

Le Bon Ronnie? Humean Themes in de Sousa

Donald C. Ainslie

ABSTRACT This paper argues that Ronald de Sousa's contemporary theory of the emotions bears a striking resemblance to aspects of David Hume's theory as presented in *A Treatise of Human Nature*. Both philosophers see the emotions as able to create new social facts, the assessment of which depends on further sentiments. And both philosophers argue that, once our emotional nature is properly understood, our moral concerns range much more broadly than other philosophical theories typically allow.

David Hume was a popular guest in the Parisian salons of the Enlightenment, where he became known as '*le bon David*' (Mossner, 1943). Ronald de Sousa – who has occasionally published using his more formal last name, Bon de Sousa (1966) – is welcomed in contemporary salons wherever he goes, but I argue below that he has more in common with Hume than their shared bon-iness. de Sousa describes his philosophy as having 'one central preoccupation: how best to understand how our lives turn about the pivot of our passions' (2011, xv)? Hume's most significant work – *A Treatise of Human Nature*¹ – also pivots on the passions, the topic of its central, second Book, the lessons of which echo back to his analysis of the understanding in Book 1 and point forward to his investigation of morality in Book 3. And both philosophers embrace a kind of naturalism, though neither falls prey to a reductionist scientism. That said, though de Sousa approaches his topics with a highly attuned historical sense, his normal interlocutors from the philosophical canon tend to be rationalists, such as Plato or Descartes. He rarely addresses Hume, and when he does, he does not pause to remark on the commonalities they share. I aim to rectify this problem in what follows.

My approach will be to bring out the Humean themes in de Sousa – or perhaps to put it more accurately, to bring out how Hume in the *Treatise* already embraces some of the theses that de Sousa champions. In particular, I will examine Hume's overall conception of the passions, including his famous claim that they enslave reason, in §1; his investigation of the person-oriented passions and their socially inflected role in recognizing and constituting selves in §2; and, in §3,

¹ References to Hume (1739-40/2007) will be made parenthetically as 'T' followed by the Book, Section, and Paragraph numbers as given in the Norton and Norton edition.

his appeal to moral sentiments in his explanation of our moral evaluations and their capacity for truth and falsehood.

§1. Passion and Reason

Hume's discusses the passions in Book 2 of the *Treatise* as part of his larger project of the 'science of man' – his attempt to discover how the human mind operates so that he can clarify what can be accomplished in the various special sciences (T Intro.4). As the subtitle to the *Treatise* indicates, he approaches his task using the 'experimental method,' that is, an empirically structured investigation of the phenomenon in question, akin to the approach pioneered by 'my Lord Bacon' (T Intro.7) in the 17th century. The problem, for Hume, is that, while it is possible to study the non-human world by means of carefully designed experiments, in the case of an investigation of the mind such

reflection and premeditation would so disturb the operation of my natural principles, as must render it impossible to form any just conclusion from the phænomenon. We must therefore glean up our experiments in this science from a cautious observation of human life, and take them as they appear in the common course of the world, by men's behaviour in company, in affairs, and in their pleasures. (T Intro.10)

And so Hume draws on history, literature, and everyday observations to justify his various claims about the mind. Notably, he does not appeal to the brain and its structures in his arguments (or at least he does so reluctantly and rarely – see T 1.2.5.20), a task he leaves to the 'anatomists and natural philosophers' (T 1.1.2.1).

de Sousa shares Hume's openness to a wide variety of evidence, especially appeals to literature, in justifying his analyses of the emotions. But he, in contrast to Hume, is also guided by the latest scientific literature on the emotions, whether by neuroscientists, cognitive scientists, or psychologists. For de Sousa, 'emotions [are] the 'frontier' between the biological and the mental' (1987, 154) and thus the biological sciences are an essential part of his account. But of course contemporary biology is far more advanced than what Hume had access to, the crucial difference being the Darwinian revolution (even if Hume was open to proto-evolutionary explanations at times [1779/1947, 182]). Hume, I think, would have taken a different attitude to the biological substrates for our mental lives if he had had access to the findings of our current science.

What might seem to be a more telling conflict between Hume's and de Sousa's views concerns their understanding of the roles of reason and representation in the emotions. Hume famously says the 'reason is, and ought only to be the slave of the passions' (T 2.3.3.4) in part because:

A passion is an original existence, or, if you will, modification of existence, and contains not any representative quality, which renders it a copy of any other existence or modification. When I am angry, I am actually possest with the passion, and in that emotion have no more a reference to any other object, than when I am thirsty, or sick, or more than five foot high. 'Tis impossible, therefore, that this passion can be oppos'd by, or be contradictory to truth and reason; since this contradiction consists in the disagreement of ideas, consider'd as copies, with those objects, which they represent. (T 2.3.3.5)

de Sousa, in contrast, takes many emotions to have 'formal' objects and he has literally written the book on the rationality of emotion (1987). But the disagreement here, I will suggest, is less than it seems, in part because Hume allows for a kind of intentionality to passions even if they lack strict representationality, and in part because the kind of rationality that de Sousa finds in the emotions is some distance from the kind that Hume denies to them.

Taking the issue of representation first, recall that, in the *Treatise*, Hume identifies the mind with a 'bundle' of what he calls 'perceptions,' where nothing ties them together in such a way that they have intrinsic synchronic or diachronic unity; in particular, the bundle lacks a subject that superintends and thereby unifies the perceptions (T 1.4.6.1-3). He describes 'the full examination' (T 1.1.1.7) of which perceptions 'are causes and which effects' to be 'the subject of the' *Treatise* (T 1.1.1.6-7), where causation is ultimately to be understood in the deflationary terms of the constant conjunction of similar types of objects and a related tendency of the mind to associate observed instances of them (T 1.3.14.31). His starting point is a division of perceptions into two kinds: impressions and ideas. The mind is wholly passive with respect to impressions – they thereby have high 'vivacity' – with sensations and emotions falling into this category. Ideas, in contrast, have lower vivacity, leading Hume to identify them with thoughts, in that we can

always conceive that things are otherwise than we think,² even if some ideas, notably causal beliefs, nonetheless seem to press themselves on us.

Hume's empiricism holds that (almost³) all thoughts ultimately get their content by 'copying' prior impressions, and he sometimes puts this point in terms of representation: 'ideas are deriv'd from impressions, and are nothing but copies and *representations* of them' (T 1.1.7.5, emphasis added; see especially T 1.1.1.7). This usage is unhelpful and has too often misled interpreters to take ideas to be *of impressions* – as if we are always thinking of our mental states – rather than to be of whatever the impressions are of. A related interpretive error takes impressions – even sensory impressions – not to be representations because they are not copies. Such readings strike me as bizarre. Hume's appeal to common life as the source of his 'experiments' means that he takes seriously our everyday presupposition that we sense objects around us: you give a child an impression of orange by 'present[ing] the object' to him (T 1.1.1.8). He thus repeatedly identifies sensations with 'images' as of external objects.^{4,5} In so far as ideas 'represent' impressions, they do so by *re-presenting* (an 18th-century usage) the content that the impression originally presented. The child sees *the orange* and then is able to think of *it*. An idea will represent

² With some exceptions, notably those relations that Hume takes to be objects of what he calls 'knowledge', in that they depend only on the relata (T 1.3.1.1).

³ Hume famously introduces the case of the 'missing shade of blue,' where someone who has experienced all shades of blue but one is able to think of the missing shade despite the lack of a prior impression. He describes this scenario as 'so particular and singular, that 'tis scarce worth our observing, and does not merit that for it alone we should alter' (T 1.1.1.10, SBN 6) his core empiricist principle that ideas ultimately acquire content only from prior impressions. But I take him to be signalling here that he is open to other novel sources for idea-content such as the 'manner' in which objects appear – the source for our ideas of space and time (T 1.2) – or the impression of reflection that emerges when we expect an event to happen on the basis of prior experiences of similar instances (T 1.3.14.22) – the source of the idea of causation.

⁴ See especially: 'When we . . . have the images of external objects conveyed by our senses; the perception of the mind is . . . an impression' (T Abs.5). In the main body of the *Treatise*, he describes sensory impressions as imagistic at 1.1.7.6, 1.1.7.10, 1.2.1.4-5, 1.2.3.7, 1.3.5.5, 1.3.9.9, 1.3.9.16, 1.3.10.2, 1.3.10.5, 1.3.11.13, 1.3.12.10–11, 1.3.12.22, 1.3.12.24, 1.4.2.33, 1.4.2.36, 1.4.2.38, 1.4.5.15, 1.4.7.3, 2.2.4.4, 2.2.5.5.

⁵ Part of the interpretive problem is that Hume's longest treatment of our beliefs about external objects, 'Of scepticism with regard to the senses' (T 1.4.2), is notoriously opaque. I hold Hume to be arguing there that the contents of sensory impressions on their own do not suffice for them to represent persisting external objects; rather, when such impressions appear in a regular pattern (showing 'constancy' or 'coherence'), the imagination responds in such a way to enable them *thereby* to be of such objects. That is, the content of an impression can be different depending on the role it plays in the mind's causal economy. Similarly, Hume holds that, though every idea has content that is particular, it can nonetheless represent universals if the imagination responds in the right way (T 1.1.7). For a full defense of these claims, see Ainslie 2015.

an impression – have an impression *as its object* – only when it results from introspection (or what, following Locke, Hume will sometimes call ‘reflection’⁶).

A related interpretive misunderstanding stems from Hume sometimes calling perceptions ‘objects,’ especially in his tortuous treatment of our beliefs about external objects (‘Of scepticism with regard to the senses’ [T 1.4.2]).⁷ Henry Allison thus points to ‘Hume’s lamentable inattention to (or perhaps exploitation of) the perceived–perceiving ambiguity’ (2008, 243). Given that Hume also intimates that by ‘perception’ he means what Locke calls ‘ideas’ – ‘whatsoever is the Object of the Understanding when a Man thinks’ (1700/1975, 1.1.8) – it is unsurprising that many interpreters attempt to resolve this ambiguity by taking perceptions to be mental objects of some kind (Penelhum 1975, 28-9; Stroud 1977, 17). But Locke’s theory of mind also includes mental activities, notably ‘*Perception, or Thinking*’ (1700/1975, II.vi.2), so that thought involves the *perception of ideas*, while Hume takes an idea to be a kind of perception, not its object. Indeed, when Hume identifies the mind with a bundle of perceptions (T 1.4.6), he thereby denies that there is some other mental action that takes them as its objects. Interpreters sometimes try to avoid this issue by suggesting that, for Hume, we ‘experience’, ‘attend to’ (Norton 2000, I23, I32), or are ‘conscious of’ (Waxman 1994) perceptions, but then the mind would not be a bundle of perceptions, but rather a bundle and the experiencings, attendings, or consciousnesses of them. A better solution is to recognize that sensory perceptions for Hume *combine* the mental activity of Lockean perceptions with the object-like character of Lockean ideas. Such perceptions are thus *perceivings* – states of awareness. A sensory perception is an awareness-of-sensory-content, where the awareness has the distinctive qualitative character associated with the relevant sense modality, either in its most intense form when the perception is an impression, or in a lesser form, when the perception is an idea. Hume does at times indicate that he understands perceptions in these terms, noting that they include both the ‘objects, of which we ... think’ and ‘the action of the mind in the meditation, that certain *je-ne-scai-quoi*, of which 'tis impossible to give any definition or description, but which every one sufficiently understands’ (T 1.3.8.16; see also 3.3.1.2). And, at

⁶ But, where Locke analogizes reflection to ‘internal Sense’ (1700/1975, II.i.4) and treats it as a source for the mind’s content that supplements sensation, Hume never takes reflection to, like sensation, yield impressions. Rather he says that we *think* of the mind’s activities by means of what he calls ‘secondary ideas,’ and he compares such ideas to the idea of the missing shade of blue, in that both involve a less-than-straightforward relation to impressions of sensation (T 1.1.1.11, 1.3.8.15-17).

⁷ See T 1.4.2.14, 1.4.2.18, 1.4.2.31, 1.4.2.36, 1.4.2.38, 1.4.2.40, 1.4.2.43, 1.4.2.44, 1.4.2.46, 1.4.2.48, 1.4.2.50, 1.4.2.53, 1.4.2.56.

several crucial points in his explanations of our ideas of the vacuum and of external objects, he appeals to such ‘actions of the mind’ in noting that ideas can be associated on the basis of a resemblance either of their objects or of the mental actions they involve (T 1.2.5.21, 1.3.8.16, 1.4.2.35n39).

Where does this leave the passions? I take Hume’s claim that they lack ‘representative qualities’ to mean that they do not have same structure as sensory perceptions. Rather than being awareness-of-sensory-contents, they are simply *ways of being aware* of whatever objects are on our minds. Hume at various points identifies the vivacity of a perception with the ‘manner’ in which we are aware of its object (see especially T 1.3.7), saying that different such manners ‘feel’ different (T 1.3.7.7). It seems likely then that feelings themselves are nothing but such manners of conception. That is, passions are modifications of how we perceive things. They are, as de Sousa argues in relation to emotions, *attitudes* (1987, 156).

When Hume’s passions are understood in terms of their nature as perceptions, some of what might seem to be his more radical claims about them fall into place. In particular, their lack of ‘representative qualities’ indicates only that they are mere qualifications of our awareness of things. They can still have objects in a looser sense, in that our anger, for example, while by itself objectless, will filter or overlay our thought of, say, a lover’s betrayal, itself exemplified in a complex idea that is separable and thereby ontologically independent of the anger. It might even be a strength of Hume’s account that it allows our emotions to range widely over different objects, including what might seem to be perverse cases, as when someone is angry at a particular marigold or a star light years away. That said, he does allow that some passions have special links to particular things: pride and humility take the person who has them as their ‘objects’ (T 2.1.2.2); benevolence and anger follow on, respectively, love and hatred, and thus focus on the persons who are their objects (T 2.2.6); and so on. But Hume takes these connections to be causal and contingent, reflecting ‘original’ qualities of human nature rather than analytic elements of the relevant emotion (T 2.1.3.2-3). Beyond that, for any given person, particular passions might end up tied up with some objects because of her acculturation, her personal or family history, or the literature that she read as a child.

Hume’s analysis of reason and passion also seems less radical when passions are understood as objectless modifiers of other object-oriented perceptions. For he defines reason in terms of the capacity to discover relations between *objects* (T 1.3.2.2, 3.2.5.4n77). Reason cannot

directly address the passions because the passions lack what reason concerns. Hume does, however, leave open for *indirect* address:

[’T]is only in two senses, that any affection can be call’d unreasonable. *First*, When a passion, such as hope or fear, grief or joy, despair or security, is founded on the supposition of the existence of objects, which really do not exist. *Secondly*, When in exerting any passion in action, we chuse means insufficient for the design’d end, and deceive ourselves in our judgment of causes and effects. Where a passion is neither founded on false suppositions, nor chuses means insufficient for the end, the understanding can neither justify nor condemn it. (T 2.3.3.6)

Passions that are accompanied by these types of false beliefs are ‘unreasonable’ in only a derivative sense, ‘and even then ’tis not the passion, properly speaking, which is unreasonable, but the judgment’ (T 2.3.3.7). Thus, though Hume says that reason is the ‘slave’ to the passions, in that only the passions can push someone to action, he ultimately acknowledges that both reason and passion are ‘requisite’ in understanding human behaviour. The passions are ‘blind’ without reason’s ‘direction’ (T 3.2.2.14).

de Sousa’s discussion of reason and the emotions disentangles several senses of ‘rationality.’ First, there is cognitive rationality, which relates to beliefs and judgements in general; this notion of rationality is perhaps closest to Hume’s use of ‘reason’, despite his officially restricting it to the recognition of relations. Second, de Sousa points to strategic rationality, which, similarly to Hume’s derivative sense of ‘reason’, concerns how best to secure the means to ends given by a person’s desires. Third, and perhaps most importantly, de Sousa introduces an axiological sense of rationality, in which emotions are understood to be perceptions of value and the value in question is explained ‘in terms of the formal objects determined by the nature of the emotion itself’ (1987, 131). Hume’s passions do not have formal objects in this sense, even if they might have contingent links to one thing or another. de Sousa’s ability to appeal to an evolutionary account of how emotions relate to their objects allows him to improve on Hume here. The contingency that Hume leaves as brute, de Sousa can explain in terms of the selective pressures that brought about our basic cognitive economy.

de Sousa’s suggestion that emotions can be a form of perception of *value* might also seem distinctively un-Humean, given Hume’s famous embrace of kind of subjectivism about value. But I suggest in §3, below, that the moral sentiments are best understood as perceptions of moral

features of the social world – not values as wholly independent of human attitudes, but values that we put into the world by means of the coordination of our emotional responses and our habituation into our culture.

Overall, then, while Hume treats reason as a slave to the passions, de Sousa argues for the not-too-distant thesis that emotions *frame* our strategic and cognitive rationality:

[E]motions set the agenda for beliefs and desires: we might say they ask the questions that judgment answers with beliefs and evaluate the prospects to which desire may or may not respond.... In this way emotions can be said to be judgments, in the sense that they are what we see the world ‘in terms of.’ But they need not consist in articulated propositions.... So long as we presuppose some beliefs or preexisting desires, the directive power of ‘motivation’ belongs to what controls attention, salience, and inference strategies preferred (1987, 197).

And in this Hume would agree. Our passions orient us to the world, with reason helping us find our way through it.

§2. *The Indirect Passions and the Self*

de Sousa perhaps comes closest to Hume when he recognizes a category of emotions that are not perceptions of objective value (in some sense of ‘objective’) but are rather investments in their objects that thereby make them of value to their subjects. These ‘self-related desires,’ he says, ‘give meaning to our lives’ (1987, 309). Hume points to a set of person-directed passions that similarly define us – in his case, the ‘indirect’ passions of pride, humility, love, and hatred. He devotes two thirds of his discussion of the passions in the *Treatise* to them (Parts 2 and 3 of Book 2), though he offers little explanation for this special focus. Interpreters have puzzled about why he takes this approach, especially in that he treats them somewhat generically, for example, showing little interest in distinguishing pride from self-conceit, vanity, self-esteem, self-possession, or other self-directed affects.

I have argued (1999) that Hume’s treatment of the indirect passions is best interpreted by taking seriously his claim that there is a ‘great analogy’ between his treatment of causation and his account of these passions, in that in both cases associations among perceptions yield an especially vivacious idea, respectively, a causal belief and an emotionally inflected conception of a person (T 2.1.5.11). Recall how Hume understands causation in the *Treatise*. We lack insight into intrinsic

necessary connections between purported causes and effects, and must instead rely on the experienced regularities of objects similar to them ('constant conjunction'). Even then, we do not have insight into nature's fundamental structure, and thus reason cannot yield a belief that the experienced constant conjunction will continue in the future. Ultimately, Hume thinks that it is only because our experience causes us to expect observed regularities to continue that we make causal inferences. A non-rational process of association leads us to think (have an idea) of an object similar to one conjunct upon experiencing (having an impression) of an object similar to the other. And the vivacity of the impression that triggers the association is transferred to the associated idea so that we *believe* the effect will happen upon observing the cause (or vice versa).

The analogous associative process in the case of the indirect passions involves an impression of pleasure (or pain) at something connected to a person, whereby that first impression is associated with the pleasure of pride or love (or the pain of humility or hatred), and the idea of the triggering object or feature is 'transfused' (T 2.1.5.11) into the idea of the relevant person, whom Hume calls the 'object' of the passion – either the self in pride and humility, or another in love and hatred.⁸ This process is a 'double relation of impressions and ideas' (T 2.1.5.5), in that the hedonic response to the trigger is associated with the hedonic quality of the passion and the idea of that trigger is associated with the idea of the person. And the association transfers the vivacity of the impressions into the vivacity of the resultant idea. We come to *believe* that the person is specially characterized by the feature that caused the passion. As Hume puts it, we think of the person in terms of her or his 'qualities and circumstances' (T 2.1.5.6); we 'form an idea of [her or his] merit and character' (T 2.1.8.8).

Though Hume's analogy between causation and the indirect passions focuses on the similarity in the psychological mechanisms that lead to the relevant beliefs, I think the analogy can be extended to shed light on why he adopts the associative explanations in the first place. In the case of causation, it is reason's incapacity on its own to discern intrinsic connections in nature.

⁸ Note that, by 'person' here, Hume is not referring back to his earlier discussion of personal identity, where his concern was only 'personal identity as it regards our thought or imagination' (T 1.4.6.5). There, the mind is described as a 'bundle' of perceptions lacking intrinsic unity at a time and across time. He notes that 'personal identity ... as it regards our passions or the concern we take in ourselves' (T 1.4.6.5) – is a different topic, pointing us forward to Book 2's discussion of the indirect passions. The relation between these two accounts of the person is complex. I have argued elsewhere (2005) that Hume recognizes that the bundle contains the perceptions it does because of the different ways we locate ourselves (and are located) in the social world. The appearance to us of this socially inflected pattern of experience amounts to an 'idea, or rather impression of ourselves [that] is always intimately present with us' (T 2.1.11.4) – a crucial element of his explanation of sympathy.

Given that we do make causal inferences – notably, when Hume offers causal explanations for our causal beliefs – causal reasoning must ultimately include arational associations. In the case of the indirect passions, Hume is mostly silent about why he gives so much time to explaining them. But the analogy with causation suggests that Hume is worried about how we can come to see only some of the many facts about persons to *define* them, to make them *who they are*. He thus points to a difference between those of our features that ‘are in a manner separated from us’ and those that ‘are ... consider'd as connected with our *being and existence*’ (T 2.1.8.8; emphasis added). What allows us to discern what we can call *existential connections*? Reason and the senses only enable us to recognize the panoply of facts that describe someone. Similarly to the case of causation and necessary connections, Hume instead explains our taking people to have existential connections to only some of their features by means of an associative process, the indirect passions. We *feel* what sets us apart, what makes us into who we are, what ‘distinguishes’ us from others (T 2.1.6.5).

Though these associative processes are arational, it does not mean that the resulting beliefs are immune to criticism. In the case of causation, Hume points out that sometimes our experiences of constantly conjoined objects are merely accidental, a product of our limited observations rather than of how things are in the world. Because our causal expectations are not infrequently dashed, we learn to distinguish ‘the accidental circumstances from the efficacious causes’ (T 1.3.13.11). We form what he calls ‘general rules’ on the basis of this experience and Hume even offers a ‘logic’ consisting of what he takes to be the best rules for causal investigation (T 1.3.15). But, given that all such rules are ultimately based on our experiences, they too depend on our arational tendency to associate ideas on the basis of experienced constant conjunction of their objects. For Hume, the causal world is one relative to the ‘human capacity’ (T 1.2.5.26n12.2), without our being able to say how things are in their ‘real nature and operations’ (T 1.2.5.26n12.1). Indeed we can only gesture towards such a fully mind-independent reality indirectly, by negating our human ways of thinking (T 1.1.5.10, 1.2.6.9).⁹

Hume’s treatment of the indirect passions similarly allows for criticism of our emotionally inflected beliefs about one another. First, he argues that someone is able to maintain an indirect

⁹ Note, however, that his scepticism means that he steers clear of idealism. He is aware that his claims about the mind’s role in our understanding the natural world applies just as much to his claims about the mind. We are not in a position to vindicate or repudiate our take on things. I have argued elsewhere (2015) that his scepticism amounts to a recognition that philosophy must ultimately rely on arational processes.

passion – especially pride – only if it is ‘seconded’ by others (T 2.1.11.1), who similarly find the feature that causes the passion to be admirable. Hume points to the role of ‘sympathy’ here, his name for our tendency to be infected with the emotions of those around us. Thus if I take pride in something wholly idiosyncratic, though I might end up with a belief about myself in terms of that feature, it will not be taken up by others and thus I will be unable to live my life in terms of that identity.

Second, he argues that ‘general rules’ apply to the passions just as much as they do to causal beliefs:

[G]eneral rules have a great influence upon pride and humility, as well as on all the other passions. Hence we form a notion of different ranks of men, suitable to the power or riches they are possess of; and this notion we change not upon account of any peculiarities of the health or temper of the persons, which may deprive them of all enjoyment in their possessions. This may be accounted for from the same principles, that explain'd the influence of general rules on the understanding. (T 2.1.6.8)

Because we have repeatedly observed people who take a particular feature of theirs to matter – in this case power and riches – we respond to those who have the feature by assuming that it will cause an indirect passion and a concomitant idea of themselves whether or not they do in fact care about it. Even if these people are indifferent to the feature, we treat them as though they do, and this, Hume thinks, is sufficient for them to be defined by it – to count as of that ‘rank’ – independently of their caring about it. Thus, while he says that ‘[c]ustom readily carries us beyond the just bounds in our passions, as well as in our reasonings’ (T 2.1.6.8), in that our human tendency to generalize goes beyond what reason countenances, he also says that ‘custom and practice ... have settled the just value of every thing’ (T 2.1.6.9). Just as our causal associations reveal to us a world with a causal structure, the indirect passions, regularized by sympathy and general rules, reveal to us a world with a social structure.

In the ‘Abstract’ to the *Treatise*, Hume considers how ‘a man, such as Adam, created in the full vigour of understanding, without experience’ (T Abs.11) would be unable to make sense of objects’ causal interaction. He makes a similar point about how experience with the indirect passions is necessary for making sense of our social interaction with one another:

[’T]is evident, that if a person full-grown, and of the same nature with ourselves, were on a sudden transported into our world, he wou’d be very much embarrass’d with every [person]¹⁰, and wou’d not readily find what degree of love or hatred, pride or humility, or any other passion he ought to attribute to it. (T 2.1.6.9)

In his view, we find ourselves in a social world the contours of which are determined by the indirect passions we feel for ourselves and others.

There is a sense, then, that Hume’s theories of both the natural and social worlds might be called ‘subjectivist’. de Sousa helpfully distinguishes several different senses of the subjective – the phenomenological, the projective, the relative, and the perspectival (1987, 145-8). Our causal and passion-induced beliefs are, for Hume, phenomenologically distinctive; they *feel* a certain way. But, as de Sousa points out, this kind of subjectivity does not threaten their being open to criticism for mischaracterizing how things are in the relevant respect. As for de Sousa’s projective sense of the subjective, Hume of course is famous for suggesting that ‘the mind has a great propensity to spread itself on external objects, and to conjoin with them any internal impressions, which they occasion, and which always make their appearance at the same time that these objects discover themselves to the senses’ (T 1.3.14.25). Thus we tend to assume that the necessity we experience in ourselves is in fact connecting the causally linked objects we consider. But such projection can be resisted, and Hume’s definition of causality makes no reference to it. Causal claims are, for him, true or false independently of whether we take the necessity we feel within us to bind the external cause and effect.

The social case, however, is more complex. Our projection here is of the kind involving general rules, where we expect people to care about those features that typically elicit the indirect passions. And, as I have emphasized, these expectations, regularized by sympathy and general rules, *create* the social world. In this case, when enough people project their feelings onto one another, they constitute themselves in terms of what they take to matter. The social world ‘in a manner receives [its] being’ (T 2.1.7.7) from our sentiments. This kind of subjectivity is best understood as a form of what de Sousa calls ‘relativity’. I noted above that the causal world is relative to what Hume calls the ‘human capacity’ (T 1.2.5.26n12.2). The social world is relative to more narrow social groupings, with what matters to people in ancient Greece or 18th-century France differing from what matters to Hume’s more local contemporaries. Despite these

¹⁰ Hume uses the word ‘object’ here, but he means it in the sense of ‘object of an indirect passion,’ i.e. a person.

differences, all of us are ultimately moved by the pleasure and pain we collectively feel when confronted with various features of persons. Even when there are dramatic cultural differences between social worlds, Hume holds they remain intelligible once we understand their relativity to human nature.¹¹

de Sousa's final sense of 'subjective' speaks to the perspectival nature of some of our beliefs and attitudes. Hume most clearly embraces this notion of the perspectival in his treatment of the moral sentiments, where our feelings of approval and disapproval count as fully moral only when we have made adjustments for the 'peculiar point of view' that shapes our perceptions of one another. We must correct our initial assessments by taking up 'some *steady* and *general* points of view' (T 3.3.1.15). I turn to a more detailed explanation of Hume's view of the moral sentiments in the next section.

§3. *Objectivity and Moral Evaluation*

When de Sousa considers Hume's moral theory, he tends to read it rather restrictively, grouping him with his Scottish contemporaries, Francis Hutcheson and Adam Smith, as 'identifying certain positive emotions such as sympathy and compassion as intrinsically moral, since they motivated us to behave in socially beneficial ways' (2011, 140). He complains that 'this leaves out the possibility that personal and social relations themselves generate a vast range of emotions that are more or less directly relevant to central ethical questions as traditionally conceived' (1987, 312). While this reading of Hume is perhaps not inaccurate as it relates to the simplified version he presents in the *Enquiry concerning the Principles of Morals*, it overlooks the richer view of ethical life found in the *Treatise*.¹² We have already seen that, in Book 2 of that work, Hume emphasizes how 'personal and social relations generate a vast range of emotions' – or rather how the crucial indirect passions generate personal and social relations. And he identifies four categories of causes for these passions: virtue and vice, bodily qualities, power and riches, and a miscellaneous grouping of whatever features happen to have been contingently caught up in a particular society's mutual concern (T 2.1.7-10). My focus here will be his treatment of the moral category in Book 3

¹¹ See 'A Dialogue,' appended to the *Enquiry concerning the Principles of Morals* (Hume 1751/1998), for Hume's clearest attempt to address cultural differences.

¹² Annette Baier (2008) and Jacqueline Taylor (2015) would disagree with this characterization of the second *Enquiry*.

of the *Treatise*, and how it encompasses ‘a vast range of emotions’ and beliefs, beyond merely compassion or benevolence.

Hume starts his investigation of morality by rejecting rationalist theories that see us as able to discover the ‘eternal fitnesses and unfitnesses of things’ (T 3.1.1.4). As I noted above, he takes reason to be in the business of discovering relations between objects, and any seemingly moral relation – say, killing one’s parent – can apply just as easily to clearly non-moral cases – the sapling that overshadows and ultimately kills the tree from which it was seeded (T 3.1.1.24). Moreover, Hume has already established that reason is motivationally inert; thus if reason alone discovered ethical truths, morality would be detached from passion and action (T 3.1.1.6). Instead, Hume argues, we make moral evaluations by means of distinctive feelings: ‘To have the sense of virtue, is nothing but to *feel* a satisfaction of a particular kind from the contemplation of a character. The very *feeling* constitutes our praise or admiration’ (T 3.1.2.3).

He then faces two challenges. First, feelings can seem to be too subjective to make sense of the seeming objectivity of our moral judgements. He himself says that virtues ‘command’ our approval and vices our disapproval.¹³ And we are to *correct* our initial judgements if they are partial or self-interested. What kind of standard is he appealing to when implying that our feelings might be wrong? Second, Hume identifies the primary objects of moral evaluation as what he calls ‘characters’ (as in the quotation above) or ‘mental qualities’ (T 3.3.1.3, 3.3.1.6, 3.3.1.9) – what we call character traits. But how do such traits fit into Hume’s bundle theory of the mind? They do not seem to be easily explicable as complexes of perceptions, especially in that he takes them to persist even when not manifested in behaviour, while perceptions ‘succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement’ (T 1.4.6.4). I will address the second question first and, in answering it, will show that it points to an answer to the first question as well.

My starting point is Hume’s tight linking of the moral sentiments with the indirect passions. Notably, he emphasizes that approval and disapproval are immediately taken up into the ‘double relation of impressions and ideas,’ in that virtue and vice are features that matter to us that are closely connected with their bearers. ‘It follows, that these two particulars are to be consider’d as equivalent, with regard to our mental qualities, *virtue* or the power of producing love and pride, *vice* and the power of producing humility or hatred’ (T 3.3.1.3). The indirect passions are ‘the most

¹³ T 3.1.2.4, 3.3.1.13, 3.3.1.18, 3.3.1.25, 3.3.1.27, 3.3.5.1.

considerable effect that virtue and vice have upon the human mind’ (T 3.1.2.5). We ‘approve of [a person’s] character [trait], and love his person’ (T 3.3.3.2). But why does he take the indirect passions to be so significant for his moral theory? Some interpreters take Hume to be *identifying* the moral sentiments with the indirect passions in an attempt to demystify and naturalize what can seem like an *ad hoc* account of our capacity for moral judgement (Árdal 1989; Korsgaard 1999).¹⁴ I think, however, that the interpretation of the indirect passions that I sketched above shows that his concern is how moral qualities come to be existentially connected to their bearers – and how they come into being in the first place.

Consider a virtuous trait such as generosity. A person displays this trait by, say, shoveling her neighbour’s walk after the snowfall, sponsoring refugees from Syria, or bringing home-baked cookies to share with others at the office. Each such action is for Hume the outcome of the causal interaction of various perceptions in her mind-bundle. And

[a]ctions are by their very nature temporary and perishing; and where they proceed not from some cause in the characters and disposition of the person, who perform'd them, they infix not themselves upon him, and can neither redound to his honour, if good, nor infamy, if evil. The action itself may be blameable; it may be contrary to all the rules of morality and religion: But the person is not responsible for it; and as it proceeded from nothing in him, that is durable or constant, and leaves nothing of that nature behind it, 'tis impossible he can, upon its account, become the object of punishment or vengeance. (T 2.3.2.6)

What can connect the action to the person so that it can be appropriately infixed? Hume rejects a conception of human agency that posits a superintending self that wills its actions or commits itself to a long-term policy or ‘maxim’ of helping others. Further, the action in question might be one-off, the result of a passing whim, rather than any sign of a persistent generosity. For no single action is necessary or sufficient for generosity, but rather the pattern of action, reaction, thought, and feeling must hang together in the right way for it to count, and even that pattern will depend on how it relates to the rest of the person’s behaviour and broader mental outlook.

¹⁴ Hume does at one point say that ‘our approbation or blame [of someone’s virtue or vice] ... is nothing but a fainter and more imperceptible love or hatred’ (T 3.3.5.1), and Árdal especially takes this passage to ‘furnish us with a key to the understanding of Hume’s moral theory’ (1989, 113).

I think that, for Hume, at the most basic level, the spectators' responses are what tie this pattern together. When we see the person repeatedly helping others – genuinely so, and not for ulterior motives – we sympathetically share in the pleasure she is creating and then feel an indirect passion – love – towards her. She thereby comes to count among the 'rank' of the generous. Of course, this basic level has abstracted from our having been brought up to have an implicit understanding of what kinds of behaviour count as generous and what kinds of motives count as ulterior. Our parents and others who raise us teach us the meaning of 'generosity' and the general rules for exemplifying it. 'The distinction [between virtue and vice] ... being so great and evident, language must soon be moulded upon it, and must invent a peculiar set of terms, in order to express those universal sentiments of censure or approbation' (Hume 1751/1998, 9.8).

Hume uses the example of wit to make his case:

No one has ever been able to tell what wit is, and to shew why such a system of thought must be receiv'd under that denomination, and such another rejected. 'Tis only by taste we can decide concerning it, nor are we possest of any other standard, upon which we can form a judgment of this kind. Now what is this *taste*, from which true and false wit *in a manner receive their being*, and without which no thought can have a title to either of these denominations? 'Tis plainly nothing but a sensation of pleasure from true wit, and of uneasiness from false, without our being able to *tell the reasons* of that pleasure or uneasiness. The power of bestowing these opposite sensations is, therefore, the very essence of true and false wit; and consequently the cause of that pride or humility, which arises from them. (T 2.1.7.7; original emphasis only on 'taste')

We see here that the disparate behaviour that can exemplify wit is consolidated by our responses to it, in particular our taking pleasure in it and coming to feel an indirect passion towards the person as a result. Because indirect passions, when spread by sympathy and structured by general rules, create our social 'ranks,' wit 'receives [its] being' as a trait that, like other virtues, 'commands' our approval. Those who overlook someone's wit will be 'embarrass'd' (T 2.1.6.9) in their encounters with her or him, not able to interpret her or his conversational contributions correctly. de Sousa makes a similar point when he says that 'the rich landscape of value would flatten into universal indifference if we were to discount our emotions. Objective moral facts exist, because

people make moral judgments that are relative to natural and cultural facts about human emotions' (2011, 146).

We can now answer the first question I noted above, about how Hume can treat moral evaluation as both a matter of feeling and yet something that is 'commanded' of us. Because traits depend on our mutual passionate responses and thereby become an objective feature of *our social world*, Hume can require that we take the relevant steps needed to assess them properly. First, we must ensure that our self-interest does not distort our understanding of a person's traits: "Tis only when a character is considered in general, without reference to our particular interest, that it causes such a feeling or sentiment, as denominates it morally good or evil' (T 3.1.2.4). Second, Hume acknowledges that sometimes we are not properly situated with respect to a person to assess her traits. For example, seeing someone shovel her neighbour's walk once does not suffice for a finding of generosity, with concomitant approval. Rather, we must defer to the verdicts of those who know her well enough to recognize if the shoveling hangs together with the rest of her behaviour to count as exemplifying virtue (rather than, say, a petty desire to make the neighbour feel indebted). And different traits require different perspectives on a person. A leader's greatness depends on the verdicts of those who follow her, while a father's goodness requires the perspective of his children (T 3.3.2-3). Finally, we must recognize that people's circumstances constrain the kinds of behaviour they display, and our moral assessments must take these factors into account. A pauper's sharing of a crust of bread might be an instance of benevolence, while a wealthy person's doing something similar might manifest contempt. Or as Hume puts it: '[V]irtue in rags is still virtue' (T 3.3.1.19). When we correct our sentiments for these three factors, we thereby judge from what he variously calls a 'steady,' 'general,' or 'common point of view' (T 3.3.1.15, 3.3.1.30).

Hume ultimately argues that, from this corrected perspective, the virtues can be identified as those traits that are useful or agreeable to their bearers or those who interact with them, with the vices those traits that are harmful or disagreeable (T 3.3.1). Through sympathy we feel pleasure or pain when the relevant behaviour is displayed, and this feeling is then taken up in the 'double relation' so that we see persons as defined by the relevant qualities of mind and these qualities in turn are solidified as elements of our social world. He takes a very broad understanding of morality to follow, including such traditional virtues as courage, prudence, generosity, and the like, but also

other socially salient traits that speak to natural abilities, such as intelligence or good humour (T 3.3.4), or even to elements of our bodily comportment such as an ‘air of health’ (T 3.3.5.4).

The result is that each of us bears many virtues and vices, namely whatever behaviour-involving features get at who we are. Hume illustrates his conception of morality by ending each chapter in his *History of England* with a description of the ‘characters’ of the relevant king or queen and some of the most influential figures of the era. For example, in a multi-paragraph character sketch, Elizabeth I is *inter alia* noted for her ‘vigour, her constancy, her magnanimity, her penetration, vigilance, [and] address,’ though she ‘guarded not herself with equal care or equal success from lesser infirmities; the rivalry of beauty, the desire of admiration, the jealousy of love, and the sallies of anger’ (1754-61/1983, IV.xliv, 351). Hume says of Charles I: ‘The character of this prince, as that of most men, if not all men, was mixed; but his virtues predominated extremely above his vices’ (1754-61/1983, V.lix, 542). Hume clearly rejects the thought that the virtues display some kind of unity.

de Sousa also calls for ‘an expansion of the ethical sphere beyond the domain of obligation, duty, and the dichotomy of right and wrong.’ Like Hume, he acknowledges a ‘plurality of values.’ And, like Hume, he sees our emotions as having a productive role, bringing value into being: the ‘axiological perspective ... does not claim that emotions simply apprehend the world as they find it. A crucial role must be granted to bootstrapping – the invention and creation of value’ (1987, 302). Indeed, de Sousa even (unknowingly?) follows Hume in including wit as an exemplar of this expanded moral conception:

‘Wit’ is standardly suspected of being frivolous; so it is not surprising to find it excluded from the privileged class of ethically relevant focal properties. But this ... begs the question of whether wit is among the intrinsically valuable ends of life. If it is, then wit, together with other ‘personal features,’ far from being excluded from ethical value, must be counted among its foundations. (1987, 308)

We must then reject de Sousa’s claim that ‘not every emotion ... receives Hume’s approval, but only those in a privileged class of moral sentiments’ (1987, 306). Rather, any emotion, so long as it appears alongside a pattern of other feelings and beliefs in such a way as to manifest a virtue, merits our approval. The privileged class of moral sentiments speaks rather to our means of recognizing the morally salient aspects of the world – the ways in which people’s behaviour makes a meaningful difference to us and them.

§4. Conclusion

Hume summarizes his moral theory with what he calls the ‘ultimate test of merit and virtue’:

We may observe in general, that if we can find any quality in a person, which renders him incommodious to those, who live and converse with him, we always allow it to be a fault or blemish, without any farther examination. On the other hand, when we enumerate the good qualities of any person, we always mention those parts of his character, which render him a safe companion, an easy friend, a gentle master, an agreeable husband, or an indulgent father. We consider him with all his relations in society; and love or hate him, according as he affects those, who have any immediate intercourse with him. And 'tis a most certain rule, that if there be no relation of life, in which I cou'd not wish to stand to a particular person, his character must so far be allow'd to be perfect. If he be as little wanting to himself as to others, his character is entirely perfect. (T 3.3.3.9)

Having known de Sousa – Ronnie – for 26 years now, I am confident in my assessment that he is a commodious colleague and an easy friend; I suspect he himself would admit that ‘husband’ was a role he found challenging, and I defer to his daughter, Ava Qingting’s, verdict on his fatherhood, though I know he has been immensely proud of her throughout her life. As to how he stands with respect to himself, Hume admits that these traits can be hard to get a grip on, in that we acquire our understanding of them through sympathetic resonance with the person in question’s own self-assessment, even though we all fall prey to self-deception (T 2.2.4.8).¹⁵ But to me, Ronnie stands out for his comfort with being who he is: brilliant, energetic, full of ideas, and with an unstoppable *joie de vivre*, even as he enters his ninth decade.

References

Ainslie, Donald C. 1999. ‘Scepticism about Persons in Book II of Hume’s *Treatise*,’ *Journal of the History of Philosophy* 37, 469-492.

¹⁵ He elsewhere points to ‘the very difficulty of judging concerning an object, which is never set at a due distance from us, nor is seen in a proper point of view’ (Hume 1757/2007, 2.33).

- Ainslie, Donald C. 2005. 'Sympathy and the Unity of Hume's Idea of Self.' In *Persons and Passions*, edited by Joyce Jenkins, Jennifer Whiting, and Christopher Williams, 143-173. South Bend, IN: University of Notre Dame Press.
- Ainslie, Donald C. 2015. *Hume's True Scepticism*. Oxford: Oxford University Press.
- Allison, Henry. 2008. *Custom and Reason in Hume: A Kantian Reading of the First Book of the Treatise*. Oxford: Clarendon Press.
- Árdal, Páll S. 1989. *Passion and Value in Hume's Treatise*, 2nd edn. Edinburgh: Edinburgh University Press.
- Baier, Annette C. 2008. 'Enquiry concerning the Principles of Morals: Incomparably the Best?' In *A Companion to Hume*, edited by Elizabeth Radcliffe, 293-320. Hoboken, NJ: Wiley.
- Bon de Sousa, Ronald. 1966. 'The Tree of English Bears Bitter Fruit,' *Journal of Philosophy* 63, 37-46.
- de Sousa, Ronald. 1987. *The Rationality of Emotion*. Cambridge, MA: MIT Press.
- de Sousa, Ronald. 2011. *Emotional Truth*. New York: Oxford University Press.
- Hume, David. (1739-40) 2007. *A Treatise of Human Nature: A Critical Edition*, vol 1, edited by David F. Norton and Mary J. Norton. Oxford: Clarendon Press.
- Hume, David. (1751) 1998. *An Enquiry concerning the Principles of Morals: A Critical Edition*, edited by Tom L. Beauchamp. Oxford: Clarendon Press.
- Hume, David. (1757) 2007. *A Dissertation on the Passions; The Natural History of Religion: A Critical Edition*, edited by Tom L. Beauchamp. Oxford: Clarendon Press.
- Hume, David. (1754-61) 1983. *History of England*. Indianapolis: LibertyClassics.
- Hume, David. (1779) 1947. *Dialogues concerning Natural Religion*, edited by Norman Kemp Smith. Indianapolis: Bobbs-Merrill.
- Korsgaard, Christine M. 1999. 'The General Point of View: Love and Moral Approval in Hume's Ethics,' *Hume Studies* 25, 3-42.
- Locke, John. (1700) 1975. *An Essay concerning Human Understanding*, 4th edn, edited by P. H. Nidditch. Oxford: Clarendon Press.
- Mossner, Ernest Campbell. 1943. *The Forgotten Hume, Le Bon David*. New York: Columbia University Press.
- Norton, David F. 2000. 'Editor's Introduction.' In *David Hume: A Treatise of Human Nature*, edited by David F. Norton and Mary Norton, 13-199. Oxford: Oxford University Press.

Penelhum, Terence. 1975. *Hume*. London: Macmillan.

Stroud, Barry. 1977. *Hume*. London: Routledge.

Taylor, Jacqueline A. 2015. *Reflecting Subjects: Passion, Sympathy, and Society in Hume's Philosophy*. Oxford: Oxford University Press.

Waxman, Wayne. 1994. *Hume's Theory of Consciousness*. Cambridge: Cambridge University Press.