

**Hume's Subjectivization of the Self:
Personal Identity As Affective Continuity
Wayne Waxman**

The problem of personal identity as Hume understood it could hardly be more different from how it is conceived today.¹ Its solution has nothing much to do with the specification of identity criteria for selves – in essence, how to count them, to tally up the individuals instantiating this category rather than, say, organisms with human genotype, Social Security card holders, movie characters, or other natural and/or non-natural kinds with which it may be liable to be confused. Nor is its solution a matter of analyzing the use of the singular reflexive first-person pronoun to single out which person *I* am from those I am not and why I am not a *we*. These and other current conceptions of the problem of personal identity take for granted the existence of complex individuals, a world of material beings subject to natural law, even a social realm governed by norms. By contrast, Hume's problem, as I understand it, takes its start from a "world" that is nothing but a "perpetual flux and movement" of perceptions – impressions and ideas – devoid of all connection and relation, whether to one another or to anything unperceived ("They are the successive perceptions only, that constitute the mind; nor have we the most distant notion of the place, where these scenes are represented, or of the materials, of which it is compos'd," THN 253/165). And Hume's solution to it consists in applying principles of associative imagination to show how, from this disjointed welter, we form the idea of our self, individual at each time and identical over time, prior to and independently of our ideas of all other complex individuals.

¹This talk is adapted from the Hume of *Kant and the Empiricists. Understanding Understanding*, New York: Oxford University Press, 2005. I abbreviate *A Treatise of Human Nature* 'THN' and refer first to page numbers in the Selby-Bigge/Nidditch edition and second to page numbers in the new Norton/Norton edition, both Oxford University Press. I abbreviate *An Enquiry concerning Human Understanding* 'EHU' and refer to passages by section and paragraph number in the new Oxford University Press edition by Tom L. Beauchamp.

I shall begin my examination of Hume's account with his explication of the idea of identity itself. While it may seem impossible to extract any idea of identity at all from a flux of non-complex, unrelated perceptions in accordance with Hume's principle that all ideas are copied from impressions, it will emerge that he traced it to an origin not in the objects present to consciousness – the flux of perceptions – but rather in consciousness itself, and in particular the affective disposition of associative imagination when contemplating this flux (what he termed "perfect identity"). Next, I will consider Hume's attempt to explain the idea of the identical self as the result of an unavoidable, ineradicable confusion between the affective disposition distinctive of identity with another affective disposition so closely resembling it that we are induced to overlook the radical qualitative disparity on the side of the appearances (objects, perceptions) and treat the non-identical as identical ("imperfect identity"). And, to conclude, I shall examine some implications of Hume's account of personal identity for the rest of his system.

I. The fiction of Identity

Hume forswore definition as a means of explicating the contents of ideas, opting instead to trace ideas to their originating impressions of sensation or reflexion (THN 157/106 and EHU VII/i ¶4). However, when he subjected identity to this treatment – the idea of an existence that continues through successive times – he could find no originating impression at all. Far from continuing the same for even two successive moments, or even seeming to, "all impressions are internal and perishing existences, and appear as such" (THN 194/129). Indeed, he deemed this to be true of perceptions generally, ideas no less than impressions: they are "interrupted, and perishing, and different at every different return" (211/140; also 252-3/165). Thus, for want of anything "stedfast and unchangeable" (37/30) in perception, he could

discover no impression from which to derive the idea of an existence that continues the same through time.

Here it is worth remarking that this result should not be confused with a semantic thesis as to the meaningfulness of our discourse concerning unchanging existents (THN 64-5/47). Hume never pretended that linguistic meaning is, should, or can be confined by the limitations of our capacity to form ideas; quite the contrary in fact. But since, for him, impressions and ideas are the only *objects* we are ever acquainted with – the only realities ever present to our senses or thought – no *objective, ontological* worth can be accorded to discourse concerning continued existents in the absence of an idea to underwrite it. Hume denied objective worth to identity on the ground “that time cannot make its appearance to the mind, either alone, or attended with a steady unchangeable object, but is always discover'd by some *perceivable* succession of changeable objects” (35/28); and “since the idea of duration cannot be deriv'd from such an object, it can never in any propriety or exactness be apply'd to it, nor can any thing unchangeable be ever said to have duration” (37/30). This principle restricting scope of application by origin only governs ideas, not discourse, and in no way conflicts with the utility, even indispensability, of our talk of unchanging objects. Only attempts to accord an extra-semantic, ideational import to it would have prompted Hume to object that, ontologically and objectively, the very notion of an enduring existence is entirely unintelligible.

But even if no idea of enduring existence underwrites our discourse, Hume did allow that there is a fictional application of time to what is unchangeable that can “make us fancy we have that idea” (THN 65/47). This fiction, which “almost universally takes place” (200/133), starts from the fact that “the idea of time [is] ever present with us” because of the “continual succession of perceptions in our mind” (65/47). Since anything co-existent with these perceptions will also partake of the idea of time, an object observed without interruption over an interval that exhibits no variation is fancied to partake in the duration marked out by our successive perceptions without at the same time partaking of their succession. Of course, if we consider any of the successive moments within that interval

together, our ideas of the object are *ipso facto* multiplied, that is, we will be conceiving not identity but *number* (201/134). To avoid multiplying the object – that is, to be able to overlook the manifest multiplicity of (qualitatively identical) objects that in fact are present to our perception – we would need to abstain from considering the successive moments together. The have-your-cake-and-eat-it-too impossibility at the heart of this exercise – Hume's proof that identity over time is a fiction and not a genuine idea at all – is the supposition that, in not considering the successive moments together, we can nevertheless conceive the time marked out by that succession, and so use it to conceive the object as enduring identically, without variation or interruption. For if Hume's explication of time is correct, we can never, without considering a succession of different moments together, even so much as conceive time or, *a fortiori*, the duration of any existence (35-7/28-30). Consequently, the true ideational outcome of the exercise called for in the fiction of identity would be to banish time altogether from our consideration of the object, yielding not identity but what Hume termed *unity*, where, in the proposition “an object is the same with itself,” nothing (including time) intervenes to distinguish ‘object’ from ‘itself’ (200/133).

Of course, the impossibility of forming the idea does not prevent us from succumbing to the fiction of identity every time we encounter *qualitatively* invariant, uninterrupted sequences of perceptions. The case here seems closely analogous to, or even an instance of, a type of fiction to which Hume returned frequently over the course of the first book of the *Treatise*: our “strong propensity to compleat every union by joining new relations to those which we have before observ'd betwixt any ideas” (THN 217/143).² The principal example Hume

²The propensity to complete the union by impossible extensions of relations itself seems to be an instance of the propensity of the imagination to continue any action “even after the reason has ceas'd, which first determin'd it to begin” (THN 48/36-7), which also contributes to the fiction of identity by means causal coherence: “the imagination, when set into any train of thinking, is apt to continue, even when its object fails it, and like a galley put in motion by the oars, carries on its course without any new impulse” (198/132). The identity fiction Hume

provided of this phenomenon is the fictitious spatiality accorded to objects that, in and of themselves, are aspatial: finding a constant conjunction between the look, feel, smell, and flavor, say, of a fig, and finding that the two former qualities are, in addition, related by contiguity in space, we proceed to complete the union by extending the contiguity relation to the two latter qualities as well: and we do this notwithstanding the impossibility of spatially situating objects that, quite literally, exist nowhere, and can no more stand in a relation of contiguity with the objects of touch and vision than moral reflections and passions can (THN 235-7/154-6). Extending this model to the case of identity over time, we may be said to complete the union by extending the properties specific to the idea of duration that we are only capable of forming by perceiving a succession of distinct objects (impressions and ideas) to something in which no succession is perceived at all, so “that the unchangeable or rather fictitious duration has the same effect upon every quality, by encreasing or diminishing it, as that succession, which is obvious to the senses” (65/47-8): and here too we do this despite the fact that anything in which no succession is perceivable is *ipso facto* impossible to conceive as enduring.

Given that the impossibility of the idea no more detracts from the ubiquity and force of fictions of identity than the impossibility of spatially situating the aspatial represents no obstacle to the universal human propensity to situate flavors in fruits or thoughts in people's heads, the question that arises is how, according to Hume, we pull off such impossible stunts: that is, how, when our object completely fails us, do we nonetheless manage to deceive ourselves that we are conceiving a genuine identity? The answer, it turns out, also holds the key to comprehending Hume's notion of personal identity.

II. Perfect and Imperfect Identity

Since associative psychology lies at the heart of Hume's account of the fiction of identity, we need briefly to consider what an associative relation is, and in particular his thesis that the nature or

describes on 65/47-8 and 200-1/133-4 seems a perfect example of a relation the imagination continues even after the object fails it.

essence of such a relation is a *facile* (easy, smooth, natural, etc.) transition of thought. There is probably no more important element of Hume's philosophy that commentators have more consistently or completely overlooked, despite the fact that on no less than four occasions Hume explicitly states that facility is the essence of relation, three of which specifically concern the identity and simplicity of complex individuals, including personal identity (THN 99/69, 204/135, 220/145, 260/169-70; see also 115-16/80, 289/189-90, 355-6/230, 378/243-4, 510n/327n, and EHU III ¶12). Although Hume never elaborated on what he meant by a *facile* transition, its equation with associative relation, together with such textual evidence as there is, all suggest that he had in mind the *affective* dimension of such transitions, i.e. the way they *feel* to the imagination that makes them. In particular, it seems to have furnished him with an answer to the question, what are we conscious of when we are conscious of a relation? If a relation is nothing more than a transition from the appearance in impression or idea of one object to the thought of some other object, what is it we are aware of in the course of such a transition – the “associating quality” (10/12) as Hume termed it – that distinguishes it from an arbitrary or random transition, in which no relation is recognized? A transition of thought relates the perceptions concerned in it just in case the transition *feels* easy (smooth, natural) to the imagination. If a transition lacks this feeling of facility (the “indifference” he ascribed to chance transitions at 124/87), or if it is characterized by its opposite (e.g. the “sensible violence” felt when contemplating transitions contrary to ingrained customary associations: 128/88), then, no matter how constantly experience may reinforce a conjunction of perceptions, or how strong the custom of making the transition may be, we will not regard them as related. Associative relation, for Hume, depends essentially on that “facility and force,” that “firm hold and easy introduction,” that alone suffices to distinguish a transition to a certain idea “from any new and unusual idea” (116/80).

Fictions of identity are products of associative relations of resemblance. Human nature is so constituted that whenever qualitatively identical objects appear to us, we feel an easy transition

from one to the next. A succession of such objects will accordingly be the occasion of a succession of facile transitions, each no different to the feeling than the last, with the result being a sustained – *unvarying, uninterrupted* – affective disposition:

When we fix our thought on any object, and suppose it to continue the same for some time; 'tis evident we suppose the change to lie only in the time, and never exert ourselves to produce any new image or idea of the object. The faculties of the mind repose themselves in a manner, and take no more exercise, than what is necessary to continue that idea, of which we were formerly possess, and which subsists without variation or interruption. The passage from one moment to another is scarce felt, and distinguishes not itself by a different perception or idea, which may require a different direction of the spirits, in order to its conception. (THN 203/135)

To understand this description of what Hume termed a 'perfect identity', recall his claim that "Where we have no successive perceptions, we have no notion of time, even tho' there be a real succession in the objects" (35/28). It derives from Locke's thesis that succession is an idea that has its source not in sensation (external sense) but in reflexion (the internal sense of the operations of one's own mind). In the present context, where we are witness to a qualitatively invariant appearance, this means that we can only become aware of the passage of time by observing a change of perceptions inwardly, in our contemplation of the appearance. That is, unless we were aware, however vaguely, of the effort, however slight, needed to continue the same, qualitatively unchanging image or idea we had formerly into the present, we would perceive no succession at all, and so too no duration. So, the phenomenon of a perfect identity depends crucially on our inwardly perceiving transitions between fleetingly existent successive perceptions, and so on the very contrary of identity. The reason it nevertheless gives rise to the fiction of identity lies in the fact that, though felt, the transitions are "scarce felt." This is because the feeling incident to producing an idea remains invariant and

uninterrupted so long as the same idea is successively repeated. The moment a new idea must be produced, requiring "a different direction of the spirits, in order to its conception," the feeling is ended and, with it, the facile transition essential to identity or any other relation. Thus, in addition to the qualitative identity of the objects from and to which the transitions are made, the transitions themselves must be qualitatively indiscernible in feeling in order to produce the impression original of the idea of identity.

Here it is important to take account of the remark with which Hume prefaced his description of perfect identity. Resemblance "not only causes an association of ideas, but also of dispositions," where by 'disposition' he means the "act or operation of the mind" in which "ideas place the mind" (THN 203/135). This is an unusual use of the term 'disposition' even for Hume, but it is not unique, as when he commented in a related context (leading up to his first enunciation of the thesis that facility is the essence of relation) that "All the operations of the mind depend in a great measure on its disposition, when it performs them; and according as the spirits are more or less elevated, and the attention more or less fix'd, the action will always have more or less vigour and vivacity" (98/69). In this use of the term, Hume seems to have had uppermost in mind the mind's *affective* disposition, its feeling in performing its actions, and it is in this sense too, I believe, that he would have us understand his use of the term in connection with perfect identity: the facile transitions constitutive of the resemblance relations conjoining successive qualitatively invariant perceptions are themselves so alike to the feeling as together to constitute a sustained – invariant, uninterrupted – affective disposition that continues just so long as no new image or idea necessitates a different direction of the spirits. What this means becomes clearer if we consider how the same appearances might be considered in the imagination if human nature were different. For example, if the action of successively repeating the same idea, instead of being scarce felt, required as great an effort as presently is required to successively produce new ideas, then we would be just as clearly aware of the nature of the former as a succession of fleetingly existent objects as we are of the latter. In particular, there would be no sustained –

invariable, unchanging – affective disposition to tempt us to confound what is, both in appearance and in reality, “interrupted, and perishing, and different at every different return” (211/140), with something in perception answering to our semantic notions of identity. Thus, for Hume, our propensity to fictions of identity is as much a function of the affective disposition of associative imagination as of the appearances before it.

Indeed, this claim is too weak. For once we turn from perfect to imperfect identities – identity even in the face of interrupted and/or variable appearances – it immediately becomes evident that, for Hume, the primary locus of all identity, perfect no less than imperfect, is the affective disposition of associative imagination. Thus, the principle invoked time and again in his explanations why things bearing little or no resemblance to perfect identity are nevertheless invariably confounded with it, and treated as identities in all our thought and action, is that they “place the mind in the same disposition, or in similar ones,” so that it “readily passes from one to the other, and perceives not the change without a strict attention, of which, generally speaking, 'tis wholly incapable” (THN 203/135). In particular, we conflate what it feels like to successively conceive the same *relation* between ideas with what it feels like to successively conceive the same *idea*:

a succession of related objects places the mind in this disposition, and is consider'd with the same smooth and uninterrupted progress of the imagination, as attends the view of the same invariable object. The very nature and essence of relation is to connect our ideas with each other, and upon the appearance of one, to facilitate the transition to its correlative. The passage betwixt related ideas is, therefore, so smooth and easy, that it produces little alteration on the mind, and seems like the continuation of the same action; and as the continuation of the same action is an effect of the continu'd view of the same object, 'tis for this reason we attribute sameness to every succession of related objects. The thought slides along the succession with

equal facility, as if it consider'd only one object; and therefore confounds the succession with the identity. (204/135)

III. Personal Identity

Personal identity provided Hume with an ideal illustration of how identity comes to be attributed to manifestly non-identical phenomena owing to the close resemblance between the mind's affective dispositions in successively conceiving the same idea and in successively conceiving the same relation of ideas. To set the stage, he began by asking what the impression original of one's idea of self is. Since the self is supposed to continue invariably the same through the whole course of our lives, an originating impression would have to be constant and invariable. Instead of this, however, we find only that “Pain and pleasure, grief and joy, passions and sensations succeed each other, and never all exist at the same time” (THN 252/164). If not any one impression, might the self not then be a thing “to which our several impressions and ideas are suppos'd to have a reference” (251/164)? Yet, our perceptions offer no evidence of any such reference:

All these are different, and distinguishable, and separable from each other, and may be separately consider'd, and may exist separately, and have no need of any thing to support their existence. After what manner, therefore, do they belong to self; and how are they connected with it? For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch *myself* at any time without a perception, and never can observe any thing but the perception. (252/164-5)

Finding no non-fictitious pre-associative impression original either for an idea of the identical self or for a real connection of our perceptions to such a self, Hume concluded that, in reality, the mind

is a kind of theatre, where several perceptions successively make their appearance; pass, re-pass, glide away, and mingle in an infinite variety of postures and situations. There is properly no *simplicity* in it at one time, nor *identity* in different... They are the successive perceptions only, that constitute the mind; nor have we the most distant notion of the place, where these scenes are represented, or of the materials, of which it is compos'd. (253/165)

Yet, while we never observe anything “that really binds our several perceptions together,” there is something that “associates their ideas in the imagination,” that is, a bond we “feel ... among the ideas we form of them,” effected by “uniting principles in the ideal world” whose “very essence ... consists in their producing an easy transition of ideas” (THN 259-60/169). Our perceptions may be fleetingly existent and qualitatively diverse but this does not prevent us from associating them according to uniform relations so ubiquitous as to put our imaginations into an affective disposition effectively indistinguishable from that characteristic of identity (253-4/165-6). In particular, the ideal associative relation of cause and effect, in being applied to all our successive perceptions, suffices to yield the idea of a constant, invariable identity that persists through the entirety of our conscious existence:

the true idea of the human mind, is to consider it as a system of different perceptions or different existences, which are link'd together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other. Our impressions give rise to their correspondent ideas; and these in turn produce other impressions. One thought chaces another, and draws after it a third, by which it is expell'd in its turn. (261-2/170-1)

Causal relations between successive perceptions are continuously and intimately present to our consciousness. The connection of impressions of the senses – what we see, feel, smell, hear, and taste –

to their idea copies is causal. The same is true when an idea, say, of an approaching danger, causes in us a reflexive impression of dread. And we reason from effect to cause or vice versa when we infer a fire from sightings of smoke, frayed wiring in the walls from the fire, the gnawing of rats from the frayed wiring, and so on. Yet, in remarking that the true idea of the human is of a system of causal relations, Hume made clear that he was doing so with the specific intent of applying what he had proved earlier in the *Treatise* about causal relations generally to the particular case of personal identity (259-60/169): that our ideas of cause and effect, and particularly their principal constituent – ideas of necessary connection – originate in customary association (153-66/105-112), and so consist in nothing more than facile transitions of thought (260/169-70). Hence, if we wonder why Hume thought that a system of causal relations counts as the source of our idea of self, the answer is to be found in the sustained affective disposition that results from the uninterrupted and invariant succession of facile transitions of thought from cause to effect and from effect to cause continuously and intimately present to our consciousness (“the smooth and uninterrupted progress of the thought along a train of connected ideas,” 260/170).

This need not have been the case. If human nature had been such that forming ideas of successive causal relations demanded a considerable effort, and required always a different direction of the spirits, then no sustained affective disposition could result from a system of causal relations and the affective basis would be wanting for confounding such a system with an identity. It is therefore not causal relations per se that Hume equated with personal identity but the affective disposition to which their continuous succession gives rise in beings with natures like ours:

That action of the imagination, by which we consider the uninterrupted and invariable object, and that by which we reflect on the succession of related objects, are almost the same to the feeling, nor is there much more effort of thought requir'd in the latter case than in the former. The relation facilitates the transition of the mind from one object to another, and renders its

passage as smooth as if it contemplated one continu'd object.
(253-4/165-6)

Though a necessary condition for personal identity, a system of causal relations is sufficient only insofar as their succession gives rise to the same affective continuity we meet with when contemplating an uninterrupted, invariant appearance. When that further condition is met, we cannot help extending the fiction of identity (unchangeable existence) from the latter to the former, heedless of how "imperfect" that identity is (owing to the constant variation in the objects conceived, and the consequently ever changing direction of the spirits requisite to their conception). For "tho these two ideas of identity, and a succession of related objects be in themselves perfectly distinct, and even contrary, yet 'tis certain, that in our common way of thinking they are generally confounded with each other" (253/165-6).

The requirement of affective continuity is often neglected or regarded as inessential to Hume's explication of "the true idea of the human mind." One might defend this position by pointing to the fact that a causal system preserves its identity even if, at some point in time, it branches off into multiple discrete, non-interacting causal lines. By contrast, if personal identity is supposed to consist principally in affective continuity, a single person can ramify into as many distinct persons as there are discrete causal lines (some of which might later converge, forming again one person). The question is whether such purely conceptual possibilities would have mattered enough to Hume to induce him to treat affective continuity as ancillary to personal identity. With no experienced instance of branching causal systems of perceptions to warrant according it the least probability, I doubt that he would have reacted to it any differently than he did to the possibility illustrated by the missing shade of blue: "the instance is so particular and singular, that 'tis scarce worth our observing, and does not merit that for it alone we should alter our general maxim" (THN 6/10).

One rarely remarked difficulty that arises if affective continuity is discounted and personal identity equated with a system of causally associated perceptions is that, in the context of Hume's associationism,

it would make *my* identity and that of the *universe as a whole* one and the same ("as these are the only ties of our thoughts, they are really *to us* the cement of the universe," THN 662/416-7). It is causal association in particular that "peoples the world, and brings us acquainted with such existences, as by their removal in time and place, lie beyond the reach of the senses and memory. By means of it I paint the universe in my imagination, and fix my attention on any part of it I please" (108/75). To confine the scope of the causal system of perceptions so that it can plausibly be equated with the self, everything we suppose to exist distinct from (independently of and external to) the mind must be excluded from the system. To be sure, there is no reason to think that Hume did not mean to include causes and effects of all kinds in his system, not excepting the continued, distinct existents that result from the fiction of imperfect identity examined in THN I/iv/§2. Exponents of the view that he equated the self with a system of causally linked perception can still plausibly maintain that he regarded the objects painted in associative imagination in either of two ways: as complex ideas with sensation and memory constituents belonging to the identity of the mind, or as objects fictitiously supposed to have an existence distinct from the mind. When the causal system of perceptions is viewed in the former regard, the result is an idea of the self rather than of the universe as a whole, and affective continuity contributes nothing essential to its content at all.

Of course, a great deal of philosophy is involved in learning to view the physical world as consisting in nothing but complex ideas with sensation and memory constituents fictitiously regarded as preserving their identity through interruptions in their appearance to the senses. Does this mean that only philosophers who succeed in convincing themselves of this are capable of forming an idea of the self? Anyone inclined to answer yes probably regards Hume's explication of the idea as a debunking exercise designed to deny the idea of the self and set that of a causal system of perception in its place. It then bears a strong resemblance to positivist readings of Hume's account of cause and effect in favor in throughout much of the twentieth century, portraying his intent as being to deny the existence of any such idea and to set constant conjunction in its place (the

“regularity view”). Yet, just as most scholars today appreciate that Hume's preeminent concern with the origin, content, and scope of application of the idea of necessary connection stemmed from his conviction that we possess an idea of cause and effect distinct from constant conjunction and that this idea is essential to the cognitive and conative economy of our minds, we should also recognize that Hume's concern with the origin, content, and scope of application of the idea of personal identity was likewise motivated by a conviction that we possess an idea of the self distinct from that of a causal system and that this idea too plays a vital role in human cognition and conation.

A causal system of perceptions is just what it is: the source not of any idea of a single complex individual but of a system of individuals, simple and complex, that “mutually produce, destroy, influence, and modify each other” (THN 261/170). Only with the addition of affective continuity is there anything present to imagination resembling perfect identity and capable of being confounded with it. Moreover, as soon as we recognize that, for Hume, the locus of identity is the affective disposition of imagination, we can see that it makes no difference which things are included in the causal system that underpins it. For if the self is equated not with the system itself but the affective disposition whereof it is the occasion, there can be no absurdity in the fact that this system includes my body, your body, mountains, cities, moons, stars, and all the other objects that populate the boundless universe we paint inside our imaginations. Finally, by restoring affective continuity to its central place in Hume's explication of “the true idea of the human mind,” we thereby retrieve it from the rarefied realm of idealist philosophy and restore it to everyone, animals included, possessed of an imagination equipped to associate perceptions relations of resemblance, contiguity, and cause and effect.³ Thus, if being a Humean on personal identity means equating the self

with a system of causally related perceptions, then Hume was not a Humean.

IV. Personal identity in Hume's system of human nature

Having now re-traced the steps of Hume's explication of “the true idea of the human mind,” I shall conclude with some observations.

(i) *Simplicity*. Identity is only one aspect of Hume's treatment of complex individuals. The other is their simplicity, or individuality at any moment of their continued existence. Hume explained this according to the same model he did identity over time: the exceedingly close relation of all the coexistent parts – the facile transitions from one to the other and back again – puts the imagination in an affective disposition closely resembling that of contemplating a genuinely simple (non-composite) object. The evident fact that there is no more simplicity in the manifoldness of perception at one time than there is identity in their flux at different times (THN 253/165) we disguise by feigning “a principle of union as the support of this simplicity, and the center of all the different parts and qualities of the object” (263/171). The one thing that distinguishes the confusion of a close associative relation of coexistent ideas with perfect simplicity from the confusion of a sequence of the same relation of ideas with perfect identity is that perfect simplicity is not, like perfect identity, itself a fiction: Hume unquestionably did affirm the existence of perfectly simple impressions and ideas (e.g. minima).

(ii) *Improvement on Locke's memory-based personal identity*. Although Hume accepted (and added to) Locke's criticisms of the substantial identity conception of the self, he disagreed with his predecessor's view regarding the role of memory in the origin of the idea of the self. Though accepting (THN 635/400) Locke's view that “Consciousness makes Personal Identity” (ECHU II/xxvii/§10), Hume rejected his predecessor's contention that memory is not just a precondition for this consciousness but essential to it, even constitutive of it:

Had we no memory, we never shou'd have any notion of causation, nor consequently of that chain of causes and effects,

³Since Hume ascribed such imaginations to animals (THN 327/212-13), he may presumed to have supposed that they too have a conception of self. This is confirmed by his attribution to animals of the indirect passions of pride and humility and love and hatred – passions causally dependent on the idea of self.

which constitute our self or person. But having once acquir'd this notion of causation from the memory, we can extend the same chain of causes, and consequently the identity of our persons beyond our memory, and can comprehend times, and circumstances, and actions, which we have entirely forgot, but suppose in general to have existed. (THN 261-2/170-1)

Hume's comparison of person identity with the identity of a commonwealth (261/170) implies that personal identity persists even if, gradually over time, each memory of past experience is effaced and supplanted with others – once, twice, even a thousand times. For so long as the *effects* of these forgotten perceptions persist (or the effects of their effects, etc.), the identity made possible by the causal system of perceptions will be preserved – just as a commonwealth continues through even the most radical revolutions in its laws and constitutions (history, however forgotten, still has its ineffaceable effects). In addition, Hume's causal association based account has the advantage of excluding extraneous elements of the kind substantialist conceptions of the self inevitably end up including: perceptions that leave no causal trace behind, and so exert absolutely no influence whatsoever on the subsequent course of the causal system, and have no causal antecedents in that system, are accorded no relation to *me* even if the same substance that supports the perceptions that do belong to me supports these as well. For Hume, then, the identity of the thinking I is defined solely by a functional continuity of consciousness, and neither by ideational contents nor factors inaccessible to consciousness.

(iii) *Hume's theory of ideas idealized.* In the very first section of the first book of the *Treatise*, Hume established the manifestly causal copy thesis, that all simple ideas are copies derived from antecedent impression originals. Then, in the second section, Hume integrated this principle into a more extensive system of causal relations premised throughout his science of human nature: impressions of reflexion (desire, aversion, etc.) are usually caused by the ideas copied from impressions of sensations, these new impressions leave behind copies in their turn, and ideas copied from impressions in their turn leave behind secondary idea copies of themselves. Although Hume

predicated these causal principles on experienced constant conjunction (e.g. THN 4-5), he left it unclear through most of *Treatise* I whether these relations are associative in nature or not. Even in part iii, where he established that the ideas of necessary connection are essential to all ideas of causal relations and that these are one and all copies from impressions of reflexion immanent to acts of customary association, he refrained (apart from scattered hints) from applying this result to the copy thesis and the other causal relations he relied on in explicating such ideas as substance, space and time, existence, and of course cause and effect itself. Though almost invariably unremarked, this situation finally changes in the section on personal identity near the end of part iv of *Treatise* I, when Hume applies to the case of personal identity “What has been already prov'd at large, that the understanding never observes any real connexions between objects, and that even the union of cause and effect, when strictly examin'd, resolves itself into a customary association of ideas,” that is, a uniting principle “in the ideal world” whose “very essence” consisting in “producing an easy transition of ideas” (259-60/169). Thus, almost in passing, Hume's account of person identity settles the issue of the status of the causal relations premised throughout his science of nature in favor of their subjective ideality.

(iv) *The identity of body presupposes that of the mind.* In his explication of the identity of body – a perceived existence that continues through interruptions of its appearances to the senses – Hume argued that the only way the appearance of a perception can be distinguished from its existence is if the mind, with respect to whose senses it appears and disappears, is itself merely a product of associative imagination. For, in that case, the mind is, in itself, nothing more than “a connected heap of perceptions,” and “there is no absurdity in separating any particular perception from the mind; that is, in breaking off all its relations, with that connected mass of perceptions, which constitute a thinking being” (THN 207/138). One of the implications of this that surprisingly has seldom if ever been remarked is that the identity of body premises the identity of the mind (self, person), and that though in the order of textual exposition body

comes first, the mind precedes and makes it possible in the order of human nature.

To confirm this, we need only recall that when Hume first canvassed the possibility that the idea of the distinct existence of bodies – their external, independent existence – has its origin in “a kind of fallacy and illusion,” he insisted that, in order for this to be so, “both the objects and ourselves must be obvious to our senses, otherwise they cou’d not be compar’d by these faculties,” so that “The difficulty, then, is how far we are *ourselves* the objects of our senses” (THN 189/126). The difficulty is not to be solved by comparing external objects to our own body, since “properly speaking, ’tis not our body we perceived, when we regard our limbs and members, but certain impressions, which enter by the senses; so that the ascribing a real and corporeal existence to these impressions, or to their objects, is an act of the mind as difficult to explain, as that which we examine at present” (191/127). It is the mind (self, person) and nothing but this that is presupposed in order to perform the comparison requisite to establish the independent external existence of bodies, including one’s own. To be sure, Hume mentions the need for such a comparison only in the course of arguing that the senses cannot be the source of the idea of body, either veridically or by fallacy and illusion. But this does not make such a comparison any less necessary to forming the idea of body no matter what its faculty-origin, so that, when Hume does eventually trace its source to associative imagination, he leaves no doubt (207-8) that this is so only because that same faculty is also the source of “the true idea of the human mind.” For associative imagination then has what it needs to perform the comparison requisite to be able even so much as to distinguish, if only fictitiously and falsely (209/139), the appearance of a perception to the senses from its real existence.

(v) *The self of imagination and the self of the passions.* Hume may well have had Locke’s narrowly circumscribed, forensic sense of ‘person’ in mind when he distinguished our “identity with regard to the passions” from our memory-transcending identity “with regard to the imagination” (THN 261/170). For a person’s passions, emotions, desires, and volitions do seem to be delimited by memory – for

example, memories of the actions responsible for the pride or humility we feel in our own character and the love or hatred we feel towards others. But this raises the question of the relation of these two selves: did Hume affirm more than one sense of ‘person’, and if so, how are they correlated?

Many interpreters regard Hume’s imagination-centered account of the self in *Treatise* I, according to which there is no constant and invariable impression we can equate with the self nor any impression of a connection of particular perceptions as inhering in a self, as not only different from but inconsistent with the self of the passions described in *Treatise* II, as when he maintains “that the idea, or rather impression of ourselves is always intimately present with us, and that our consciousness gives us so lively a conception of our own person, that ’tis not possible to imagine, that any thing can in this particular go beyond it” (THN 317/206). This appearance of inconsistency vanishes, however, as soon as we recognize that Hume’s denial of an impression of the self in *Treatise* I concerns only the kind of impression that originates in the senses prior to and independently of associative imagination, not the kind that arises from its action (the affects immanent to it), which prove to be the source not only of ideas of identity but also of our ideas of necessary connection (the impressions of reflexion incident to customary association of the type founded on experienced constant conjunction). To be sure, the self of imagination is an idea, not an impression. But just as Hume sometimes accounted memories impressions rather than ideas because their vivacity is so great that their effects on thought and action are comparable to impressions (82/58), the idea of the self merits this appellation for the same reason: first, because, unlike other ideas, it is never absent from our purview (the causal relations underpinning it are continuous), so that, being always front and center, “the present situation of the person is always that of the imagination” (430/276), and, second, because the continuous presence of these associative relations serves to continuously reinforce the vivacity of this self – so that, more nearly than any other idea (even memories) it approaches the vivacity of an impression (317/206). Thus, far from conflicting with the imagination-centered account of the self in *Treatise* I, the

perpetually and intimately present self at the focus of the passions in *Treatise* II is its direct consequence.

(vi) *Hume's quandary concerning personal identity.* In an appendix published one year after the first volume of the *Treatise* (comprising books I and II), Hume confessed to having discovered a fatal flaw in his account of personal identity:

having thus loosen'd all our particular perceptions, when I proceed to explain the principles of connexion, which binds them together, and makes us attribute to them a real simplicity and identity; I am sensible, that my account is very defective, and that nothing but the seeming evidence of the precedent reasonings cou'd have induc'd me to receive it... [A]ll my hopes vanish, when I come to explain the principles, that unite our successive perceptions in our thought or consciousness. I cannot discover any theory, which gives me satisfaction on this head.

In short there are two principles, which I cannot render consistent; nor is it in my power to renounce either of them, viz. *that all our distinct perceptions are distinct existences, and that the mind never perceives any real connexion among distinct existences.* Did our perceptions either inhere in something simple and individual, or did the mind perceive some real connexion among them, there wou'd be no difficulty in the case. For my part, I must plead the privilege of a sceptic, and confess, that this difficulty is too hard for my understanding. Others, perhaps, or myself, upon more mature reflection, may discover some hypothesis, that will reconcile these contradictions.

Interpreters have long been puzzled by Hume's claim that his two principles are inconsistent. They certainly are not intrinsically inconsistent, and it is by no means obvious how the quandary he describes renders them so. Many, widely divergent solutions to this puzzle have been proposed, and I should like to finish today by proposing one in light of the interpretation of Hume's explication of "the true idea of the human mind" in *Treatise* I as a matter of

confusing distinct, yet closely resembling affective dispositions of associative imagination.

Hume's claim that the effect of his account in *Treatise* I is to have "loosen'd" all our particular perceptions follows directly from his denial that anything "really binds our several perceptions," and that "the understanding [ever] observes any real connexion among objects." For, by substituting causal association – a uniting principle in "the ideal world" of imagination – in place of the real bonds of cause and effect and inherence in a substrate posited by his predecessors (including Locke), Hume thereby left perceptions entirely loose and separate in the real world (that is, perceptions as they are prior to and independently of the fictitious identity and simplicity reflected on them by the imagination). Why does this create a problem? Hume seems not to have realized until after the first volume of the *Treatise* that there is something entailed by his thesis that the "very essence" of causal and other associative relations "consists in their producing an easy transition of ideas" (260/169) that cannot be explained associatively, namely, that the successive perceptions to and from which such transitions are made must already be present in one and the same consciousness, which therefore counts as a *preassociatively* single, unchanging, enduring existence.

The easy transitions of thought in terms of which Hume explained personal identity in *Treatise* I cannot be used to explain this numerically identical enduring consciousness for the simple reason that they presuppose it: to perceive a transition, we must perceive the perceptions concerned in it; but since these are separated in time, our consciousness of them must endure identically in order to contain them both. The exclusion of association from its explanation would thus have obliged Hume to look outside "the ideal world" of imagination for some *real* principle of connection "to explain the principles, that unite our successive perceptions in our thought or consciousness." This, however, is where his two unrecounceable principles present seemingly insuperable obstacles. For while a real substantial substrate in which successive perceptions might inhere would provide an exit from the quandary, this route is blocked by the principle that all our distinct perceptions are distinct existences. Similarly, although

directly perceptible real connections would suffice to show that present perceptions depend for their existence on past ones, this path is blocked by the principle that the mind never perceives any real connection among distinct existences. Thus did Hume find himself stymied: only an associationist account of person identity is compatible with the fundamental principles of his philosophy, but such an account is powerless to provide a way out of his quandary.