

# Hume's Dispositional Account of the Self

## 1. Overview

Hume puzzlingly distinguishes two notions of the self in 'Of Personal Identity' (THN 1.4.6):<sup>1</sup>

...we must distinguish betwixt personal identity, as it regards our thought or imagination, and as it regards our passions or the concern we take in ourselves. The first is our present subject... (THN 1.4.6.5)

And in this view our identity with regard to the passions serves to corroborate that with regard to the imagination, by the making our distant perceptions influence each other, and by giving us a present concern for our past or future pains or pleasures. (THN 1.4.6.19)

Call 'ST' the self 'as it regards our thought or imagination', and 'SP' the self 'as it regards our passions or the concern we take in ourselves'. My focus in this paper is on explicating SP in relation to ST. For my purposes, I will operate with a straightforward understanding of ST as a bundle (c.f. THN 1.4.6.4) or succession (THN 1.4.6.18) of perceptions.<sup>2</sup>

In this paper, I will argue that SP seems subject to two constraints. One, it should be a succession of perceptions (that is, identical to ST), since Hume characterises it as such (THN 2.2.1.2; THN 2.1.2.3). Second, it should have durable constituents in virtue of the roles it plays with regard to pride and humility, as well as normativity. However, I argue that these two constraints are in tension, since our perceptions are too transient to play these roles. As the durability constraint seems more central to Hume's framework in Books 2 and 3, I proceed to characterise SP in light of it, arguing that it is a bundle of dispositions to our perceptions. I then explain why Hume might have mistakenly characterised SP as a succession of perceptions.

---

<sup>1</sup> 'THN' refers to the *Treatise of Human Nature*, 'EHU' to the *Enquiry Concerning Human Understanding*, 'EPM' to the *Enquiry Concerning the Principles of Morals*, and 'EMPL' to *Essays Moral, Political, and Literary*. Arabic numerals refer to section and paragraph numbers (EHU and EPM); to book, part, section, and paragraph numbers (THN); or to page numbers in the Miller edition of the *Essays* (EMPL).

<sup>2</sup> One might take 'bundle of perceptions' to refer to the synchronic self and 'succession of perceptions' to refer to the diachronic self.

## 2. The Perceptions Constraint

Frustratingly, Hume does not enter into an explicit discussion of SP.<sup>3</sup> Hume does argue that the indirect passions of pride and humility are intimately related to a notion of self; this notion is presumably SP, since it ‘regards our passions’ of pride and humility, which embody ‘the concern we take in ourselves’. In this vein, SP also seems to be the morally crucial conception of self. Hume often notes the intimate connection between virtue and pride, as well as vice and humility (THN 2.1.5.4, THN 2.1.7.5, THN 2.1.7.8). Most explicitly:

It follows, that these two particulars are to be consider'd as equivalent, with regard to our mental qualities, *virtue* and the power of producing love or pride, *vice* and the power of producing humility or hatred. (THN 3.3.1.3)

Given the moral valence of hatred, it seems that for Hume, moral judgments regarding vice (and by parallel, virtue) are directed at persons, since hatred is directed at persons (THN 2.3.2.6). Thus, insofar as SP is the notion of self that pride and humility concern, it is correspondingly the notion of self that is involved in our moral judgments.

One question is why Hume does not bother with an explicit discussion of SP if it plays these crucial roles in his philosophy. Two possibilities present themselves: perhaps SP is essentially the same as or very similar to ST, with any differences illuminated by Hume’s explication of the operations of pride and humility; alternatively, perhaps SP is entirely different from ST, but is unanalysable.

The latter position has its defenders. Purviance (1997, p.203) argues that the passions present a ‘moral fact’ of SP, which is ‘necessary for practical activity but not sufficient for theoretical knowledge’; this fact ‘is no longer a philosophical problem’. Lecaldano (2002,

---

<sup>3</sup> This is noted by Penelhum (1976, p.9, n.19), Baier (1979, p.235), and McIntyre (1989, p.545).

p.182) cites THN 2.1.5.3: ‘the peculiar object of pride and humility is determin’d by an original and natural instinct...’. Lecaldano takes this as evidence that SP is original (that is, unanalysable and inexplicable – c.f. THN 1.1.4.6 and EHU 1.2).

However, I think such positions lack textual substantiation. Purviance’s position is unlikely because Hume never acknowledges any ‘moral facts’ which lie beyond the domain of reason and philosophy.<sup>4</sup> On the other hand, Lecaldano misconstrues Hume’s meaning in THN 2.1.5.3; what is original is not SP, but rather pride’s *intentionality* towards the self.<sup>5</sup> HThis paragraph ends: ‘I... consider such a peculiar direction of the thought as an original quality’ (THN 2.1.5.3). Clearly, what is an original quality is the *direction* of the thought (i.e. pride’s intentionality) rather than the *object* of pride (i.e. SP) – see also THN 2.1.3.2 and THN 2.1.3.3.

Thus, SP does not seem to be wholly different from ST. Indeed, it is often argued that the two are very similar, if not identical.<sup>6</sup> The few times Hume mentions the self in the context of pride and humility, he describes it as a succession of perceptions:

This object is self, or that succession of related ideas and impressions, of which we have an intimate memory and consciousness. (THN 2.1.2.2).

But tho’ that connected succession of perceptions, which we call self, be always the object of these two passions... (THN 2.1.2.3)

---

<sup>4</sup> Moral sentiments are outside the domain of reason, but would not constitute ‘facts’.

<sup>5</sup> I say more about this in <omitted>.

<sup>6</sup> See Alanen (2014), Baier (1991, p.130), Carlson (2009), Garrett (1997, p.168), McIntyre (1989), Rawls (2000, p.87), and Schmitter (2009, p.235).

Similarly, Hume's claim that the self is nothing independent of our perceptions (THN 2.2.2.17) suggests the bundle theory.<sup>7</sup> All in all, Hume seems to take SP to be essentially a succession of perceptions, that is, ST.<sup>8,9</sup> Call this the Perceptions Constraint.

A question arises. If Hume thinks that SP and ST are so intimately related, then why would he bother to distinguish them in the first place?<sup>10</sup> Why are there two selves, rather than one? I will briefly return to this question at the end of the paper.

### 3. The Durability Constraint

SP undeniably requires durability. Hume explicitly states that pride and humility (and love and hatred) are only stirred by durable qualities:

What is casual and inconstant gives but little joy, and less pride.... We compare it to ourselves, whose existence is more durable; by which means its inconstancy appears still greater. (THN 2.1.6.7)

If any action be either virtuous or vicious, 'tis only as a sign of some quality or character. It must depend upon durable principles of the mind, which extend over the whole conduct, and enter into the personal character. Actions themselves, not proceeding from any constant principle, have no influence on love or hatred, pride or humility; and consequently are never consider'd in morality....

---

<sup>7</sup> Hume also notes our 'intimate... consciousness' of the self (THN 2.1.2.2); as I argue in <omitted>, Hume's usage of 'conscious' and its cognates often relates to that immediate awareness we have of our perceptions.

<sup>8</sup> One seeming difference between ST and SP is that the latter is sometimes described as encompassing body (e.g. THN 2.1.9.1); Penelhum (1976, p.10) takes Hume to be inconsistent in this respect. In contrast, I think Hume does not explicitly refer to body in THN 1.4.6 because of the sceptical context of this section, which occurs just a few sections after a discussion of scepticism about the external world (THN 1.4.2). Thus, he probably treats body as he does any other external object, that is, as a series of perceptions. In Book 2 Hume is less strict about this issue, and so he sometimes mentions body as distinct from a series of perceptions. Another possibility that an anonymous referee suggests to me is that Hume's notion of the imagination is borrowed from Malebranche (Wright 1983, Kail 2008), and thus involves the effect of the motion of animal spirits in the brain – given that ST is the self 'as it regards our thought or imagination', it might be seen to incorporate some notion of body. For disagreement, see Capaldi (1989, p.168) and Alanen (2014, p.13), who takes body to be crucial and novel to SP.

<sup>9</sup> Also puzzling is why Hume speaks of 'the idea, or rather impression of ourselves' (THN 2.1.11.4), since THN 1.4.6 argues that we have no impression of the self. Garrett (1997, p.169) argues that the idea of a self is an abstract idea, and any impression we have could serve as this abstract idea. Carlson (2009, p.177) argues that the ordinary person's vague idea of themselves as their perceptions might be enlivened to a degree that it approaches the status of an impression.

<sup>10</sup> Thanks to <omitted> for pressing this point.

These [qualities] alone are *durable* enough to affect our sentiments concerning the person. (THN 3.3.1.4-5)

Similarly, Hume argues that moral approbation and blame require durability:

Actions 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... 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. (THN 2.3.2.6)

Thus, fleeting things are not a source of pride for us; for instance, a once-off achievement is derided as a 'flash in the pan', a dismissal implying that it is unworthy of genuine pride.

As the passions of pride and humility, as well as our self-concern, regards our durable traits, SP must therefore involve a conception of ourselves as comprising durable traits, on pain of denying the commonsensical fact that we are often proud or humble of mental traits we possess. When one is proud because of a durable mental quality, one considers oneself as the possessor of this durable mental quality. What is distinctive about SP is that it is the notion of self that our passions are concerned with, and this notion is durable in nature. Call this the Durability Constraint.

As McIntyre (1990, p.200) notes, there is a tension between the Durability Constraint and the Perceptions Constraint; transient perceptions do not seem durable enough to stir pride or humility, nor provide the grounds for normativity. One might reply that although perceptions are not durable, the bundle which they constitute might be. For instance, each second is certainly a transient period of time, but a million seconds is not.<sup>11</sup>

---

<sup>11</sup> Thanks to an anonymous referee for pressing me on this.

However, the durability of the bundle is not sufficient to ground moral attributions, nor to stir pride and humility. We are not simply proud or humble of ourselves in itself, but are proud or humble of ourselves *in virtue of* certain features, as Hume's double relation of impressions and ideas highlights – what is required to stir pride or humility is an object or quality that bears some association with us. Indeed, in THN 2.1.6.7 above, Hume explicitly notes that a transient quality fails to stir pride or humility precisely because we compare it to the greater 'durability' of 'ourselves'. In short, what is crucial is not so much that the *self* is durable (which Hume explicitly recognises to be the case), but that the *qualities* that we are proud or humble of are.<sup>12</sup> In the case whereby we are proud or humble of qualities of ourselves, this requires that we view ourselves as comprising durable qualities. Similarly, our moral evaluations do not simply hold people to be laudable or blameable merely in themselves, but only in virtue of certain character traits, i.e. their virtues and vices. Hume emphasises that the 'quality or character' must be 'durable' (THN 3.3.1.4) for attributions of virtue or vice to occur. Rachel's durability is not in itself sufficient to morally laud her for her generosity; what is further required is that the generosity itself is a durable trait. Again, it is not only SP that must be durable, but also its constituents. A conception of SP as merely a succession of transient perceptions seems to lack the resources for explaining these durable features of the self which are central to normative evaluations, and pride and humility. The constituents of SP must be durable if they are to play the key explanatory role in these respects.

This leaves us with an interpretive puzzle. Can we have a notion of SP that accommodates both the Durability Constraint and Perceptions Constraint, or does one have to give?

---

<sup>12</sup> Of course, if the self comprises durable qualities, the self will correspondingly be durable.

#### 4. A Passionate Account of the Self?

McIntyre's (1990) solution to the above puzzle is to argue for the 'identification of character traits with passions' (p.201), which are taken to be durable (p.200). Thus, we can hold that SP is a succession of perceptions, while still maintaining that some of these perceptions (viz. the passions) possess the requisite durability.

One of McIntyre's arguments for passions being durable is that Hume treats them as more durable than other perceptions (THN 2.3.9.12).<sup>13</sup> I argue that even if passions indeed persist for longer than other impressions, they are nevertheless insufficiently durable to evoke pride and humility, or be subject to moral evaluation. This result generalises to all Humean perceptions. Thus, the Durability Constraint and the Perceptions Constraint are incompatible.

Humean passions uncontroversially possess qualitative character, as they are impressions of reflection (THN 1.1.2.1; THN 1.1.6.1; THN 2.1.1.1). For Hume, impressions are the building blocks of experience, and must therefore be experiential in nature. Moreover, all perceptions are felt to some extent insofar as they have some degree of vivacity, which is undoubtedly an experiential notion.<sup>14</sup> Furthermore, with regard to pride and humility, Hume takes 'their sensations, or the peculiar emotions they excite in the soul' – that is, their qualitative characters – to 'constitute their very being and essence', and moreover notes that 'upon the removal of the pleasure and pain, there is in reality no pride nor humility' (THN 2.1.5.4). This seems to generalise readily to all perceptions given that all perceptions are felt to some extent in virtue of their vivacity. So according to Hume's framework, if a perception is not felt, it does not exist.<sup>15</sup>

---

<sup>13</sup> In response, Pitson (2002, p.90) objects that passions do not have 'genuine duration'.

<sup>14</sup> Loeb (2002) sees vivacity as fundamentally being a dispositional rather than experiential notion. See Marušić (2010) for some excellent criticism of this view.

<sup>15</sup> Stroud (1997, pp.164-5), Smith (1994, p.112) and Shaw (1989) claim that the calm passions are completely insensible; I argue against this in <omitted>.

However, this makes it impossible for passions to have any significant durability, since they are not felt for meaningfully lengthy durations. After all, passions are prone to vary: a broken plate may suddenly dissipate a subtle passion of longing, or a well-timed witticism might drive away melancholic passions and engender good-humoured ones. Moreover, Hume explicitly recognises that when we are in ‘sound sleep’, we are totally devoid of perceptions (THN 1.4.6.3). Thus, the passions, like other perceptions, lack any meaningful duration. I argue that this precludes them from playing the roles requisite to SP.

First, this precludes them from providing a basis for moral attributions. Hume notes that an action which did not proceed from a durable quality ‘leaves nothing of that nature behind it’ (THN 2.3.2.6), and so cannot be something we are held accountable for; the same seems true of a passion when no longer felt. Can we no longer resent Roger for his cruelty when he is asleep, or distracted by a sitcom? Surely we can, and should! But on McIntyre’s account it is difficult to see how Hume could account for this commonsensical fact: Roger’s passion of malice, like his malicious actions, ‘leaves nothing of that nature behind it’ (THN 2.3.2.6) when he is sound asleep, and so we are unable to resent him for this passion at such times.

Second, this transience entails that the passions cannot stir pride and humility. When discussing how pride and humility respond only to durable qualities, Hume notes that in the case of a transient quality,

We compare it to ourselves, whose existence is more durable; by which means its inconstancy appears still greater. It seems ridiculous to infer an excellency in ourselves from an object, which is of so much shorter duration, and attends us during so small a part of our existence. (THN 2.1.6.7).



However, the passions seem to fall foul of this, since the felt experience of a passion is of negligible duration in comparison to the duration of our lives (unlike possessions or mental dispositions). Thus, they are too transient to evoke pride and humility.

Moreover, such an interpretation seems unable to account for Hume's claim that it is 'almost impossible for the mind to change its character in any considerable article' (THN 3.3.4.3). For the passions constantly change: for instance, in describing the interaction of two contrary passions, Hume says that they sometimes 'destroy each other', and sometimes 'remain united in the mind' (THN 2.3.9.13). Moreover, even relatively durable passions are prone to change when we hear a joke, or watch a sad movie, or fall asleep. And this is true of all our perceptions. In contrast, character does not change so easily and often.

In short, a characterisation of SP in terms of passions is unable to adequately accommodate the Durability Constraint.<sup>16</sup> And this objection applies to any straightforward identification of SP with a succession of perceptions, since, like the passions, perceptions must be felt to exist. In short, the Durability Constraint is incompatible with the Perceptions Constraint, and we have to give up one or the other.

In support of her position, McIntyre (1990, p.201) cites THN 2.3.4.1 as claiming that a calm passion can become a 'settled principle of action'. In response, it should be noted that Hume's usage of 'calm passions' is ambiguous between tokens and types. 'Calm passions' sometimes refers to calm passions tokens, such as when Hume speaks of calm passions changing into violent ones (THN 2.3.8.18). Elsewhere, 'calm passions' refers to types of passions which are prone to being calm, for instance when he describes 'benevolence' and 'the

---

<sup>16</sup> A further problem for McIntyre's account is that it struggles to account for natural abilities, which Hume considers 'on the same footing' (THN 3.3.4.1) to moral virtues, but many of which have no obvious relation to the passions. In a recent talk, McIntyre (2015) states that she now believes Humean character to encompass other mental qualities, such as beliefs and natural abilities.

general appetite to good' (THN 2.3.3.8) as calm passions. In THN 2.3.4.1, Hume is quite clearly referring to a passion-type rather than a passion-token: his point is simply that 'repeated custom' from multiple instantiations of the same passion-type have rendered the mind such that any subsequent token of the calm passion in question 'produces no... sensible agitation' (THN 2.3.4.1). Thus, this passage does not establish that calm passion tokens have significant duration, only that such tokens occur in a certain pattern relative to one another.

It is true, as McIntyre (1990, p.200) points out, that virtues and vices such as 'Ambition, avarice, self-love, vanity, friendship, generosity, public spirit' are referred to as passions (EHU 8.7); similarly with benevolence (THN 2.3.3.8). Again, 'excessive pride' is a passion, but can be 'esteem'd vicious' and 'universally hated' (THN 3.3.2.1). I think here Hume is being loose with his terminology in using the same term to describe either a passion or a disposition to the same passion. 'Benevolence' may refer to either an occurrent, transient passion of benevolence, or a durable disposition to instantiate that passion. When Hume describes benevolence as a virtue, he is referring to the disposition to instantiate the passion of benevolence.

Admittedly, we do sometimes loosely describe passions as being hated (e.g. THN 3.3.2.1), in the same way that we sometimes loosely speak of an action being considered vicious or hated, but what is strictly speaking attended with hatred is the underlying durable disposition. Perceptions, like actions, are simply too transient to possess the requisite durability. This conflation also occurs when Hume speaks of moral sentiments being stirred by 'characters and passions' (THN 3.3.1.27). Strictly speaking, 'personal character' (THN 3.3.1.4) and 'characters and disposition' (THN 2.3.2.6) are the only things that are fundamentally considered in moral judgments. However, in the case of dispositions to passions, we infer these dispositions from the occurrent passions, which is why Hume speaks of moral sentiments being

stirred by passions as well as characters. Thus, it seems plausible that when Hume seems to speak of passions constituting our character, he really refers to dispositions to these passions.

## **5. A Dispositional Account of the Self**

Given the arguments in the previous section, the Durability Constraint and Perceptions Constraint are incompatible. Of the two, which should be retained? There are reasons to think that the Durability Constraint is the more fundamental one, being crucial to Hume's framework of pride and humility, as well as normativity, in Books 2 and 3. In contrast, not much seems at stake whether or not SP is a succession of perceptions – if we reject the Perceptions Constraint, we don't seem to have to give up much, if any, of Hume's framework in Books 2 and 3. In this section, I will construct a notion of SP based on the Durability Constraint.

If I am right that when Hume seems to describe passions as virtues, he is really referring to dispositions to these passions, then dispositions seem the natural candidate for constituting SP. First, I will explore the nature of the relevant dispositions; then, I investigate what these dispositions are dispositions *to*.

### **a. The Nature of Dispositions**

Hume's general notion of dispositions clearly cannot be anything like 'psychologically real dispositions with causal powers', as Marušić (2010, pp.162) points out. Any such account is incompatible with Hume's claim that there is no distinction between power and its exercise:

The distinction, which we often make betwixt *power* and the *exercise* of it, is equally without foundation. (THN 1.3.14.34)

...the distinction, which we sometimes make betwixt a *power* and the *exercise* of it, is entirely frivolous, and that neither man nor any other being ought ever to be thought possess of any ability, unless it be exerted and put in action. (THN 2.1.10.4)

Similarly, in a footnote in the *Enquiry* discussing gravitational attraction, Hume treats this disposition as reducible to its effects:

...when we talk of gravity, we mean certain effects, without comprehending that certain power...

(EHU 7.25n)

All this suggests an extremely thin notion of Humean dispositions. If for Hume, dispositions are not anything over and above their effects, it seems a disposition to X is just identical to X's actual occurrences, i.e. its regularly occurring.

There are two problems with such an account in relation to SP. First, such a notion of dispositions does not satisfy the Durability Constraint.<sup>17</sup> It seems that for a disposition to be meaningfully durable, it typically needs to be present even when not manifested. But this seems impossible if the disposition is identified with its effects – if a disposition is literally nothing above and beyond its effects, then it cannot be sufficiently durable in this regard.<sup>18</sup> There is admittedly a looser sense in which such a disposition is durable. If Fred is angry at 10am and 9pm, then on the 'regular occurrence' view of dispositions, his disposition to anger is present at both times. Thus, Fred's disposition to anger occupies multiple time-periods, and is durable in this sense. However, this durability is not sufficient for SP's purposes. Given that dispositions in this sense are *exactly identical* to their effects (any distinction between the two being 'frivolous [THN 2.1.10.4] and 'without foundation' [THN 1.3.14.34]), it follows that they cannot exist when their effects do not exist. It is true that Fred has a disposition to anger

---

<sup>17</sup> Thanks to an anonymous referee for raising this crucial point to me.

<sup>18</sup> Of course, if the effect of the disposition were durable, then the disposition would be durable itself, since the two are identical. I will argue that the relevant effects are perceptions, which are transient.

at 10am and 9pm; however, Fred lacks this disposition at 12pm, 5pm, etc. Yet we wish to be able to morally evaluate him on the basis of his anger even at 12pm, and he might feel humility about this feature in such calmer moments. The ‘regular occurrence’ notion simply cannot account for this.

The second worry is that such a thin notion of dispositions seems unable to support robust counterfactuals. If a disposition to X is just identical to X’s regular occurrence, then we cannot make sense of how something can be disposed to X under merely hypothetical circumstances. This is particularly true with respect to the type of dispositions we are normatively concerned with. For instance, say that Tim lives in a provincial village, and has never encountered anyone from a different race; consequently, Tim has never reflected on his attitudes towards other races. Nevertheless, had Tim encountered any members of a different race, he would have formed discriminatory thoughts and behaved in abusive ways towards them. We might want to say that Tim has a disposition to racism, even if such a disposition was never instantiated due to merely contingent factors. Similarly, if Mary has a genuine deathbed repentance (with associated change in character – c.f. THN 2.3.2.7) but then passes away before having opportunity to exhibit her improved character, she might truly be said to have changed her ways. Indeed, that the dispositions we are concerned with are counterfactual-supporting is illustrated by Hume’s claim that ‘Virtue in rags is still virtue’ (THN 3.3.1.19). We recognise that someone might have a virtuous disposition even if it is never in fact useful or agreeable, because we understand that under typical circumstances, such a disposition would be useful or agreeable.

Marušić (2010, pp.160-1) suggests that Hume’s general account of dispositions takes them to be ‘lawlike patterns of thought and behavior’, which are ‘counterfactual-supporting’; this is a somewhat metaphysically thicker notion of a disposition than the ‘regular occurrence’

notion offered above (since they are counterfactual-supporting), although notably thinner than the ‘psychologically real entities’ notion initially mooted (since they are not reified). As argued above, Hume’s official notion of powers is simply too weak to be either meaningfully durable or counterfactual-supporting. However, Hume nevertheless recognises that the type of dispositions we are normatively concerned with are in line with Marušić’s conception as being both. In short, I think Marušić’s account, while unsuitable as an account of Hume’s official theory of dispositions, correctly describes the notion of dispositions to which our passions respond. Hume notes that our passions respond to a different, richer conception of dispositions than his official one:

But though this [equivalence between powers and their exercise] be true in a just and *philosophical* way of thinking, ’tis certain it is not *the philosophy* of our passions, but that many things operate upon them by means of the idea and supposition of power, independent of its actual exercise. (THN 2.1.10.4)

This notion is as follows:

Since therefore we ascribe a power of performing an action to every one, who has no very powerful motive to forbear it, and refuse it to such as have; it may justly be concluded, that *power* has always a reference to its *exercise*, either actual or probable, and that we consider a person as endow’d with any ability when we find from past experience, that ’tis probable, or at least possible he may exert it. And indeed, as our passions always regard the real existence of objects, and we always judge of this reality from past instances; nothing can be more likely of itself, without any farther reasoning, than *that power consists in the possibility or probability of any action*, as discover’d by experience and the practice of the world. . . . The passions are not only affected by such events as are certain and infallible, but also in an inferior degree by such as are possible and contingent. And tho’ perhaps I never really feel any harm, and discover by the event, that, philosophically speaking, the person never had any power of harming me; since he did not exert any; this prevents not my uneasiness from the preceding uncertainty. (THN 2.1.10.6-7)

Thus, our passions treat someone as possessing a power to perform action X not only when this person has actually exercised this power, but also when this power is only probably or possibly exercised. More generally: to have a disposition to X is simply for X to possibly or probably occur. Given Hume's commitment to determinism, the only way for a possible but non-actual event to have occurred is for different circumstances to have obtained. Thus, these dispositions might also be characterised counterfactually – to have a disposition to X is simply for X to occur under certain suitably circumscribed circumstances.<sup>19</sup> These are the dispositions that constitute SP.<sup>20</sup>

This notion of power succeeds where Hume's philosophical notion failed. First, these dispositions are counterfactual-supporting. To have a disposition to X is simply for X to occur under certain circumstances, which may not be actual. Thus, these dispositions can account for counterfactuals – ways in which the world *might* have been, given different circumstances. Thus, we can consider Tim to have a disposition to racism even if it has never been actualised, because he would have formed racist thoughts and behaved in racist ways under different circumstances.<sup>21</sup>

Second, unlike the philosophical notion of power, such dispositions are sufficiently durable to ground normative attributions. Given that these dispositions can be said to exist when there is probable or possible occurrence, they are clearly present even when

---

<sup>19</sup> The notion of possibility Hume appeals to in THN 2.1.10.6-7 is not anything like metaphysical or even nomic possibility, which would be too weak. What Hume means is rather a loose, everyday sense of 'possible' as having a non-negligible chance of occurring – e.g. 'it is possible that the medication will have side-effects', or 'it is impossible for Liverpool to win the Premiership'. The circumstances pertinent to the relevant counterfactual will correspond to this form of possibility. To say that someone has a disposition to X is to say that X would occur under circumstances that are possible in this sense.

<sup>20</sup> It is true that Hume cannot accommodate such dispositions in his metaphysics. I will return to this issue at the end of the paper.

<sup>21</sup> Indeed, Hume's distinction between the philosophical and passional conceptions of power reveals just how thin the philosophical notion is. Notably, Hume claims in THN 2.1.10.7 that a person who did not *actually* harm me 'philosophically speaking' had no power to do so even if this person *could* have harmed me. This clearly indicates that Hume's official notion of powers is not counterfactual-supporting.

unmanifested, so long as there remains the probability or possibility of their being exercised. If I say that Jane is generous, I am not necessarily claiming that Jane is behaving generously now, but that Jane would possibly display generosity, and this fact will likely still be true tomorrow, and next year, and even beyond.

Indeed, it is exactly this durability of SP enables it to facilitate a present concern with ourselves even in other times:

And in this view our identity with regard to the passions serves to corroborate that with regard to the imagination, by the making our distant perceptions influence each other, and by giving us a present concern for our past or future pains or pleasures. (THN 1.4.6.19)

McIntyre (1989) does much excellent work in explaining how various associative relations (particularly causation and resemblance), acting alongside sympathy in the future case, enable us to feel pride and humility with regard to past and future ‘perceptions, thoughts, and actions’ (p.193). While I agree with her account, conceiving of SP as a bundle of dispositions gives us additional resources to explain this fact. The durability of dispositions means that our past and future characters usually reflect on our present character. While a passion of benevolence from yesterday might be gone today, a disposition to instantiate passions of benevolence is likely to remain for some time; as Hume recognises, character is ‘almost impossible’ to change (THN 3.3.4.3). In short, SP enables a present concern with ourselves at other times via the durability of the dispositions that constitute it.

Thus far, we have provided an account of the dispositions that constitute SP. These dispositions, which Hume takes our passions to respond to, are durable and counterfactual-supporting. From here, we can infer the nature of SP itself. Hume dismisses the very coherence of a substratum when it comes to qualities, and we have no reason to think that he would be



any friendlier towards a notion of substratum for dispositions. Given this, SP seems most likely to be a bundle of dispositions.<sup>22</sup>

### **b. Dispositions to...?**

The next question: what are these dispositions *to*? Character traits are most naturally seen as dispositions to *actions*, or to *perceptions*. Take generosity, for instance. The following two characterisations seem plausible: either generosity is a disposition to generous *actions*, or it is a disposition to the *passion* of generosity. I believe that properly construed, character traits should be seen as dispositions to perceptions rather than actions. For Hume, motivation to behaviour consists in two elements: an end-setting element (the passions), and a means-end element (beliefs). Thus, any disposition towards a certain action will only generate this action by producing certain perceptions (whether passions or beliefs). To examine the case of generosity: generosity is not plausibly seen as affecting behaviour via producing certain beliefs; thus, it must affect behaviour via producing generous passions. Therefore, it seems correct to say that a disposition to generosity is, at the heart of it, a disposition to passions of generosity. Of course, a disposition to a passion can entail a disposition to certain behaviour corresponding to these passions, *ceteris paribus*. When such a disposition produces the relevant passion, this passion will interact with the relevant beliefs, if available, and produce the appropriate behaviour. Nevertheless, the disposition is more fundamentally construed as a disposition to perceptions rather than actions. Moreover, character traits seem to be a wider category than dispositions to action. For instance, the ‘natural abilities’ are included in a person’s ‘character’

---

<sup>22</sup> Purviance (1997, p.208) argues that a bundle theory of character is problematic, because it would face the same problems that ST faces in the Appendix. I do not find this objection compelling, since Hume finds that ‘this difficulty is too hard for my understanding’ (THN App.21). It would be surprising if Hume did have a quick and easy resolution to this difficulty in Book 2; it seems Hume simply brackets the issue and moves on with life. Ainslie (2005, p.152-3) argues that one must give an account of why Hume’s second thoughts in the Appendix are written as directed only at his Book 1 account of the self. I think that this is simply because the problem directly arises from the theses explored in Book 1, which is not to say that he does not think it might have deeper ramifications.

(THN 3.3.4.1), but traits such as a ‘clear head’, or ‘copious invention’, or ‘profound genius’ do not seem plausibly construed as a disposition to actions. Thus, we should see the dispositions that constitute SP as dispositions to perceptions rather than actions.

The next question is this: given that the dispositions that constitute SP are dispositions to perceptions, which perceptions? Pitson (2002, p.90) argues that character traits are dispositions to impressions of reflection, such as the passions. This manages to incorporate traits that are obviously morally relevant, such as benevolence. However, while clearly crucial, I do not believe that such dispositions are sufficient to constitute SP. Rather, the account I intend to defend is that SP is a bundle of dispositions to all our ideas and impressions (i.e. all our perceptions).

If a quality is a virtue or a vice, it seems that it should be considered as part of SP; after all, the subjects of moral evaluation are the persons who possess these qualities (c.f. THN 2.3.2.6). In other words, SP at least encompasses virtues and vices.<sup>23</sup> I will argue that dispositions to all our perceptions can constitute virtues or vices, and so are to be included within SP. For Hume, mental qualities that inspire pride and love are virtues, while those mental qualities that inspire humility and hatred are vices (THN 3.3.1.3). Thus, if a mental quality is capable of producing pride or humility, it should be considered a virtue or vice, and hence should constitute SP. Although dispositions to impressions of reflection are indeed mental qualities that are capable of producing pride, these are not the only such mental qualities.

---

<sup>23</sup> There is a complication that I am skating over. Hume thinks that certain qualities must be possessed to an uncommon degree for us to be proud or humble of them (THN 2.1.6.4). Strictly speaking, it is benevolence *to an uncommon degree* that we are proud of, and that is a virtue. Nevertheless, if benevolence to an uncommon degree is a virtue and is a part of SP, then surely benevolence to even a common degree should also be part of SP.

First, we are often proud of dispositions to instantiate sensory impressions. Consider one's pride in having excellent eyesight, or having discerning taste buds; indeed, Hume speaks of Sancho Panza's pride in his 'great nose' and 'judgment in wine' (EMPL 234). Thus, such dispositions are virtues when exceptional; this means that they are part of SP.

Equally, we can be proud of dispositions to instantiate beliefs, for instance in being proud of my astute causal reasoning, due to my well-functioning belief-forming disposition of custom;<sup>24</sup> or in being proud of my unusual ability to follow complex and difficult chains of demonstrative reasoning. These traits are mentioned in THN 3.3.4.5-6, where Hume discusses how the superiority of reason can be a point of virtue; likewise with natural abilities such as 'a clear head' and 'a sure judgment'. Similarly, Hume speaks of 'the character for judgment and veracity' (EHU 10.25), and castigates a 'blundering understanding' and 'wrong judgment' as defects in character (THN 3.3.1.24); insofar as such belief-forming dispositions are part of our character, they should be included in SP.

Moreover, we can also be proud of dispositions to instantiate non-vivacious ideas, in being proud of one's active and powerful imagination (which facilitates creativity); see Hume's admiration of a powerful and responsive imagination as 'properly what we call genius' (THN 1.1.7.15), or 'a profound genius' and 'a copious invention' (THN 3.3.4.6). Thus, dispositions to instantiate ideas may also be virtues, and are part of SP.<sup>25</sup>

Indeed, in THN 3.3.4 and EPM App.4, Hume's ecumenical conception of virtue is demonstrated by his insistence that there is no material distinction between traditionally considered virtues and natural abilities. Such a distinction is merely verbal ('a dispute of

---

<sup>24</sup> See Endnote H to EHU 9.5, where Hume discusses how people may differ in their powers of causal reasoning.

<sup>25</sup> Hume points out in THN 3.3.4.13 that the memory does not particularly impact our attributions of virtue and vice. Nevertheless, Hume only says that memory 'has the least virtue or vice in its several degrees', and not that it has none: he notes that when unusually good or bad, it still merits praise and blame.

words'), since the two 'are both of them equally mental qualities: And both of them equally produce pleasure; and have of course an equal tendency to procure the love and esteem of mankind' (THN 3.3.4.1). Given this, there seems little reason to exclude these dispositions from our conception of the morally relevant notion of personhood (i.e. SP), since these are all dispositions that our passions are concerned with, and that we care about in ourselves and others.

Since SP involves dispositions to impressions of reflection, impressions of sensation, beliefs, and also non-belief ideas, it seems as though SP is constituted by dispositions to ideas and impressions, which is to say, any perception at all.

## **6. An Understandable Conflation**

This leaves us with a final question. If the above account of SP is correct, then why does Hume nevertheless characterise SP as a succession of perceptions?

We should begin by noting that SP and ST are very closely related: while ST is a bundle of ideas and impressions, SP is a bundle of dispositions to the *very same* ideas and impressions. Despite this, SP is not exactly identical to ST. The dispositions that constitute SP are lawlike patterns of mental states that are durable and counterfactual-supporting. Were the relevant dispositions instead in line with Hume's 'philosophical' conception of powers (THN 1.3.14.34), then SP *would* be identical with ST: there would be no distinction between these dispositions and their effects, and hence no distinction between SP and ST. However, we have seen that Hume thinks our passions respond only to a thicker notion of dispositions, and hence SP must comprise this richer notion of dispositions. Since such dispositions are durable and counterfactual-supporting, a disposition to X is not equivalent to X regularly occurring, and although they are very intimately related, SP and ST cannot be identical.

This closeness between SP and ST explains why Hume mistakenly characterises SP as a succession of perceptions. Hume commits the perfectly understandable mistake of conflating his philosophical notion of power with the thicker notion of power to which our passions respond. As we have seen, the dispositions that constitute SP must be durable and counterfactual-supporting. However, in claiming that SP is a succession of perceptions, Hume confuses these thicker dispositions for the thinner philosophical notion of dispositions. As argued above, if SP were constituted by thin dispositions to perceptions, since there is no distinction ‘betwixt power and the exercise of it’ (THN 1.3.14.34), SP would be metaphysically identical to a succession of perceptions. This conflation is why Hume sometimes characterises SP as a succession of perceptions.

Indeed, that this mistake is understandable can be seen from the extent of the closeness between these two conceptions of self. ST, while not strictly identical to SP, is nevertheless intimately related to it: ST is simply the *manifestation* of SP, the way in which SP takes form. On the whole, this might explain why Hume does not bother with an explicit discussion of SP, despite it being distinct from ST – given the closeness between the two, Hume mistakenly took them to be identical, and perhaps thought that spelling out the nature of SP would be superfluous.

Nevertheless, my account also explains why we need two notions of self, rather than one. Despite the two being so intimately related, SP plays a role that ST could not. SP – that is, the self considered as a bundle of dispositions – allows for our passions of love and hatred, pride and humility to get a foothold, and for the normative approbation and blame of persons to occur. ST – the transient manifestation of these dispositions as a succession of perceptions – cannot facilitate this passionate reaction, since it is comprised of elements that are insufficiently durable to do so. A pre-condition of our passional and normative worldview that

Hume explores in Books 2 and 3 is for us to conceive of both ourselves and others as durable beings, in possession of durable dispositions. In this, SP is the conception of self that is, unlike ST, fundamentally social in nature – it grounds the socially vital passions of love and hatred, and moral approbation and blame. These roles that SP plays in Hume’s framework explains why a distinction between the two is necessary in the first place, even though the latter is merely a manifestation of the former. Conversely, given that SP relies on a philosophically problematic conception of dispositions, ST plays a role that SP cannot, as the philosophically correct (but non-normative) notion of the self.

However, even if we require two conceptions of the self, it is still unclear why Hume would have distinguished them, given that he mistakenly identifies SP and ST.<sup>26</sup> The answer to this will necessarily be speculative. Perhaps, although Hume thought that SP and ST were identical, he might have conceived the two as different modes of presentation of one and the same self – in short, the distinction between SP and ST might be seen as a distinction of reason. On such a conception, the following picture emerges. Under the ST mode of presentation – the self considered *qua* perceptions – we have a philosophical explanation for the metaphysics of the self. Meanwhile, under the SP mode of presentation – the self considered *qua* (thin) dispositions – we have the basis of our normative worldview, as a source of pride and humility, love and hatred. Although Hume might believe the two selves to be identical at the heart of it, nevertheless, a distinction between them might be thought edifying in light of the differing conceptual roles each mode of presentation is expected to play. Of course, as argued above, Hume was wrong to think that one and the same notion of self could play both roles; perhaps in believing a distinction of reason necessary here, he was grasping at his mistake.

---

<sup>26</sup> Thanks to <omitted> for pressing me on this point.

There is an implication of my interpretation that I will briefly note here. SP can only be a fictional notion, since the dispositions that constitute it do not exist, ‘philosophically speaking’ (THN 2.1.10.7). Similarly, he notes that the thicker notion of dispositions relevant to SP does not accord with the ‘just and *philosophical* way of thinking’ (THN 2.1.10.4). The problem is that Hume’s regularity account of causation is an austere one, and cannot sharply distinguish between counterfactual-supporting regularities and mere regularities; this explains the thinness of his official account of power.<sup>27</sup> Thus, the durable and counterfactual-supporting dispositions that constitute SP do not strictly speaking exist – a strict metaphysical perspective of the world has no room for such notions in Hume’s book.

Nevertheless, this strict metaphysical perspective does not exhaust our worldview. We have to view people as possessing such dispositions in order to feel pride and humility, love and hatred. In short, SP is something akin to a necessary presupposition for a normative worldview; a fictitious presupposition, perhaps, but one that our social natures nevertheless compels. Does this necessitate an error theory of morality? I think not – just as one might correctly make causal judgments even if one mistakenly thinks necessary connection to exist in objects, one might still correctly make moral judgments even if they rest on a philosophically erroneous conception of the self. But to discuss this matter in fullness would take me beyond the scope of this paper.

---

<sup>27</sup> Thanks to an anonymous referee for discussion on this.

## **References**

- Ainslie, Donald. (2005). "Sympathy and the Unity of Hume's Idea of Self", in *Persons and Passions: Essays in Honour of Annette Baier*, J. Jenkins, J. Whiting, and C. Williams (eds.), University of Notre Dame Press: Notre Dame Indiana, 143-173.
- Alanen, Lilli. (2014). "Personal Identity, Passions, and 'The True Idea of the Human Mind'", *Hume Studies* 40(1): 3-28.
- Baier, Annette. (1979). "Hume on Heaps and Bundles", *American Philosophical Quarterly* 16:4.
- Capaldi, Nicholas. (1989). *Hume's Place in Moral Philosophy*, New York: Peter Lang.
- Carlson, Åsa. (2009). "There is Just One Idea of Self in Hume's *Treatise*", *Hume Studies* 35.
- Garrett, Don. (1997). *Cognition and Commitment in Hume's Philosophy*, Oxford/New York: Oxford University Press.
- Hume, David. (1739-1740). *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, 2nd ed., revised by P.H. Nidditch. Oxford: Clarendon Press, 1975.
- Hume, David. (1748). *Enquiries concerning Human Understanding and concerning the Principles of Morals*, edited by L. A. Selby-Bigge, 3rd edition revised by P. H. Nidditch, Oxford: Clarendon Press, 1975.
- Hume, David. (1777). *Essays Moral, Political and Literary*, revised edition edited by E. Miller, Liberty Fund Inc., 1985.
- Kail, Peter. (2008). "On Hume's Appropriation of Malebranche: Causation and Self", *European Journal of Philosophy* 16(1): 55-80.



- Lecaldano, Eugenio. (2002). "The Passions, Character, and the Self in Hume", *Hume Studies* 28:2.
- Loeb, Louis E, 2002. *Stability and Justification in Hume's Treatise*, Oxford University Press.
- Marušić, Jennifer S. (2010). "Does Hume Hold a Dispositional Account of Belief?", *Canadian Journal of Philosophy* 40:2.
- McIntyre, Jane. (1989). "Personal Identity and the Passions", *Journal of the History of Philosophy* 27:4.
- McIntyre, Jane. (1990). "Character: A Humean Account", *History of Philosophy Quarterly* 7:2.
- McIntyre, Jane. (2015). "'Repentance wipes off every crime': Character, the self and moral responsibility in Hume's philosophy", presented at the Hume Society Conference, Stockholm.
- Penelhum, Terence. (1976) "The Self in Hume's Philosophy", in *David Hume: Many-Sided Genius*, K. Merrill and R. Strachan (eds.), University of Oklahoma Press.
- Pitson, Tony. (2002). *Hume's Philosophy of the Self*, Routledge.
- Purviance, Susan. (1997). "The Moral Self and the Indirect Passions", *Hume Studies* 28:2.
- Rawls, John. (2000). *Lectures on the History of Moral Philosophy*, Harvard University Press.
- Schmitter, Amy M. (2009). "Making an Object of Yourself: On the Intentionality of the Passions in Hume", in *Topics in Early Modern Philosophy of Mind*, J. Miller (ed.), Springer.
- Wright, John. (1983). *The Sceptical Realism of David Hume*, Manchester: Manchester University Press.

