

How realist was Hume's self?: A critique of Kristjánsson on Hume

Andrew OBERG¹⁾

(Received : September 29, 2016, Accepted : December 15, 2016)

ヒュームの自己はどこまで現実主義的であったか

—クリスティアンソンのヒューム解釈へのクリティーカー—

オバーク アンドリュウ¹⁾

(2016年9月29日受付, 2016年12月15日受理)

Abstract

The question of the self, of whether or not there is a core part to us that forms our inner essence or absolute nature, has been with us almost since our beginning. For centuries philosophical arguments over the self took the form of discussing the nature of the self for its existence was taken to be a given. This assumption has been increasingly called into question, however, resulting in the current climate in which the absence of a self is presumed. Contrary to this trend Kristján Kristjánsson has recently proposed a realist self that allegedly rests upon the emotions, claiming a Humean foundation for his account. In the following that claim is called into question through a close examination of Hume's approach to the self in his *A Treatise of Human Nature* (1739), and an interpretation of Hume's work that differs both from Kristjánsson's reading and from the traditional reading is offered.

Key words: anti-realist/realist, David Hume, Kristján Kristjánsson, personal identity, the self,
whole person

要 旨

自己への問い—つまり、我々の内なる本性もしくは絶対的性質を形成する核心部分が我々にあるか否か—は、人類とともに始まったと言える。何世紀にもわたって、自己に関する哲学の議論は、自己が存在することを前提に行われ、自己の性質に関する議論という形をとった。ところが、この前提がますます疑問視された結果、現今の時勢において、自己の不在が推定されるのに至っている。クリスティアン・クリスティアンソンはこの潮流に対して感情に基づくとされる現実的自己を提言し、自らの論の基礎をヒュームに求める。この論文では、ヒューム著『人間本性論』(1739年)における自己へのアプローチを詳細に検討することによって、その主張に疑問が投げかけられ、また、クリスティアンソンの見方とも伝統的なそれとも異なる解釈が提示される。

キーワード：非実在論的／実在論的、ヒューム、クリスティアンソン、人格的同一性、自己、全人

1) 高知県立大学文化学部 講師
Faculty of Cultural Studies, University of Kochi, Assistant Professor

1. The self, personal identity, and whole persons

Traditionally the self has been considered to be one's inner core, the true essence of what one is and, typically, immortal in one sense or another, the body's commanding unit or general headquarters that is nevertheless not so strongly associated with the body that it ceases upon the body's death; rather instead it continues on in some manner. This view has understandably had a bumpy ride in the modern era, particularly within academic circles in which the currently predominant view of the self is to deny that it exists and then from there to attempt to explain (or explain away) that inner sense of a 'me' with which we are all so familiar and which had previously simply been labeled one's soul. One influential self-concept in this non-realist vein is Derek Parfit's neo-Lockean account, which gives us the self as being one's psychological connectedness and/or continuance,¹ another is Barry Dainton's account which posits the self as our streams of consciousness,² or Galen Strawson's account which does give the self substance in an almost realist way but only for the duration of an experience, meaning that each new experience generates a whole new self.³ On all of these accounts a more long-lasting, truly realist self is seen as being more or less an illusion. In some ways each of these concepts also mirror the standard Buddhist arguments against the existence of a soul⁴ though it must be noted that in the Western philosophical tradition arguments against a realist self have since the time of René Descartes (1596-1650) been aimed more at the Cartesian ego than at the traditional soul.

Accordingly, it will be important here, before continuing on to the main topic of our discussion, to briefly delineate the Cartesian account from traditional ideas about the soul for the two are often conflated. What Descartes accomplished was to change the emphasis of traditional ideas by shifting the focus from the soul (and much that that term implied in traditional thinking) to the 'mind' while still maintaining the overall idea of a pilot in a vessel, or, in more modern terms, a driver in a vehicle or (stretching it) a ghost in a machine. Descartes' view itself was considerably more complex than this, however;⁵ to him one's true self was the thinking thing, a thing which must be substantive as the acts of thought have to be based in something, and that substance was an intellectual substance, the pure intellect.⁶ Descartes 'equated *soul* with *mind* and preferred the latter term to avoid ambiguity.'⁷ In Cartesian thinking then the mind and body are wholly separate entities but clearly interact as is evidenced by sensations like pain or hunger being felt by 'me' (the mind) and not merely observed by 'me'. The nature of this interaction, the manner in which the two communicated, and how the

¹ Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984).

² Barry Dainton, *The Phenomenal Self* (Oxford: Oxford University Press, 2008).

³ Galen Strawson, *Selves: An Essay in Revisionary Metaphysics* (Oxford: Oxford University Press, 2009).

⁴ For some examples see Glenn Wallis, trans. and comm., *The Dhammapada: Verses on the Way* (New York: The Modern Library, 2007); Asanga Tilakaratne, *Theravada Buddhism: The View of the Elders* (Honolulu: University of Hawai'i Press, 2012); Steve Hagen, *Buddhism: Plain and Simple* (New York: Broadway Books, 1999); and Alan Watts, *The Way of Zen* (New York: Vintage Books, 1989).

⁵ Descartes was the first major thinker to use the term 'mind' instead of 'soul' and saw the relation between the mind and body not exactly as a pilot and a vessel but one where 'they systematically affect each other, but not other things, in ways that make the two of them together function as if they were one.'

John Barresi and Raymond Martin, 'History as Prologue: Western Theories of the Self' in *The Oxford Handbook of the Self*, ed. by Shaun Gallagher (Oxford: Oxford University Press, 2011), pp. 33-56 (p. 39).

⁶ Gary Hatfield, *The Routledge Guidebook to Descartes' Meditations* (Abingdon, OX: Routledge, 2014).

⁷ *ibid.*, p. 124.

mind was able to move the body, is not clear in Descartes' account, and he was unable to adequately answer objections of this kind that were raised by contemporaries.⁸ It is also somewhat unclear on the Cartesian account what the nature of the intellectual substance is for it is arguable that, although immaterial, it is not necessarily immortal in the way a traditional soul has been thought to be. Does a Cartesian mind travel postmortem in the way that a soul does in its journey through the otherworld and/or back into this world again? Or does a Cartesian mind rather stop existing when the human body with which it was associated stops existing? Descartes himself wrote in such a way that makes the reader think he favored the former (though not, of course, the reincarnation option), although given the pains with which he sought to avoid displeasing the Church and thereby inviting harm to his person (given the historical context in which he wrote with the Inquisition still active) this may have been simply a means of protecting himself.⁹ The case therefore is, and will likely remain, an open one, but for our purposes here the distinction between the Cartesian account and the traditional one should be sufficiently clear.

How then ought we to understand the terms that form the current section's heading? How does a 'self' differ from one's 'personal identity' and how does each of them differ from a 'whole person'? The answer is largely a matter of degree. The self in its realist sense is a mental entity or psychological datum that defines and informs a person's 'true nature' or 'core identity', rooting all other psychological aspects. Anti-realist or non-realist accounts will on the other hand often use the terms 'self' and 'personal identity' interchangeably, illustrating the purported way in which the brain's standard functioning gives rise to the *notion* of a self (some cognitive accounts have also suggested that the self may be rooted in the brain in a modular or functional manner¹⁰) but that such is merely that, an idea and nothing substantively further. Since the self does not exist on such accounts what appears to be the self may be discussed in the same manner as all other 'merely' psychological phenomena that play into one's sense of identity. Personal identity then is either, on realist accounts, the self plus other psychological conditions (including one's personality, traits, characteristics, goals, etc.), or, on anti-realist accounts, only those same psychological conditions. The whole person, in contrast, consists of personal identity (with or without the self) in addition to the kind of embodied issues that are raised by our physical features and physical aspects of being. To put it simply, if there is a self then it anchors personal identity and both of them plus the body yields the whole person; if there is no self then only personal identity plus the body yields the whole person. With these definitions and preliminary discussions in mind it should be clear the kind of background image that is raised when one speaks of a 'self', realist or not, and we may therefore turn now to the specific self-concept and its alleged foundation that this paper wishes to consider.

⁸ *ibid.*

⁹ Hatfield, *op. cit.*; on some thoughts regarding what else Descartes may have meant see Strawson, *op. cit.*

¹⁰ Readers interested in pursuing this avenue may wish to see António Damásio, *Descartes' Error: Emotion, Reason, and the Human Brain* (New York: Putnam, 1994) or Daniel C. Dennett, *Consciousness Explained* (New York: Little, Brown and Co., 1991) to get started; Strawson, *op. cit.*, also discusses cognitive theories, labeling them 'brain-system' views.

2. Kristján Kristjánsson's self-concept and claim to its Humean basis

What is novel about Kristján Kristjánsson's account, particularly given the current academic climate, is that it is realist about the self – that it claims that there is in fact a substantive nature to selfhood without further making either the Cartesian ego or the traditional soul claims that we tend to associate with the realist stance. Kristjánsson finds the basis for his account in the work of David Hume, and although he admits that there is no consensus on the definitive way to interpret Book Two of Hume's *A Treatise of Human Nature*,¹¹ Kristjánsson suggests that Hume 'seems to be arguing that, whereas the self as a succession of related ideas and impressions cannot be a direct object for the understanding, the self of whose moral actions each of us is intimately conscious can be a direct object for our emotions'.¹² Such a stance makes the self's realism – the sense in which it can be argued to be 'real' – to consist of emotional activity and to thereby be generative of a self-concept that is dependent on reinforcement from others via social interactions and societal rules and conventions regarding emotions. Put more simply, how we feel, how we are taught to feel, and how others make us feel all inform what is our substantive 'self'. Although it is not entirely clear from his account just what Kristjánsson is driving at, he appears to be using Hume's idea as a means of moving away from the Lockean concept of a self-account that is based on the continuity of consciousness and conscious access to memory (or at least the possibility of being able to consciously access memory (thus accounting for the breaks that we are experientially familiar with such as forgetfulness, sleep, and the like))¹³ and instead spring boarding to the alternative concept he wishes to promote that has the self as more properly based on one's shifting emotions, subject to both personal reflection and societal support/censure, and using this interpretation of Hume to justify or ground the account. The self here, on this alternative account, is each person's moral being, the day-to-day psychological unit of reference that we go by, akin to 'the voter' or 'the citizen' or 'the taxpayer'; it is seeing oneself from an affective and morally related point of view.¹⁴ Kristjánsson summarizes the self on this account as being composed of three sets of self-related emotions: 1) Self-constituting emotions: those that define us, our 'core commitments, traits, aspirations or ideals', 2) Self-comparative emotions: those that take the self as 'an indirect object' or 'a reference point' for 'comparison with a baseline of expectations', and 3) Self-conscious emotions: those that are in the self they're about, that take the self as 'their direct attentional and intentional object'.¹⁵ The self here is that which we maintain as a personal psychological point of reference and that more or less matches what we mean when we speak of ourselves, it is the culmination of the creature performing the actions, thinking the thoughts, and having the feelings that we internally associate with those actions, thoughts, and feelings, and that others associate with them too. On this account there is no mental 'pure ego' that takes up residence

¹¹ David Hume, *A Treatise of Human Nature*, intro. and ed. by David Fate Norton and Mary J. Norton (Oxford: Oxford University Press, 1739/2000).

¹² Kristján Kristjánsson, *The Self and Its Emotions* (Cambridge: Cambridge University Press, 2010), pp. 47-48.

¹³ See Chapter 27, especially sections 11-12, 'Of Identity and Diversity' in John Locke, *An Essay Concerning Human Understanding*, 2nd edn (1689/1690). Available on Project Gutenberg: <<http://www.gutenberg.org/cache/epub/10615/pg10615.html>>. Accessed 08 August 2016.

¹⁴ Kristjánsson, *op. cit.*

¹⁵ *ibid.*, pp. 75-77.

in the physical body (the 'psychological unit' Kristjánsson refers to should be understood almost as a side effect instead of as the core that the 'pure ego' would be; note his comparisons to such definitional/functional terms as 'the citizen', 'the taxpayer'). Rather the physical body, along with our unique sense of personal identity, is all that the self is and part of that body's normal functioning is to have the emotion-based ongoing characteristics that make up Kristjánsson's first emotive set above and the self-referencing features of his second and third sets. Of note on this account is that our self-concept as a moral being is particularly subject to the social; it can correspond to reality or fail to correspond, and it can be judged objectively through the lens of its quality as other-dependent in the public realm. Our daily interactions will either reinforce the way in which we see ourselves as being or they will show us where we have been in error.

I think there are a number of problems with this account – and a number of meritorious points as well –, but rather than dwell on each of them here (I have done so elsewhere¹⁶) I wish instead to focus purely on Kristjánsson's interpretation of Hume as that is a foundational point for the self-concept as proffered and because I think that what I have to say on the matter may be different enough from both Kristjánsson's and from the traditional reading of Hume to perhaps be of some small interest. The paper will conclude with some avenues opened by Kristjánsson's account that may nevertheless be worth pursuing.

3. Rereading Hume

The traditional interpretation of Hume has him as arguing that the self is nothing more than a constantly varying bundle of experiential input whose nature is bound up with one's reflections on one's states.¹⁷ Kristjánsson disagrees with this, as outlined in the preceding, and I, in turn, disagree with Kristjánsson's reading of Hume. My own interpretation of the way that Hume presents the self understands his view as essentially equating the self with personal identity, which would not allow the self to have the deep realist foundation that Kristjánsson gives it and which does agree with the focus of the traditional reading of the Humean account as essentially being an anti-realist one though it goes further in the thickness (or richness) that it gives to the Humean self-concept. (One's personal identity being more substantive than a 'bundle of experiential input' would allow; see Section 1 above.) In this section I will make my case for this reading of Hume by going back to the end of Book One in his *A Treatise of Human Nature*, in which personal identity is addressed directly (in 1:4:6) and then looking at all the places in Book Two where Hume refers to the self to try and understand what he meant by that. Prior to examining Hume's quotes though we will need to give a brief background on his notion of the 'double relation of ideas and impressions' to see where Kristjánsson likely got the purported Humean basis for his emotional self as described above.

¹⁶ See Andrew Oberg, 'A Realist Self?', *Journal of Applied Ethics and Philosophy*, 7 (2015), 24-33; there are points that I would change about this paper but the overall critique of Kristjánsson's self-concept found therein remains more or less in place in my mind.

¹⁷ Brie Gertler, 'Self-Knowledge', *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), ed. by Edward N. Zalta. <<http://plato.stanford.edu/archives/sum2015/entries/self-knowledge/>>. Accessed 08 August 2016

Hume's system in Book Two depends on the aforementioned 'double relation of ideas and impressions' in which an idea of something (Hume gives the example of one's own house) produces an impression (pleasure, assuming it is a nice house) but also correlates to another idea (my self, the owner of the house) that is connected with another impression (the pleasurable feeling of pride in owning something nice); this latter impression, it will be noted, resembles the earlier impression. The association of ideas here grants to one the idea of one's self: 'That is, the subject brings to mind the idea of precisely that individual who is always the object of pride.'¹⁸ It is probable that Kristjánsson took his notion of an emotionally generated self from this relationship, as he specifically references Book Two of *A Treatise of Human Nature* when he writes that Hume 'seems to be arguing that, whereas the self as a succession of related ideas and impressions cannot be a direct object for the understanding, the self of whose moral actions each of us is intimately conscious can be a direct object for our emotions.'¹⁹ We must tread lightly here though as what Hume is presenting in his 'double relation of ideas and impressions' is not a theory of self but rather the common element to be found between a feeling and the various sources of that feeling, linked also to certain ideas (of which the self is one) via association. I believe that the standard anti-realist reading of what Hume understood the self definitionally to be is closer to what we can find in Hume than Kristjánsson's reading is.

To see that we turn to what Hume actually wrote. In Book One Part Four Section Six (1:4:6 'Of personal identity'), he displays a skeptical general view in the following: 'For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other ... When my perceptions are remov'd for any time, as by sound sleep; so long am I insensible of *myself*, and may truly be said not to exist.'²⁰ To this he adds, 'they [i.e. 'we', people] are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement.'²¹ Here we have Hume's famous 'bundles' concept, and with his foundation laid he summarizes his account thusly:

identity is nothing really belonging to these different perceptions, and uniting them together; but is merely a quality, which we attribute to them, because of the union of their ideas in the imagination, when we reflect upon them ... 'Tis, therefore, on some of those three relations of resemblance, contiguity and causation, that identity depends; and as the very essence of these relations consists in their producing an easy transition of ideas; it follows, that our notions of personal identity, proceed entirely from the smooth and uninterrupted progress of the thought along a train of connected ideas, according to the

¹⁸ Hume, *op. cit.*, Editor's Introduction p. 151.

¹⁹ Kristjánsson, *op. cit.*, pp. 47-48.

²⁰ Hume, *op. cit.*, Text p. 165.

²¹ *ibid.*, Text p. 165.

²² *ibid.*, Text pp. 169-170.

principles above-explain'd.²²

What Hume appears to be giving us here is an anti-realist account that equates the self (or, rather, sense of self) with personal identity and being dependent on psychological continuity; note that without perceptions (i.e. psychological stimuli) one 'may truly be said not to exist'. Personal identity then, by its nature, makes us think that there is a self. Furthermore, on Hume's account, this mistaken supposition is additionally supported by memory which discovers and helps produce our identity, aided by our emotions working with our imagination in linking present concerns with past and future,²³ and by the idea of causation which allows us to take identity beyond memory and suppose the same identity even during forgotten times.²⁴ Having established sufficient doubt in any realist concept of the self (he supposes), and having described what personal identity really is and why it erroneously makes us think there is a self resting beyond it, Hume concludes his thoughts with this gem of a statement: 'all the nice and subtile questions concerning personal identity can never possibly be decided, and are to be regarded rather as grammatical than as philosophical difficulties.'²⁵ We, however, may still have our doubts about just what Hume actually meant and so for the moment I ask only that we keep the preceding in mind as we turn to Book Two to see how Hume wrote of the self there and to try to ascertain if it is at heart an anti-realist account. If my reading of Hume is right and he does give to personal identity the weight that Kristjánsson wants a realist self to carry, then the self view on Hume's account – on its own based on what Hume writes about it – may not be able to be 'a direct object for our emotions' as Kristjánsson has it. However, if we think that *personal identity* could be a direct object for our emotions, and I do not see why it could not, Kristjánsson's core idea does still stand, but there is now within it a fatal flaw: although personal identity can (conceivably) be an object for one's emotions it would not thereby generate a self able to underpin personal identity. This is important because, as remarked in the first section above, the whole point of a realist self is to argue that the self is an entity or function (depending on how realist one wishes to be) that is more foundational and that ensures personal identity continues even when psychological continuity (/perceptions/memory) does not. Such reasoning would not, though, in itself mean that the self is necessarily non-existent (that is, anti-realist), nor would it mean that Kristjánsson's self-concept is without value, it would only mean that Kristjánsson is mistaken in what he attributes to Hume. If reading Hume did give Kristjánsson his ideas about how a realist self could be established then certainly Hume does deserve the credit given him, I do not mean to object to that, perhaps just not in the way that Kristjánsson seems to be saying.

The first mention of the self in Book Two comes in Part One Section Two (2:1:2 'Of pride and humility; their objects and causes') where Hume gives us a definition of the self that seemingly just substitutes the term 'self' for 'personal identity' as he previously summarized it: 'Tis evident, that

²³ *ibid.*, Text p. 170.

²⁴ *ibid.*, Text pp. 170-171.

²⁵ *ibid.*, Text p. 171.

pride and humility, tho' directly contrary, have yet the same OBJECT. This object is self, or that succession of related ideas and impressions, of which we have an intimate memory and consciousness. Here the view always fixes when we are actuated by either of these passions.'²⁶ This is followed by a statement that shifts the focus somewhat, moving back to perceptions (and not 'ideas and impressions') that closely matches Hume's 'bundles' line quoted above: 'But tho' that connected succession of perceptions, which we call *self*, be always the object of these two passions, 'tis impossible it can be their CAUSE, or be sufficient alone to excite them.'²⁷ Following on along these lines of distinguishing between object and cause we have this thought, which appears to be the root of Kristjánsson's reading of Hume:

Pride and humility, being once rais'd, immediately turn our attention to ourself, and regard that as their ultimate and final object; but there is something farther requisite in order to raise them: Something, which is peculiar to one of the passions, and produces not both in the very same degree. The first idea, that is presented to the mind, is that of the cause or productive principle. This excites the passion, connected with it; and that passion, when excited, turns our view to another idea, which is that of self. Here then is a passion plac'd betwixt two ideas, of which the one produces it, and the other is produc'd by it. The first idea, therefore, represents the *cause*, the second the *object* of the passion.²⁸

This seems clear enough; Hume here has the self as the object of the passion (the emotion) of pride or humility which was caused by an idea (say, of one's house) and then via the related ideas raised by that rests on the idea of the self. Yet the self being described is just that, an idea, incapable of being a cause in itself, and given what Hume has previously written now reading a realist conception into this may not be entirely warranted. Interestingly, though, this passage does seem to suggest that the self (understood as personal identity) may not necessarily be tied to inner reflection as Hume is usually read to mean. It could be that Hume considered the passions to play into one's personal identity even when one did not thereafter reflect all the way out to that personal identity (that is, down the chain of related ideas). If so then Kristjánsson's alternative interpretation of Hume might have some merit to it though not, perhaps, to the extent that Kristjánsson claims. We continue our exploration.

Part One Section Five (2 : 1 : 5 'Of the influence of these relations on pride and humility') tells us that "tis absolutely impossible, from the primary constitution of the mind, that these passions [pride and humility] shou'd ever look beyond self, or that individual person, of whose actions and sentiments each of us is intimately conscious.'²⁹ The self now appears to be something more akin to a whole

²⁶ *ibid.*, Text p. 182.

²⁷ *ibid.*, Text p. 182.

²⁸ *ibid.*, Text p. 182.

²⁹ *ibid.*, Text p. 187.

³⁰ Quoted above: '[Hume] seems to be arguing that, whereas the self as a succession of related ideas and impressions cannot be a direct object for the understanding, the self of whose moral actions each of us is intimately conscious can be a direct object for our emotions', Kristjánsson, *op. cit.*, pp. 47-48.

person, and we note too that this is the statement referenced by Kristjánsson when he gives the basis for his self-concept as Book Two of *A Treatise of Human Nature*.³⁰ Then a little later in the same section Hume describes the self as an idea attached to pride, a 'contrivance of nature' like lust or hunger and their objects.³¹ This is not terribly helpful as the objects of lust and hunger are certainly very real (in contrast to how Hume has earlier described the self), yet the thrust here seems to be the notion of an idea necessarily attached to a passion, that you cannot have one without the other. The picture is now murkier as we find in Part One Sections Six and Nine Hume making two further rather cryptic remarks about the self; in 2:1:6 ('Limitations of this system'), he contrasts pride with joy, the former having two objects necessary to its production (the cause of pride and the self) and the latter having only one (the cause), although he admits that joy must have 'some relation to self' but such is only needed to 'render it agreeable; nor is self, properly speaking, the object of this passion.'³² This seems to hint that joy can exist in the absence of a self, but if Hume's account is anyway non-realist about the self then such would be perfectly possible while maintaining that the joy must be enjoyed by some living creature and not merely some thing (even though that enjoyer, the object, is not required for joy's existence by this account). In 2:1:9 ('Of external advantages and disadvantages') Hume states that, 'But tho' pride and humility have the qualities of our mind and body, that is *self*, for their natural and more immediate causes'³³ (he then notes that there are other causes which obscure the primary), suddenly seeming to introduce the body into the concept of the self, which, if it was important for Hume, would be something else entirely from his earlier 'bundle' of perceptions or series of connected ideas explanations. This is an isolated comment though and is not repeated, whereas in Part Two Section One (2:2:1 'Of the objects and causes of love and hatred') Hume does repeat an earlier self definition, from 2:1:5, writing that, 'As the *object* of pride and humility is self or that identical person, of whose thoughts, actions, and sensations we are intimately conscious.'³⁴ Here is our apparent whole person again; just what are we to make of all this? Taken with Hume's earlier description of what personal identity is (an attribution we assign to our continuing perceptions (see the first long quote above)) and how our perceptions generate the sense of self (again, that same attribution given to perceptions: personal identity = (sense of) self), which when absent ('as by sound sleep') leads to a condition equated with non-being (see the first quote from 1:4:6 above), gives a picture of the self that is either an illusion generated by perceptions (= personal identity)/the flow of other ideas and impressions relating to the person who is their object, or is simply the whole person 'of whose thoughts, actions, and sensations we are intimately conscious' (but then who is the 'we' in that?). Kristjánsson seems to opt for the latter, perhaps considering the referenced 'we' to be his emotionally-based realist self (that is, what it would really be; Hume of course did not have access to the account) although it is questionable if such would work in the sense carried by this 'we' given the picture of

³¹ Hume, *op. cit.*, Text p. 188.

³² *ibid.*, Text p. 191.

³³ *ibid.*, Text p. 198.

³⁴ *ibid.*, Text p. 214.

the self as a day-to-day psychological unit that Kristjánsson describes it as being. My own view is that judging just from the space Hume gives to each the whole person notion seems much less important to him (if indeed it was a separate concept and not merely a case of incautious writing³⁵), and in both senses of the self as an illusion (via perceptions or via the train of other ideas) the view of the self is anti-realist and tied to psychological continuity. A realist self would require more.

I should stress that this is only my reading of Hume and that it is difficult to know exactly what he meant, but when what he does write about personal identity and the self is laid out in this way it seems clear that Kristjánsson's interpretation is at least weaker than he claims. Granted, even if Hume's self is an anti-realist one, as the traditional reading (and as my own) has it, the sense of self that develops in the way Hume describes could still be an object for emotional activity, and that it could be so does lend credence to Kristjánsson's three self-related emotional sets. It is here, in fact, that my reading of Hume parts ways very slightly with the traditional interpretation as I can see how Hume might have considered there to be some sense in how a self/personal identity may continue even when not purposely reflected upon, and that such might therefore have more weight than has typically been given to the self on Hume's account. Kristjánsson too may have simply wished to indicate that he used Hume as a jumping off point, but even so the problem of how the self stands in relation to personal identity in Kristjánsson's self-concept remains, especially given our conclusion that in Hume's account, as in many anti-realist accounts, the terms are interchangeable. Hume's work may be an inspiration for a realist self-concept but it cannot be the basis for one.

4. Closing remarks

Where does this leave us? I find Kristjánsson's account to be an intriguing one and to possibly point to a way in which a realist basis could be found for the self that is understood in neither the sense of a soul nor that of a Cartesian ego. There may indeed be layers of realism within a realist concept and if such could be convincingly argued for then further alternatives to the realist/anti-realist dichotomy might well become attractive in studies on the self. We clearly approach our world and daily lives from the perspective of an enduring self or identity and for that reason alone the subject warrants much further investigation. Of particular interest in this regard might be the use of phenomenological methodology to augment whatever theoretical basis may be found for realist proposals.³⁶ The question of the self may well prove to be an ultimately unanswerable one and it seems very likely to continue to perplex all those who take an interest in just what kind of creature we human animals are, yet its

³⁵ Although Hume does repeat this 'case of incautious writing' (if it was such), we should not be put off by that; he may simply have jotted down a note or referred back a few pages in his writing, keeping his focus on the main topic he was addressing and not worrying overmuch about exactly what he was saying about the self (perhaps because he considered the matter put to rest by his earlier investigations). At any rate, Hume's text is not authoritative and all we are trying to find out is what he most likely had in mind when he wrote of the self; in answer to that the evidence seems to point to an illusion.

³⁶ The account is far from complete but see again Oberg, *op. cit.* for a tentative start in this direction, as well as Andrew Oberg, 'Dreaming of AI Lovers' (forthcoming) for additional thoughts on the orientation such an account might take.

intractability should not deter us from its study for at heart this is a question of being and there can be no more foundational philosophical quest than being itself. If selves exist, how and in what way? Are selves necessarily tied to consciousness? Do nonhuman animals also have selves and if so do they differ from our own or are they the same? Furthermore, do we *have* selves or *are* we selves? The questions are as important as the answers.

References

- Barresi, John, & Martin, Raymond, 'History as Prologue: Western Theories of the Self' in *The Oxford Handbook of the Self*, ed. by Shaun Gallagher (Oxford: Oxford University Press, 2011), pp. 33-56
- Dainton, Barry, *The Phenomenal Self* (Oxford: Oxford University Press, 2008)
- Damásio, António, *Descartes' Error: Emotion, Reason, and the Human Brain* (New York: Putnam, 1994)
- Dennett, Daniel C., *Consciousness Explained* (New York: Little, Brown and Co., 1991)
- Gertler, Brie, 'Self-Knowledge', *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), ed. by Edward N. Zalta. <<http://plato.stanford.edu/archives/sum2015/entries/self-knowledge/>>
- Hagen, Steve, *Buddhism: Plain and Simple* (New York: Broadway Books, 1999)
- Hatfield, Gary, *The Routledge Guidebook to Descartes' Meditations* (Abingdon, OX: Routledge, 2014)
- Hume, David, *A Treatise of Human Nature*, intro. and ed. by David Fate Norton and Mary J. Norton (Oxford: Oxford University Press, 1739/2000)
- Kristjánsson, Kristján, *The Self and Its Emotions* (Cambridge: Cambridge University Press, 2010)
- Locke, John, *An Essay Concerning Human Understanding*, 2nd edn (1689/1690). Available on Project Gutenberg: <<http://www.gutenberg.org/cache/epub/10615/pg10615.html>>
- Oberg, Andrew, 'A Realist Self?', *Journal of Applied Ethics and Philosophy*, 7 (2015), 24-33
- Oberg, Andrew, 'Dreaming of AI Lovers', (forthcoming)
- Parfit, Derek, *Reasons and Persons* (Oxford: Oxford University Press, 1984)
- Strawson, Galen, *Selves: An Essay in Revisionary Metaphysics* (Oxford: Oxford University Press, 2009)
- Tilakaratne, Asanga, *Theravada Buddhism: The View of the Elders* (Honolulu: University of Hawai'i Press, 2012)
- Wallis, Glenn, trans. and comm., *The Dhammapada: Verses on the Way* (New York: The Modern Library, 2007)
- Watts, Alan, *The Way of Zen* (New York: Vintage Books, 1989)

