even if Hume does hold that impressions of sensation represent, he can reconcile this claim with CTR by distinguishing original from derived representationality. Something has derived representationality iff it inherits its representationality from something else. Many philosophers today think this is true of words and sentences: they inherit their representationality from the thoughts of people who utter and interpret them. In contrast, something has original representationality iff it is representational and does not inherit its representationality from anything else. CTR is most plausibly construed as a theory of original representationality: words, sentences, and other things with derived representationality need not be copied from the things they represent.

George Berkeley held that impressions of sensation (what he called “ideas of sense”) represent in the manner of a language: like words, they “suggest” ideas to our minds, and inherit their representationality from these ideas; so, their representationality is derived, not original.24 Nabeel Hamid (2015) argues that Hume accepts this part of Berkeley’s view. If Hamid is correct, then Hume can consistently accept CTR as a theory of original representationality, while holding that impressions of sensation have derived representationality and suspending judgment about what causes them.

---


---

does give Hume’s considered view, it provides no evidence against the view that he accepts CTR, for it explicitly equates being a copy with being a representation: “fleeting copies or representations” (E 12.9/152, emphasis added). (For the view that impressions of sensation represent iff they are copies, see Schafer 2015, 999n6.)
So, Hume’s claims about impressions of sensation provide no decisive evidence against the view that he accepts CTR as a theory of original representationality. Let us now turn to the second kind of perceptions: ideas.

§4 Ideas

As formulated in the *Treatise*, Hume’s Copy Principle states that all simple ideas are copied from, and “exactly represent,” simple impressions (T 1.1.1.7/4). I take this to mean that every simple idea is copied from, and is of or about, a simple impression. For example, a simple idea copied from a particular redness-instance (an impression of sensation) is of or about that particular redness-instance.

Hume uses the Copy Principle for both constructive and destructive purposes. *Constructively*, he uses it to clarify what certain ideas represent, by tracing them to the impressions

---

25 As formulated in the *Enquiry*, the Copy Principle states that “all our ideas or more feeble perceptions are copies of our impressions or more lively ones” (E 2.5/19). Hume must be speaking loosely here: in his own view, the (complex) idea of a golden mountain is not, itself, copied from any impression (E 2.5/19). To have any plausibility, the Copy Principle must be restricted to simple ideas. I therefore focus on the *Treatise’s* formulation of this principle, which contains this restriction.

26 Ruth Weintraub (2005) and David Landy (2012, 2015) distinguish two principles here: i) a principle concerning the causal origins of simple ideas, and ii) a principle concerning what simple ideas represent. Landy reserves the name ‘Copy Principle’ for (i).

27 Donald Ainslie (2015, 127) distinguishes two “kind[s] of copying relation,” in the context of a Dual-Aspect interpretation of Humean perceptions: i) a kind whereby a copy re-presents the object-aspect (in Ainslie’s vocabulary, the “image-content”) of its original; ii) a kind whereby a copy’s object-aspect is its original. A type-(ii) copy presents both the act- and object-aspects of another perception. Ainslie calls a type-(ii) copy a “secondary idea,” citing T 1.1.1.11/6–7.
from which they are copied (T 1.2.3.1/33, 1.3.2.4/74–75; E 7.4–5/62–63). For example, he uses it to argue that our idea of causation represents the relations of contiguity, succession, and necessary connection between cause and effect (T 1.3.2.5–12/75–77), and that this necessary connection is a mental disposition to infer effects from their causes and vice versa (T 1.3.14.15–23/162–66; E 7.28/75–76). *Destructively*, he uses it to argue that we lack certain alleged ideas, because we have no impressions from which to copy them (for examples, see §2).

Hume’s acceptance of the Copy Principle, and his two ways of using it, are consistent with the hypothesis that he accepts CTR. If a simple perception represents only that from which it is copied (per CTR), and all simple ideas are copied from simple impressions, then: i) all simple ideas represent simple impressions (per the Copy Principle); ii) we can clarify what a simple idea represents by identifying that from which it is copied (per Hume’s constructive uses of this principle); and iii) if we have no simple impression with certain characteristics, then we have no simple idea of something with those characteristics (per his destructive uses).

However, Hume’s theory of ideas provides at least two reasons to think he *does not* accept CTR. First, his discussion of modes and substances (T 1.1.6) implies that two ideas may represent different things, despite being copied from the same collection of impressions or quality-instances. For example, an idea copied from instances of yellowness, weight, malleableness and fusibility will represent a piece of gold—a particular *substance*—if it is suitably associated with a name and a disposition to incorporate other quality-instances into the collection bearing that name, under certain circumstances (T 1.1.6.2/16). But an idea copied from the very same collection of impressions will represent an instance of metallicness—a particular *mode*—if it is differently associated with a name, such that no other quality-instances may be added to the collection bearing
that name (T 1.1.6.3/17). These claims seem to conflict with CTR, which implies that any two ideas copied from the same thing(s) represent the same thing(s).

Second, Hume’s theory of “abstract ideas” (T 1.1.7)—i.e., ideas that represent general kinds of things—implies that ideas sometimes represent more, or other, than the thing(s) from which they are copied. For example, consider a simple idea copied from a particular redness-instance. Hume’s theory implies that this idea represents redness in general, if it is associated with a “general term” (for example, the word ‘red’) that is, in turn, associated with many other ideas of particular redness-instances (T 1.1.7.7–10/20–22; for the view that simple ideas can become abstract or general, see T 1.1.7.7n/637). This implication seems to conflict with CTR, which implies that a simple idea represents only that from which it is copied. Hume’s theory of abstract ideas also implies that two ideas copied from the same thing sometimes represent different things. For example, consider two simple ideas copied from the same redness-instance: one associated with the word ‘red’ and many ideas of other redness-instances; the other associated with the word ‘color’ and many ideas of other color-instances, including yellowness-instances, blueness-instances, and so forth. Hume’s theory implies that these ideas represent different things: the first, redness in general; the second, color in general. Again, this implication seems to conflict with CTR.

Faced with these passages, we have three interpretive options. First, we might argue that, appearances notwithstanding, these passages are consistent with CTR. For example, Landy (2017) argues that an abstract idea is a complex idea, composed of the many ideas associated with the relevant general term. This complex idea represents the various impressions from which its parts (the many associated ideas) are copied. So, Hume’s theory of abstract ideas is compatible with CTR, after all. This approach has problems. For example, some of the ideas associated with a
general term exist only potentially (T 1.1.7.7/20), and it is not clear that merely potential entities can compose a complex idea. Also, it is not clear that this approach can be applied to Hume’s discussion of modes and substances.

Second, we might conclude that Hume completely rejects CTR. For example, Garrett (2006 and 2015, 67–77) claims that Hume explains representation in terms of causal role. This differs from copying: causal role need not involve resemblance, and can involve a perception’s effects and well as its causes. Various causal role theories of representation are possible. Garrett’s specific proposal is that \( x \) represents \( y \) iff \( x \) reliably indicates \( y \) or \( x \) plays the functional role of \( y \). (Roughly, to say that Fs reliably indicate Gs is to say that an F occurs iff a G is present. And to say that something, \( x \), plays the functional role of something else, \( y \), is to say that \( x \)’s causal relationships are parallel to \( y \)’s.) A challenge for such interpretations is to explain the passages where Hume argues that we lack ideas of certain things because we lack impressions from which to copy them. Why couldn’t a perception reliably indicate and thus represent (say) a permanent self, without being copied from such a thing? (Compare Schafer 2015, 989–90).

Third, we might infer that Hume holds a “hybrid” theory on which there are two kinds of representation, one explained by copying, the other by causal role. Hume distinguishes i) the representation of particular quality-instances (T 1.1.1) from ii) the representation of substances, modes, and general kinds (T 1.1.6–7).\(^{28}\) He holds that (i) is necessary for (ii): in order to represent a particular substance or mode, an idea must first represent a collection of particular quality-instances composing that substance or mode (T 1.1.6); and in order to represent a general kind of quality-instance, substance or mode, an idea must first represent a particular member of that kind (T 1.1.7). So, if Hume holds a hybrid theory on which copying is necessary for (i), he will hold

---

\(^{28}\) This distinction is orthogonal to that between original and derived representationality, introduced in §3.
that being a copy is necessary for being a representation. He will also hold more specific claims about the relation between copying and representation, e.g. that being copied from a simple redness-instance is necessary for representing redness in general. These commitments may explain the passages about copying and representation that pose a challenge for non-hybrid causal role interpretations, like Garrett’s.

Karl Schafer (2015, 990–98) offers a hybrid interpretation of the kind just sketched. Donald Ainslie (2015, 64–69) offers a similar interpretation, in the context of a Dual-Aspect interpretation of Humean perceptions. For Ainslie, a kind of copying explains how an idea re-presents the “image-content” (i.e., the object-aspect) of its corresponding impression. Associations with words and other perceptions then explain how an idea with this image-content serves to represent other things, including bodies and universals. Ainslie (2015, 65) and Schafer (MS) both add that, for Hume, our uses of ideas must be subject to certain norms, in order for those ideas to represent at least some of the things they do—for example, general kinds. This additional claim is compatible with, but not required by, the kind of hybrid theory sketched above.

All three options attribute naturalistic theories of representation to Hume. CTR explains representation in terms of causal relations and resemblances, which are found throughout the natural world and are not, individually, peculiar to representations. So, it is naturalistic. Causal role theories explain representation entirely in terms of causal relations, so they too are naturalistic. And hybrid theories explain representation in terms of copying, causal relations, and—in Ainslie’s and Schafer’s versions—norms of use for ideas. While I cannot argue the point here, I think that any plausible interpretation of Hume will say that he aims to naturalize normativity. So, whether or not it appeals to norms of use, a Humean hybrid theory of representation will also be naturalistic.

---

29 Ainslie distinguishes two kinds of copying: see note 27, above.
And so, without deciding which of these theories Hume accepts, we can conclude that he aims to naturalize the representationality of ideas.

§5 Impressions of Reflection
Hume divides impressions of reflection (or “secondary impressions”) into those that are generally more violent, and those that are generally more calm (T 2.1.1.3/276). The former include the passions. The latter include certain moral and aesthetic sentiments; presumably, they also include the impression of necessary connection, which Hume says is an impression of reflection (T 1.3.14.22/165). Hume subdivides the passions into direct passions, which “arise immediately from good or evil, pain or pleasure,” and indirect passions, which arise from pain or pleasure “by the conjunction of other qualities” (T 2.1.1.4/276).³⁰

Recent discussions of Hume’s views about representation and impressions of reflection have focused on his “Representation Argument” (T 2.3.3.5/415): reason cannot compete with the passions for the control of the will, because a passion “contains not any representative quality, which renders it a copy of any other existence or modification,” and involves no “reference to any other object”; so, passions cannot be true or false; and so, they cannot conflict with “truth and reason.”³¹

---

³⁰ Hume’s distinction between direct and indirect passions raises interesting issues that are beyond this essay’s scope. For helpful discussions, see Cohon (2008a) and McIntyre (2000).

³¹ Hume writes that moral approbation and blame are “nothing but a fainter and more imperceptible love and hatred” (T 3.3.5.1/614). If this is his considered view, then the Representation Argument applies to these moral sentiments, as well as to the non-moral passions.
That Hume gives this argument may seem to show that he accepts either CTR or a hybrid theory of representation, rather than a causal role theory. He seems to claim that passions do not represent, on the grounds that they are not copies (Cohon and Owen 1997, 50–58; Schafer 2015, 984). If so, he is assuming that passions represent only if they are copies. CTR and hybrid theories explain why he should make this assumption: as we have seen, these theories imply that being a copy is necessary for being a representation.

However, the Representation Argument also alludes to the passions’ distinctive causal role. Hume writes that “when I am angry, I am actually possest by the passion, and in that emotion have no … reference to any other object” (T 2.3.3.5/415). To be “possest” by something is to be controlled by it, and “emotion” can mean an excited mental state, i.e. one that tends to cause disturbance in the mind. So, Hume may mean to argue that passions do not represent, on the grounds that their distinctive causal role—their controlling us and causing mental disturbance—renders them unable to do so. And so, the Representation Argument does not constitute decisive evidence for, or against, any of the three ways of interpreting Hume’s theory of representation that we have canvased: they all remain live options.

Whichever interpretation we accept, the Representation Argument should appear puzzling. Hume repeatedly claims that passions have “objects,” i.e. things to which they are “directed” (T 2.1.3.4/280, 2.2.1.2/329). For example, the object of pride and humility is oneself (T 2.1.2.2/277); the object of love and hatred is another person (T 2.2.1.2/329). So, Hume seems to hold that passions have intentionality. But in the Representation Argument, he seems to deny this. How, if at all, can we reconcile this argument with his other claims about the passions?

---

32 See the Oxford English Dictionary, ‘possessed’, entries 1 and 2; and ‘emotion’, entry 3.a.

33 For a proposal along these lines, see Cohon (2008b, 68n).
First, we might simply dismiss the Representation Argument. For example, Annette Baier (1991, 160) claims it is a “very silly paragraph” that is “unrepresentative” of Hume’s considered view that passions have intentionality. As Cohon (1994, 184–87) argues, this option faces a serious problem. Hume relies on the Representation Argument when arguing against moral rationalism (T 3.1.1.9–10/458). By dismissing it, Baier undermines Hume’s grounds for his rival, sentimentalist theory of morals.

Second, we might read the Representation Argument as being restricted to some part or aspect of the passions, allowing that other parts or aspects of them have intentionality. For example, Lilli Alanen (2006, 193) takes Hume to accept the Dual-Aspect View of perceptions. On her reading, his claim that passions involve no “reference to any other object” is restricted to their act-aspects, considered apart from their object-aspects (2006, 193–94). So, his Representation Argument allows that a passion has intentionality, when all its aspects are taken into account. If Hume accepts the Object View of perceptions, as I have argued (§1), then this reading is untenable. Alanen also claims that passions are “complex” (2006, 186). This suggests another, similar reading that does not depend on the Dual-Aspect View: perhaps the Representation Argument is restricted to a part of the passions, allowing that a passion as a whole involves “reference” to another object. However, I am unsure how to square Alanen’s claim that passions are “complex” with Hume’s repeated insistence that the indirect passions (at least) are “simple” (T 2.1.2.1/277, 2.2.1.1/329).34

Third, we might deal with the seeming tension between Hume’s Representation Argument and his claim that passions are “directed” on “objects” by distinguishing intentionality from representationality. On this approach, being a representation is one way, but not the only way, of

---

34 Merivale (2009) argues that Hume regards the indirect passions as simple in the Treatise, but as complex in his later Dissertation of the Passions.
having intentionality. The Representation Argument claims only that passions are not representations. This allows that they have intentionality or, as Hume says, have “objects.” So, there is no genuine conflict among his views.35

Some scholars who take this third option say that the passions’ intentionality is intrinsic to them, e.g. that a particular impression of pride is intrinsically directed on oneself.36 But if Hume accepted this, it would significantly complicate his arguments in *Treatise* Book 1. A copy shares the intrinsic properties of the original from which it derives, save for its degree of force and liveliness (T 1.1.1.5/3–4). So, if an impression (such as a passion) could be intrinsically directed on x without representing x, then so could an idea copied from this impression. Hume shows no recognition of this in *Treatise* Book 1: he claims that “ideas always represent their objects or impressions” (T 1.3.14.6/157), which does not seem to allow that an idea might be intentionally directed on something without representing it. In particular, when arguing that we lack ideas of certain things—objects specifically different from impressions, causal power, and a self that continues invariably throughout our lives—he does not consider the possibility that our ideas might be directed on these things without representing them. This suggests that he does not regard this as a genuine possibility, hence does not allow that impressions can be intrinsically directed on things without representing them.

---

35 For versions of this approach, see Cohon (1994) and (2008a), Garrett (2006), Qu (2012), Schafer (2015), Schmidt (2003, chapter 6), and Schmitter (2009). Although this is not her way of handling the Representation Argument, Alanen (2005, 124) distinguishes intentionality from representation, so this approach is open to her, as well.

36 For example, see Qu (2012). Schmidt (2003, 175) may also take this view: she writes that, for Hume, “a passion is inherently directed toward a certain type of object.”
I infer that, when Hume says passions are “directed” on “objects,” he means to attribute an extrinsic feature to them: specifically, a causal relation to an idea. When he says that a direct passion, such as a desire, has a certain “object” (T 2.3.3.3/414, 2.3.9.27/446), he means that the passion is caused by an idea of that object.\textsuperscript{37} When he says that an indirect passion, such as pride or humility, has a certain “object,” he means that the passion causes, or “turns our view to” an idea of that object (T 2.1.2.4/278). In each case, the idea is representational—it “represents … the object of the passion” (ibid.)—but the passion itself is not. This proposal is a version of our third option: passions do not represent, but they are extrinsically intentional or “directed.”

How, then, should we understand Hume’s claim that passions involve no “reference to any other object,” which seems to imply that they lack intentionality altogether? We can read this claim, more weakly, as meaning that passions are not representational\textsuperscript{38} or that they are not intrinsically intentional. Read in either way, Hume’s claim allows that passions are extrinsically intentional.

Conclusion
This essay has surveyed some of the main issues and positions in the exciting new body of literature on Hume’s views about mental representation and intentionality. Along the way, I have

\textsuperscript{37} This may seem to conflict with Hume’s claim that direct passions “arise from … pain or pleasure” (T 2.1.1.4/276), not from ideas of pain or pleasure. But elsewhere, Hume says that direct passions are caused by ideas (T 1.1.2.1/8), or by “the prospect of pain or pleasure” (T 2.3.3.3/414), i.e. an idea of some future pain or pleasure. And his discussion of pride and humility suggests that “the cause” of a passion is “that idea, which excites [it]” (T 2.1.2.4/278). I therefore think that, in Hume’s considered view, the direct passions result from ideas of pleasurable or painful things.

\textsuperscript{38} For example, Schmidt (2003, 174) claims that “Hume is using the word “reference” here for the relation of “copying” or representation between an idea and its object.”
argued that Hume aims to naturalize the intentionality involved in the mind’s perceiving of its perceptions, and that, on any viable interpretation, he aims to naturalize mental representation. I have also argued that he regards passions as non-representational but extrinsically intentional. But I do not intend, or expect, to have the last word on these issues. No doubt the literature on them will continue to grow, and others will uncover new interpretive options, or new ways of defending options that I have rejected here.39

---

39 I thank Donald Ainslie, Don Garrett, David Landy, Hsueh Qu, Lisa Shapiro and an anonymous referee for helpful feedback on earlier drafts, and Rachel Cohon, for helpful discussions.
REFERENCES


