

## Hume's Causal Account of the Self

Nathan Brett

Volume 9, 1990

URI: <https://id.erudit.org/iderudit/1012607ar>

DOI: <https://doi.org/10.7202/1012607ar>

[See table of contents](#)

### Publisher(s)

Canadian Society for Eighteenth-Century Studies / Société canadienne d'étude  
du dix-huitième siècle

### ISSN

0824-3298 (print)

1927-8810 (digital)

[Explore this journal](#)

### Cite this article

Brett, N. (1990). Hume's Causal Account of the Self. *Man and Nature / L'homme et la nature*, 9, 23–32. <https://doi.org/10.7202/1012607ar>

## 2. Hume's Causal Account of the Self

### I. INTRODUCTION:

What concept of mind is presupposed by Hume's account of human nature? In this discussion I will attempt an interpretation of Hume's theory of the self which I believe to be plausible in itself and as an account of what Hume intended. Briefly stated, the view that I end up with is this: a person is a combination of body and mind. Bodily continuity provides the basis for individuating the series of perceptions which are human minds. But bodily continuity is not sufficient as an account of the identity of mind. A mind is a causally integrated system of perceptions. A mind is not something which could exist apart from a body. A mind is a series of perceptions (a succession of states of consciousness). But it is misleading of Hume to say that the mental self is merely this. We can conceive of a series of states of consciousness in a single human body under conditions where this series does not constitute a human mind.

Hume is often forced to assume what he has taken some pains to deny. To take one example: the copy principle<sup>1</sup> which he frequently employs as a critical tool, could not be intelligible if it were correct. For, it is inconceivable that there could be 'some one impression' from which the ideas signified by the universal quantifiers 'all' and 'every' are derived — and yet these are embedded in various statements of this principle. In reflecting on Hume's comments about personal identity, I will be testing to see whether it suffers from this kind of fault. These problems will function in part to drive my discussion forward, as I attempt to show how his account can escape from inconsistency or circularity.

### II. MINDS:

A. *False Ideas of the mind*: I shall begin with a brief outline of the interpretation that I give to Hume's discussion of personal identity. My primary concern in this section is with the identity of minds rather than that of

persons taken as a whole. It is clear that Hume intends to deal with this restricted topic in the essay on personal identity. It is not clear just what his view is concerning the relation of the concept of 'person' and (human) 'body'. I shall return to that subject in the next section.

(1) There are, first, some *relatively* uncontroversial points about what Hume *denied*. In the essay on the immateriality of the soul he denies that it even makes sense to ask whether the soul is an immaterial substance. In rejecting this model, Hume argues that the substance/attribute formula cannot resolve any of the questions about the mind's identity. The theory of substance merely crowds the mental realm with something totally inaccessible, whose relation to perceptions ('inherence') remains utterly mysterious. I have previously defended this critique of the concept of mental substance.<sup>2</sup>

(2) Hume also denied that the self exists as a content of consciousness, as something of which we are every minute 'intimately conscious' (T, 251). This is his opening move in the essay on personal identity. There are two complaints that he has against this view which makes the self observable: (a) that he finds no introspective evidence for it and (b) that even on the assumption of its existence, it would only be another perception, and not 'that to which our several impressions and ideas are suppos'd to have a reference' (*ibid*). Hence, the hypothesis of an observed continuous self is, in Hume's view, both false and useless.

(3) If the supposition of a persisting unchanging subject of experience (an immaterial substance) is nonsense, and the hypothesis of an unvarying *content* of consciousness is both false and pointless, then it seems that there is no invariant and uninterrupted entity either beyond or within our train of perceptions. Hence, the question arises as to how we come to suppose that we are 'simple' and individual. How do we come to ascribe identity to a series of perceptions which is so obviously diverse — consisting as it does of separate perceptions which are qualitatively diverse?

In attempting to give Hume's answer to this question, one leaves behind any claim to an *uncontroversial* interpretation. The question of what enables us to attribute identity to a *series* of perceptions, seems straightforward enough. But Hume's answer to this question is most baffling. He seems to simply fill in here with some material left over from his discussion of the belief in 'body' (matter). We ascribe identity to this series of perceptions because we are led to overlook the distinctness and interruption of the perceptions that comprise it. How this now yields an idea of *self* identity, where it led to the idea of *bodily* identity before, is left unexplained. Professor Penelhum has remarked that it is almost beyond belief that anyone could make this mistake or be reassured by it.<sup>3</sup> Yet the claim seems to be made too clearly to allow us to sweep it

under the rug which covers the products of Hume's own carelessness and inattention.

I am no longer convinced, as I once was, that Penelhum is justified in dismissing this part of Hume's account as unnecessary and unbelievable. It is bizarre of course to suppose that anyone could come to believe that her entire array of changing sensations, thoughts, feelings, etc., was really just a single unchanging and uninterrupted perception. But we need not suppose that Hume is attributing to us any such grotesque fallacy. He is, I believe, looking for an account of a fictitious self-awareness which one might arrive at while fully conscious of the changes of (some of) his perceptions. Hume's account at this point displays once more his commitment to a certain form of empiricism. Earlier he had puzzled over the question of where the idea of 'identity' could have come from (T, 200-1). It is through comparison that we arrive at identity. The comparison of something unchanging amidst some other series of changes gives rise to the idea of persistence through time. One is looking at a peach (has an invariant and uninterrupted perception of the peach) while a tune is running through his head. From this array of perceptions there emerges the idea of peach as an object which persists through time. By a distinction of reason one can arrive at the general idea of identity through time — though of course there is no idea of identity apart from specific instances of non-variation<sup>4</sup> through change. In puzzling over the idea of self-identity Hume is looking for something which *is perceived to be* invariant through some changes which generate an awareness of the passage of time. But, we have already seen that his introspective search yields no such unchanging perception. Unless he is simply blind to something that others can notice, then, this sort of self-identity is just a fiction.

Even a *fiction* of self-identity presents Hume with a problem, however. Completeness demands that he give some account of the origin of this fiction. Had Hume been a twentieth-century British philosopher, he would have almost certainly have filled-in with a linguistic diagnosis. In fact, Hume himself does make occasional use of this type of diagnosis. However, in addressing the diagnostic task he has assigned himself here, Hume has recourse to his much more characteristic phenomenalist account. The idea of a persisting self arises from our carelessness and inattention.<sup>5</sup> We ignore the interruption and variation of our perceptions and thereby arrive at the fiction of an invariant and uninterrupted content of consciousness. Presumably we *also* have an awareness of the opposite. There are still perceptions of the buzzing fly and the tune running in the auditory imagination. Otherwise we would not have the comparison necessary to arrive at identity. But along side all this is developing the fiction of a uniform awareness which comes to be re-

ferred to as 'self-awareness.' People — or at least some people — have come to believe that along with all of the other contents of their consciousness, the booming, buzzing confusion of experience (to use James' expression) there is also a reflexive awareness of self. Hume insists that there is not such awareness; there is *just* the series of changing perceptions. Those who suppose that they have such an awareness derive this notion from stretches of their experience which are relatively uniform. Relatively uniform somatic sensations, for example, could be what an individual is in fact depending upon when he claims to have an awareness of a self which persists through time.

B. *The True Idea of the Mind*: We need to consider at this point a puzzle which arises from the fact that Hume is presenting this as a diagnosis of the origin of an error. If we take Hume to be denying either the existence or the identity of the self or mind when he talks about fictions and mistakes, then the account seems degenerate. We end up with nothing which can function as the possessor of this fiction and nothing to which we can ascribe the activities or dispositions which generate this error. We can hardly suppose that it is the fictitious self that falls into error; nor can one suppose that a self or mind which lacks identity could be anything but fictitious.

Almost the whole of the essay on personal identity is written as if it were pursuant to the diagnostic question *How do we arrive at the fiction of self-identity?* As Hume moves forward, however, his argument gets caught in an undertow created by the fact that 'identity' — even as he uses the term — does not mean just what he claims it does. The last major argument of the section begins as a further elucidation of the associations that lead us to ascribe a fictitious identity to 'the mind of man' (p. 2569). It ends, however, in a discussion of 'the true idea of the human mind,' and with the conclusion that a person, like a republic, can undergo radical changes 'without losing his identity' (pp. 260-61). The mind has identity because it is a *system* of related perceptions. A mind is a set of perceptions which are causally integrated and whose causal integration is specified by the principles of association. Thus, there is a sense in which a mind is, not just a succession of perceptions; it is a causal association of perceptions. This, of course, is perfectly consistent with his earlier denial that a mind is some entity (substance) which *has* perceptions. Since the connections which give mental identity to any series of perceptions are *causal*, this also means that a human (or Humean) mind can only exist as something spread out across time.<sup>6</sup>

It is worth remarking here that Hume's initial introduction of the principles of association contains at least the seeds of this theory. At the very beginning of the *Treatise* (p. 10) he speaks of them as 'universal principles which render (the imagination) in some measure uniform

with itself in all times and places.' This suggests that the concept of mind which is articulated in Section VI is not an *ad hoc* invention which is designed merely to fill the gap left by his rejection of the immaterial self.

### III. MIND AND BODY:

Let me assume for the moment that this associationist view is not obviously self-defeating and that it does provide conditions under which a mind can preserve its identity over time. I shall consider briefly the questions of whether we can derive any insight from this concept of mind and how this account could be employed in a fuller account of persons which includes their bodies. My own interest in the Humean concept of mind stems from the belief that adequate accounts of persons must dispense with the 'ghost in the machine,' but that it must not dispossess those conscious states which Hume refers to as perceptions. Now, the view of self which Gilbert Ryle derides as a ghost is not unlike the view which Hume attempts to debunk in his critique of the immaterial self. But the positive account of mind which Ryle advances is one in which a great many questions arise about the status and even reality of the sensations, images, thoughts, etc., which are the raw materials of the Humean mind. In reading Ryle, I find myself unsure whether these mental items have made their exit with the ghost, or somehow remained with the machine - to put it crudely, perhaps misleadingly. The Humean concept of mind as a causal system offers at least the possibility of making non-ghostly sense out of the consciousness of persons.

On the other hand, Hume does leave us in some doubt as to the status of human bodies. Nearly the whole of the essay on personal identity proceeds as if this were entirely a question about *mental* identity. But it is equally true that he comes to include the body in the idea of the self in the discussions of Book II (Cf. T, 298). In fact, the human body has really been in the background of many of the discussions in Book I. This is evident when he talks of the 'dissolution of the body' when he considers death, when he talks of the organs of sensation, and when he suggests an account of what happens in the brain when we fall into certain errors (T, 64). Even though the essay on personal identity itself is primarily a discussion on mental identity, Hume does briefly discuss the retention of identity through change of bodily characteristics. Perhaps the most important claims in this regard are made in Hume's discussion of the mind-body problem, a discussion which occurs almost immediately prior to the analysis of personal identity. (See T, 246-48.) Unlike the other questions which he considers in the chapter on immateriality, he does not reject this one as ultimately unintelligible. The

problem he considers is whether bodily changes could possibly give rise to anything as unlike themselves as perceptions. The answer he supplies is that

everyone may perceive that the different dispositions of his body change his thoughts and sentiments; . . . we may certainly conclude that motion may be, and actually is, the cause of thought and perception. (T. 248)

The conclusion which I am inclined to draw from these passages is that Hume is in fact working with a concept of person which does include the body as well as the successive states of consciousness which are (again) causally related to it. It is bodily continuity that provides him with the *individuation* of the series of perceptions, which he simply assumes when he looks within himself and when he imaginatively looks within the breast of another. And it is this assumption that becomes explicit in the essay on personal identity when he remarks that 'An infant becomes a man, and is sometimes fat, sometimes lean, without any change in his identity' (T, 257).

Although Hume does assume bodily identity in his discussion of personal identity, he does not assume that this is *sufficient* to answer the questions that can be raised about the identity of minds, or selves. The question as to why, and indeed whether we should consider the conscious states of this body to be contents of a single mind, he supposes to be a separate issue. Is he right about this?

Let us see if we can conceive of a case in which we have both bodily continuity and consciousness but where mental identity is not present. Assume that there is a human body which (in some sense) has perceptions. Assume also that these states occur in a wholly random fashion. (We might give this a greater sense of reality if we imagine that this body is festooned with brain probing electrodes whose electrical potential is controlled by a randomizing device). The resulting 'stream' of consciousness might be as follows: a glimmer of hope is succeeded by the smell of parsley; an expanse of green is followed by a dull ache, which yields to the thought of triangularity, and so on. These sensations and thoughts continue to occur in such ways that they never have any connection with one another; their only connection is that they are all related to the same body. Only these stimuli can generate conscious states in this body.

Under these conditions would there be any non-trivial sense in which the resulting succession of perceptions is something to which one could significantly attribute identity? Does one have here either a mind or a person? <sup>7</sup> I can see no reason to think so. If there is not any reason then we can assume that one does not answer questions about the identity of

minds or persons simply by giving perceptions a reference to a body. Thus, it seems both possible to raise the question of mental identity as an issue separate from the bodily criterion of individuation. It also appears to be necessary to raise this question if we are to arrive at an adequate theory of persons. Whether Hume's own causal answer, based on the principles of association was adequate or even *coherent* is another matter. I will turn now to the latter of these issues, to the criticism that his account is self-referentially absurd.

#### IV. HUME'S APPENDIX — A PARADOX OF THE HEAP OR BUNDLE:

If, as I have argued, Hume has neither denied the existence of a continuous self, nor the possibility of a genuine concept of self, there still remains a paradox when the causal account of the mind is considered in relation to Hume's account of causality itself. It is this difficulty that Passmore focuses upon in his critique of Hume's theory of the mind. And, as I see it, it was the realization of a possible clash between these two that led Hume to his confession of failure in the *Appendix* to the *Treatise*:

...upon a more strict review of the section concerning *personal identity*, I find myself involv'd in such a labyrinth, that, I must confess, I neither know how to correct my former opinions, nor render them consistent. (T, 633)

If perceptions are distinct existences, they form a whole only by being connected together. But no connexions among distinct existences are ever discoverable by human understanding. We only feel a connexion or a determination of the thought, to pass from one object to another. (T, 635)

The present philosophy, therefore, has so far a promising aspect. But all my hopes vanish, when I come to explain the principles, that unite our successive perceptions in our thought or consciousness.... In short there are two principles, which I cannot render consistent, nor is it in my power to renounce either one of them, viz. that all our perceptions are distinct existences, and that the mind never perceives any real connections among distinct existences. (T, 635-6)

Suppose that the true idea of the mind is that of a causal system of perceptions. These perceptions themselves are 'distinct existences' — as they must be if they are causally related. So far, this seems respectable enough: the fact that these are distinct entities does not exclude the possibility that their relation gives rise to another kind of thing, viz., the mind or self. But at this point Hume is struck by the thought that the discernment of causal relations is not the discovery of 'real connexion'



(T, 636). He concludes from this that the whole account is unsatisfactory, though he leaves open the hope that there may be some hypothesis that can save it.

I believe that the key to understanding this much debated passage and the embarrassment which Hume expresses, lies in the interpretation of the word 'real'. With what does the 'real' contrast in this context? It is not that which is fake or illusory, nor with what is contingent, as opposed to necessary. Nor is it true that *real* connections (as opposed to the causal ones in question) are bonds which remove numerical distinctness, as Professor Penelhum has recently argued.<sup>8</sup> It is true that causal relations are not like some epoxy which can make one object out of many. (Parent and child are causally related, but that does not make them one object.) This, however, does not seem a good reason to suppose that they are not real *connections*.

The appropriate contrast does emerge, I believe, when one reflects on what Hume has said about causal relations: they are not 'real', but *mind-dependent*.<sup>9</sup> The definitions of 'cause' which Hume had provided at the end of his discussion of the idea of necessary connection emphasize that the connection we have noticed when we think of events as causally related is, at bottom, an association of impression and idea.<sup>10</sup>

Consider, now, the section of the essay on personal identity which becomes the object of Hume's criticism in the *Appendix*. A question arises, he says

...whether in pronouncing concerning the identity of a person, we observe some *real* bond among his perceptions, or only feel one among the ideas we form of them. This question we might easily decide, if we would recollect what has already been prov'd at large, that the understanding never observes any real connexion among objects, and that even the union of cause and effect, when strictly examined resolves itself into a customary association of ideas. (T, 259-60; italics mine)

It is clear that this view of causality will yield a paradox when it is combined with the causal account of the mind. If the mind must add something to the regularities that are discerned in order to arrive at causal connections; and if the mind is itself a causal system, then *the mind must superimpose order upon itself in order to exist*. But there is nothing left over to bear this task of superimposition. Hence the causal account of the mind runs afoul of the mental account of causality. We must either say that this is an exception, in which *real* causal relations exist and are discerned; or we must say that the mind which the account of causation presupposes has a unity that is non-causal, e.g., through the supposition that perceptions inhere in something simple and individual. It was not

within Hume's power, i.e., within the boundaries of his system, to make either of these adjustments.

## V. CAUSAL CONNECTION:

In this final section I want to defend Hume's account of the mind against this line of attack by arguing that the trouble is with the supposition that causation can have the status of a secondary quality. I do not believe that this will yield an interpretation that is particularly damaging to his system as a whole — though it would take another long paper to fully explore the implications of this maneuver. I would argue that the analysis of causation as 'mind dependent' involves a confusion of psychological and epistemological problems. The result is an account that is incoherent and hence untenable irrespective of what one says about minds. To be more specific: the view that we arrive at causal relations *because* the mind projects its own tendencies into the regularities that it discerns can only be taken as an account of the *genesis* of the idea of causal connection. If this genetic account is taken to be an analysis of causation it becomes a self-destructive account. For, it is itself a causal account, an answer to the question, viz., 'What causes us to think in causal terms?' This is not an absurd question; but it is one which presumes the reality of causal relations. Now, it may well be that causal relations are only discernable by a mind which acquires certain patterns of transition. And, again, it may be that the causal properties of *minds* — or systems of perceptions — are only to be discerned when the complex idea of such ordered transitions becomes available. It remains true that this is a discovery or discernment, not an invention or superimposition of causal connection. To interpret Hume as holding that causal relations are not 'real' but are mind dependent is to take literally some of the things that he says.<sup>11</sup> But it makes nonsense of a very great deal more than the theory of the self; and ultimately it produces an account of causation that is itself incoherent.<sup>12</sup>

The idea that causality could be shown to have the status of a secondary quality was a skeptical thesis that fascinated Hume. But Hume also conceived of himself as engaged in a Newtonian enterprise, providing an account of the mental realm comparable to that which Newton had provided for the physical world. What I have attempted to sketch in this discussion is the concept of mind that emerges from Hume's attempt at this Newtonian project.

NATHAN BRETT  
Dalhousie University

## Notes

- 1 The principle that 'every idea is derived from a precedent impression' (T, 232). References to the *Treatise* (T) are to the (1888) Selby-Bigge version as revised by P.H. Niddich (Oxford: Clarendon, 1978). In this second edition, the pagination of the earlier edition is preserved.
- 2 In 'Substance and Mental Identity in Hume's *Treatise*,' *The Philosophical Quarterly*, 1972.
- 3 In *David Hume* (London: Macmillan, 1975), pp. 77-9.
- 4 I have simplified the description here by leaving out 'uninterruption.' In doing this I have deliberately avoided some special problems which this introduces, problems connected with the fact that Hume moves back and forth between attributes of the objects of perception and attributes of the process of perceiving an object.
- 5 This is my application of Hume's phrase, not his.
- 6 What was said above about identity entails that anything to which one can meaningfully attribute identity must be assumed to have duration in time. But there is a difference between mind and body. We could (through a distinction of reason) think and speak meaningfully of the existence of a body at a moment of time. There is no comparable thought about (Humean) minds.
- 7 If one ascribes these states to a mental substance then, of course, there is a disordered but conscious mind. This clearly is not relevant to the problem under discussion, viz., whether using bodily continuity as a criterion for individuating series of perceptions is itself sufficient for mental *identity*.
- 8 Penelhum, *David Hume*, p. 80.
- 9 Compare Hume's use of the word 'real' in his discussion of the doctrine of secondary qualities (T, 227). Here the contrast is clearly between the 'real' and the mind-dependent qualities of objects.
- 10 Here is the second definition: 'A CAUSE is an object precedent and contiguous to another, and so united with it, that the idea of one determines the mind to form the idea of the other, and the impression of one to form a more lively idea of the other' (T, 170).
- 11 See for example T, pp. 165, and 167, as well as the definitions of 'cause' (p. 170).
- 12 One of the odd properties of Hume's definition of causation is that it has embedded in it what is quite obviously a causal conception: 'A CAUSE is an object precedent and contiguous to another, and so united with it, that the idea of one determines the mind to form the idea of the other . . .' (T, 170; italics mine). Taking this as an analysis of the term 'cause' will thus mean crediting Hume with an analysis that is circular. Taking this as a phenomenalist account of the origin of the idea of necessary connection, on the other hand, yields an account that is neither self-destructive nor entirely implausible.