



### **Hume's Recantation Revisited**

Vijay Mascarenhas

*Hume Studies* Volume XXVII, Number 2 (November, 2001) 279-300.

Your use of the HUME STUDIES archive indicates your acceptance of HUME STUDIES' Terms and Conditions of Use, available at

<http://www.humesociety.org/hs/about/terms.html>.

HUME STUDIES' Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the HUME STUDIES archive only for your personal, non-commercial use.

Each copy of any part of a HUME STUDIES transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

For more information on HUME STUDIES contact [humestudies-info@humesociety.org](mailto:humestudies-info@humesociety.org)

<http://www.humesociety.org/hs/>

# Hume's Recantation Revisited

VIJAY MASCARENHAS

In the Appendix to the *Treatise of Human Nature*, Hume famously recants his position on personal identity. There he confesses: “upon a strict review of the section concurring *personal identity* I find myself involv'd in such a labyrinth that . . . I neither know how to correct my former opinions, nor how to render them consistent.”<sup>1</sup> By his own admission, then, something has gone wrong in Hume's account of personal identity, something that, at least to his eyes, did not go wrong in his accounts of body and necessary connection. For those accounts were not grossly inconsistent or patently absurd. The case is different, however, with personal identity. There his philosophical enterprise suffers shipwreck, and it is important to understand why. Unfortunately, however, Hume confesses *that* he finds his former opinions false as well as inconsistent, but neglects to specify *how* or *why* he came to this conclusion. This paper is an attempt to address just that question.<sup>2</sup>

First we must observe the general philosophical tasks that Hume takes on in the *Treatise*. These I take to be three. First, he must assess and delineate the landscape of the human mind. In what does consciousness consist? Here his answer is simple: perceptions and perceptions alone, that is, impressions and ideas. To this Hume adds only a handful of associative processes by which the mind navigates from perception to perception: these are contiguity, resemblance, and causation. Having assayed the psychological apparatus available to the human mind, Hume's second task is to determine whether certain beliefs—i.e., those concerning body, causation, and personal identity—are epistemologically justified, that is, well founded and rationally grounded. His answer, of course, is no.<sup>3</sup> No matter how irresistible and useful such be-

liefs may be, nothing available to human experience justifies our accepting them. The last task Hume takes upon himself is more properly psychological than philosophical: this is to provide an account of how we are able to form such beliefs given their irrationality and given the rather narrow confines of human experience. Impression, ideas, and the association of ideas are all that Hume can rely upon in constructing an account of the psychological origin of these beliefs.

Now, I believe that Hume was able to deal fairly deftly with beliefs in body and necessary connection. Those beliefs were unfounded, but their psychological origin could be explained by “idealizing” them, that is, by attributing their formation to the association of ideas. There is no real necessary connection between perceptions, but we can come to believe in causality, because we associate the *ideas* of those perceptions. Similarly, perceptions do not enjoy continued existence when unperceived, but we come to believe they do because of the easy transition between the ideas of those perceptions. In short, Hume’s tactic seems to be to deny real connections between perceptions and then to resolve them into associations between the *ideas* of these perceptions. This ploy, however, fails him when it comes to the issue of personal identity.

To see exactly why this is so, one must turn first to the “former opinions” that Hume ultimately came to reject. In Book I, he asks how we can form the belief in personal identity despite the absence of any constant and invariable impression of the self. To arrive at an answer, he says “we must take the matter pretty deep, and account for that identity which we attribute to plants and animals; there being a great analogy betwixt it, and the identity of a self or person” (T 253). Here Hume takes a wrong turn in the labyrinth: in positing a “great analogy” between personal identity and the identity of animals, plants, and the like, he adopts an objective view with regard to something essentially subjective, the self. But more on this later. For now, let us follow Hume, even where he strays. Concerning the identity of plants and animals, he mostly just repeats what he said in the section on body: the only truly identical things are those perceptions that enjoy constant and invariable existence; but we often attribute a false identity to a succession of perceptions where they are so related as to give the imagination a smooth transition between the ideas of these perceptions. Now Hume believes that a similar smooth transition is at work in the formation of the belief in personal identity; and the two relations he cites as responsible for this smooth transition are resemblance and causation.<sup>4</sup>

Let us first take resemblance. Hume asks us to “suppose we cou’d see clearly into the breast of another” (T 260). In that case we would be privy to

a succession that includes frequent repetition of resembling perceptions due to the operation of memory, which churns out copy after copy of impressions and ideas. The resemblance between these perceptions allows for the easy transition of the imagination, which, in turn, leads to the ascription of identity. As Hume puts it:

For what is memory but a faculty, by which we raise up the images of past perceptions? And as an image necessarily resembles its object, must not the frequent placing of these resembling perceptions in the chain of thought, convey the imagination more easily from one link to another, and make the whole seem like the continuance of one object? (T 261)

The most obvious problem with this account, however, is that many, if not most, of our ideas and impressions do not resemble each other. True, once I have an impression, the ideas I form when recalling it by means of memory or imagination will resemble each other and the original impression. True, also, that when I survey my perceptions (as if I were peering into my own breast), I would find an easy transition between the idea of the initial impression and those of the subsequent recollections or copies. But the resemblance here would only be among these memory-ideas and between them and the original impression. For each non-resembling original impression, there would be a separate and distinct chain of memory-ideas; ideas within each chain would indeed resemble each other, but cross-chain ideas would only resemble each other in the case that their initial impressions did. Personal identity, however, is supposed to cover the entire ambit of perceptions, not merely some subsets. Hume gives no reason why the resemblance of *some* of my perceptions should make *all* of them seem like "one object," nor why the imagination should find easy passage through the rough terrain of non-resembling perceptions that lies between the smoother regions of resembling perceptions.<sup>5</sup>

The other relation that Hume thinks helps dupe the mind into thinking of itself as constant and invariable is causation. "The true idea of the human mind," Hume says, "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" (T 261). Now this is a rather curious claim coming from someone who spent more than an entire chapter quietly destroying the notion of power or causality, at least in so far as those notions imply a necessary connection between perceptions. We only feel a determination to pass from idea to idea and then

project this impression upon the perceptions themselves, attributing to them a necessary connection whose foundation is in the operations of the mind rather than in the perceptions themselves. And Hume also took great pains to prevent his readers from thinking that by his idealization of the causal relations he meant that there was anything in the mental realm that would answer to the beliefs in necessary connection we apply, erroneously, to the objective realm. Thus he warned: "The uniting principle among our internal perceptions is as unintelligible as that among external objects . . ." (T 169). By characterizing the human mind as a system of causally related ideas and impressions, Hume could not have meant that these perceptions were bound together by necessary connections or that they somehow overcame their distinctness from each other. The causal connections among the mind's perceptions must then be of a type congruent with his new definition of cause and effect. Now Hume gives two such definitions. One defines a "cause" as "an object precedent and contiguous to another, and where all the objects resembling the former are plac'd in like relations of precedence and contiguity to those objects, that resemble the latter"; the other defines it as "an object precedent and contiguous to another, and so united with it, that the idea of the one determines the mind to form the idea of the other, and the impression of the one to form a more lively idea of the other" (T 170). Hume claims that these two definitions are essentially the same, and, given his philosophical system, they are indeed so. The former merely focuses on the relations between the perceptions themselves, the latter upon the effect of these perceptions on the mind. But as it happens the human mind is such that where perceptions come in certain relations, i.e., contiguity, constant conjunction, etc., it is affected in a manner consistent with the second definition of causation.

The question then is how exactly these causal relations could lead the imagination to conceive of the self as constant and invariable. Unfortunately, Hume does not explicitly specify the mechanisms by which this result is to be achieved. One may presume without too much danger, however, that the power of causal relations lies in their producing an easy transition of the imagination; after all, Hume repeatedly states that the very essence of relations, of whatever type, lies in their producing an association of ideas in the imagination.<sup>6</sup> Barry Stroud agrees with this interpretation. Of a person surveying his perceptions, he writes, "if his thought 'slides easily' along that series of perceptions because of the causal connections believed to hold among them, then he will get the idea of himself as one mind."<sup>7</sup>

The problem with this account of the belief in personal identity is that the easy transition afforded by causal relations is called upon to do double

duty. For, from the above definitions of cause, it can be seen that the belief in causal relations is itself conditioned upon the ease of transition produced by constant conjunction of resembling perceptions with other resembling perceptions. After the mind is suitably conditioned by the repetition of effect following cause, "the idea of the one determines [it] to form the idea of the other, and the impression of the one to form a more lively idea of the other" (T 170). But now it seems Hume wants the ease of transition to do double duty, that is, to produce both a determination in the mind which is the origin of the belief in causation *and* a conflation of the perceptions which is the origin of the belief in personal identity. It is too much to ask of one easy transition, however, that it produce two beliefs. This is especially so as the two beliefs in question are contradictory. If one believes perception *A* to be the cause of perception *B*, then it is very unlikely indeed that one would also believe the two perceptions to be identical; for, we normally do not tend to regard effects as being the same as their cause; indeed, we tend to view them as quite separate and distinct.

### Standard Critiques

Hume's attempt to account for the psychological origin of the belief in personal identity thus does not meet with the same success that he enjoyed where necessary connection and body were concerned. In those cases, he presented accounts that were plausible, coherent, and did not conflict with his general philosophical commitments. By his own admission, however, his treatment of personal identity is lacking. Though we must be thankful for this candid self-criticism, we may still fault Hume for confessing *that* he finds his former opinions untrue as well as inconsistent, while neglecting to say *how* or *why* he finds them so. There may be some discussion of propositions that are impossible to render consistent, but most commentators agree that the obscurity of the recantation precludes any sure interpretation of why Hume chose to repudiate the beliefs he once held. We cannot then be certain that Hume was at all aware of the problems involved with basing the belief in personal identity upon resemblance and causation that have been outlined above. Why then did Hume recant? Or, more broadly, where exactly did he go wrong with his account?

Many Hume scholars have found the discussion of personal identity to be deficient and this not merely due to its rather unsettling conclusions. Though each one has a different description of the deficiency, they do fall into a handful of major camps. Perhaps the most prevalent criticism is that Hume denies the reality of personal identity in Book I yet explicitly uses this

concept of the self elsewhere in the *Treatise*, namely, in Book II.<sup>8</sup> There, much as if he had forgotten the conclusions of the last book, Hume writes such things as “[the object of pride is] the self, or that succession of related ideas and impressions, of which we have intimate memory and consciousness” (T 277). But the contradiction is in fact only seeming. As Terence Penelhum notes, Hume does not deny that we have *any* idea of self, only that we have an idea of self “after the manner it is here explain’d” (T 251), i.e., as something constant, invariable, simple and identical.<sup>9</sup> The idea of self which surfaces in Book II is not this type of idea of self; it is not, as Penelhum puts it, “the idea of the pure ego constructed by the rationalist philosophers.”<sup>10</sup> After all, Hume is quite careful to define the self as a “succession of related ideas and impressions” and not as a self-identical substance that unifies all perceptions. There is, however, at least one other passage that cannot be explained away so easily. Further on in Book II, Hume declares: “’Tis evident, 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” (T 317). Now this is a rather strong statement, one, in fact, that resembles exactly the type of statement Hume discredited in his chapter on personal identity; for, here something very much like a constant and invariable impression of self is posited; and it was precisely this kind of idea of self that Hume denied we ever had.

Some commentators such as Lawrence Ashley and Michael Stack have striven to exonerate even this passage from the charge of contradiction by interpreting the phrase “after the manner it [the idea of self] is here explained” in the chapter on personal identity to mean only “a self which has *perfect identity and simplicity*.”<sup>11</sup> But on the same page Hume expressly denies that there is any impression constant and invariable throughout our lives, and he begins the chapter by stating, “There are some philosophers, who imagine we are every moment intimately conscious of what we call our SELF” (T 251), a claim that he most obviously seeks to dispute. Now if we were truly to have *an* impression that was *always intimately present*, I cannot see how it could be anything other than a constant and invariable perception, i.e., one of which we were continually conscious. Hume would thus be flatly contradicting himself. Fortunately, Don Garrett has supplied an alternative interpretation of the troubling passage: Hume does not mean that there is a single impression (of the self) that remains constant and invariable; all that is required is that at any one time there be “an idea of a bundle of perceptions that are related by resemblance and causation.”<sup>12</sup> Just as with the idea of space there is not one impression of space that is constant and invariable, but rather

many spatial impressions, any of which could, when used as an abstract idea, give rise to the idea of space, the idea of self can be formed from a variety of impressions. In fact, it could be formed by any impression at all, since all impressions are supposed to be included in the bundle of perception constituting the self.<sup>13</sup> Besides clearing Hume of obvious contradiction, this interpretation has the advantage that it is precisely this kind of self that is integral to Hume's arguments concerning the love of fame, pride, humility, etc.<sup>14</sup>

The second criticism is that even in Book I Hume relies *implicitly* on the very conception of self that he denies and, worse still, that this surreptitious use extends to the very arguments by which that conception of self is supposedly denied. Wade Robison, for example, accuses Hume of making illicit recourse to a substantial or at least identical self whenever he has the mind perform some activity—associating, observing, mistaking, confounding, etc.—that it could not possibly perform if it were merely a loosely related bundle of perceptions.<sup>15</sup> The human mind, the argument goes, cannot possibly confuse the diversity of its perceptions with the simplicity of self unless it were, in fact, an integrated whole. But even if this claim is true, it cannot be used to charge Hume with inconsistency. This is because Hume himself does not think that the existence of mental activity necessarily implies the existence of a metaphysically simple *actor* behind the action. Mental feats, as Garrett notes, are, on Humean analysis, reduced to the “occurrence of particular kinds of perceptions.”<sup>16</sup> The attribution of a necessary connection between perceptions is nothing more than the propensity for the idea of the cause to occur after the idea or impression of the effect has occurred; similarly, the belief in personal identity may amount to nothing more than the mere fact that perceptions in question are so related by resemblance and causation that the mental state of surveying them is virtually the same as that of regarding a single identical perception.<sup>17</sup> None of this is to say that these defenses are themselves defensible, i.e., that the theory of mind they embody is in the end sound. It only means that Hume has armor enough to protect himself against the salvos of those who would accuse him of gross self-contradiction.

The absence of proper criteria by which one could individuate persons is another criticism often waged against Hume. If the self is a bundle of perceptions, these critics ask, how exactly are we supposed to distinguish which perceptions belong in which bundles; or, phrased another way, if the self is supposed to be “that to which all perceptions have a reference,” how are we to determine the extend of the “all.” Surely, some perceptions must belong to you and some to me. Now it is quite true that Hume never even

attempts to offer a means of individuation, either for the set of perceptions to which our belief in personal identity is supposed to refer, or for the set of perceptions available as psychological material for the formation of the belief. This failure to respond would be damning, indeed, if only he had himself posed the question others take him to have posed. But many readers, including Penelhum, feel that the chapter on personal identity concerns something quite different from mental individuation.<sup>18</sup> After all, the central question of the chapter is not, how do we come to believe that so-and-so is the same today as he was yesterday, but “what then gives us so great a propension to ascribe an identity to these successive perceptions, and to suppose *ourselves* possess of an invariable and uninterrupted existence thro’ the whole course of *our* lives?” (T 253, emphasis added). Now one could suppose that Hume’s use of the first person plural is merely impersonal, i.e., he means why do we believe that ourselves and others (included in the “we”) maintain personal identity; and Hume does often speak as if his concern were third person rather than first person ascriptions. All the same, the general tenor of the chapter seems to indicate that he was primarily interested in first person singular ascriptions of identity, or, to borrow Penelhum’s terminology, in “mental unity” rather than “mental individuation”: the question is “what is it that makes *me* think *I* am entitled to ascribe unity to the perceptions that made up *my* mental history?” rather than “how one distinguishes one mind from another.”<sup>19</sup> Moreover, the curious fact that all perceptions seem to come to us already bundled up in discrete units, your perceptions never creeping into my experience or vice versa, coupled with Hume’s phenomenalistic empiricism, render moot the very question of individuation: what needs to be explained is not how we can individuate our minds from others but why it is never necessary to do so. This may well mean, as some have contended, that Hume must implicitly assume some type of robust personal identity in order to be certain that every perception in the succession of perceptions that I believe to constitute my self is in fact *my* perception to begin with. It does not mean, however, that Hume has failed to provide an explanation for a problem that he himself thought it was his task to address.

The last criticism often launched against Hume is that his treatment of personal identity founders because he, like Descartes before him, treats persons as purely mental entities, excluding the body from all discussion. This claim is usually given as the reason behind Hume’s failure to provide proper criteria for the individuation of person.<sup>20</sup> As Penelhum puts it: “It is at least extremely likely that the reference to the public world of separate human bodies has something essential to do with our capacity to distinguish persons from one another.”<sup>21</sup> But Penelhum correctly adds that however true

this may be, it has little bearing on how we come to believe we ourselves are constant and invariable over time. Moreover, in denying that we have actual experience of the continued and distinct existence of unperceived objects, Hume denies that we have experience of our having bodies in the sense that we normally mean that claim. Each time I see a certain someone I experience a brand new impression; and it is only because this impression resembles the one I had some time ago that I believe I am seeing the same person I saw and that he or she continued to exist in the meantime. The situation is similar with regard to our own bodies. The first time Hume addresses the problem of personal identity is not in the chapter so named but in the chapter concerning body; there he correctly identifies the intimate connection between personal identity and body by remarking that the question of how we can perceive objects as distinct from ourselves really amounts to the question of "how far we are ourselves the object of our senses" (T 189). And the answer to this is: not at all. Any metaphysical, substantial self would be, by definition, insensible; and any physical self would also be insensible, because, according to Hume's phenomenalism, "'tis not our body we perceive, when we regard our limbs and members, but certain impressions, which enter by the senses" and do not enjoy "real and corporeal existence" (T 191). Without this type of existence bodily impressions cannot possibly serve as a foundation for beliefs concerning either mental unity or mental individuation. Mere spatial location without the other characteristics usually attributed to body is also inadmissible. In his famous theater metaphor for the mind, Hume carefully avoids possible misunderstandings by informing his readers that his comparison is not quite exact, there being no analogon in the human mind to the physical stage that furnishes a permanent ground upon which the continual change of settings and scenes transpires; in the mind, Hume cautions us, there "are the successive perceptions only . . . 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" (T 253).

This is entirely of a piece with Hume's treatment of space. Space is neither a perception nor something which preexists perceptions and into which they are placed, but rather an idea arising from the manner in which certain perceptions, visual and tactile, come to the mind. Spatial location, or place, cannot be separated from the impressions that come to us in a "spatial manner"; and, since the spatial location of bodily impressions is on all fours with that of other impressions, it cannot play the privileged role of binding together all perceptions, whether spatial or not. The impression of my hand may well be in certain spatial relations with my impression of my desk, but I cannot say that the one provides the basis upon which both are attributed to

the same self; nor can I say that the one belongs to the other in the way some think perceptions belong to the self. Far from committing an unconscious blunder, therefore, Hume is acutely aware of the mutual conceptual dependence of personal identity and body (and/or spatial location) and furthermore realizes that he *cannot* look to the latter to play any role in the formation of the belief in personal identity without contradicting the phenomenalist empiricism upon which he bases all of his reasonings. Again, this does not mean that he succeeds in his tasks, only that he has not fallen into gross inconsistency and that his resistance to employing the body is not only justified but inevitable, given his prior philosophical commitments.

### Toward a More Penetrating Critique

What then is the problem with Hume's account of personal identity? The critical blunder, I believe, comes in his insistence on conceiving the identity of the self as being analogous to that of plants, animal, churches, rocks, and rivers, that is, on conceiving the self in essentially objective terms. The nature of the blunder becomes clear if we imagine what would obtain if resemblance and causation could produce so smooth a transition in the imagination that perceptions are run together and taken to constitute a single, self-identical entity. Would not this entity be an object like any other object and, hence, something other than the self? If I see two resembling impressions interrupted by a non-resembling one (e.g., I blink), then the resemblance of the two impressions may make me believe that they are one constant and invariable impression, but I would certainly never take that one impression to be *me*. Resemblance, in other words, cannot be expected to produce both the belief in the distinct and continued existence of an object and the belief in personal identity.<sup>22</sup>

Hume's objectification of self finds expression in the fact that, though his question is posed in subjective terms—what leads me to conceive of myself as an enduring, invariable existence—his answer oscillates between first and third person perspectives. This sleight of hand reaches its craftiest when Hume asks that we suppose ourselves able to peer “into the breast of another, and observe that succession of perceptions, which constitutes his mind or thinking principle” (T 260). The problem here is that the perceptions in question become both his and ours: *ex hypothesi*, they are his; but they are also ours since we too experience them and they therefore must be elements of the set of perceptions that constitute ourselves. This confusion almost seems designed to lull the reader into forgetting that the smoothness of transition produces beliefs about objects not about subjects. It might seem

reasonable that if I were to have a private screening of someone else's perceptions and there observed a resemblance between them, that I might take all these perceptions for one object, that person's self. On closer examination however, things turn out differently. If I were really to peer into Mr. *A*'s breast and see an impression of Ms. *B* followed by several memory-ideas of this impression, no amount of resemblance between these perceptions would get me any closer to forming a belief about Mr. *A*'s personal identity. The resemblance of Mr. *A*'s perceptions may well produce an easy transition for the imagination; but such easy transition would make me believe in the constant and invariable existence of Ms. *B* rather than Mr. *A*. After all, I would not have had even a single perception which could answer to the description "Mr. *A*," whether he be constant and invariable or not. Hume's argument seems to ask us to imagine that we see the resembling perceptions and simultaneously notice that the perceptions are Mr. *A*'s perceptions and not our own, so that the identity we attribute will be to Mr. *A* and not to Ms. *B*. But as the original question concerned the personal identity of Mr. *A*, Hume seems to have caught himself in a *petitio principii*, assuming at the outset that we already have a working belief in the constant and invariable existence of Mr. *A*.

Now we may well be mistaken in thinking Hume was talking about *our* belief in Mr. *A*'s identity. Perhaps the issue at hand is Mr. *A*'s beliefs about his own self. Even more damaging results, however, follow upon this interpretation. The set of perceptions which Mr. *A* is supposed to run together and conceive as a single entity covers *every* impression and idea he has ever had. The object whose identity is in question here would therefore include everything that Mr. *A* has ever seen, heard, touched, tasted, felt, remembered, or imagined. In other words, he would come to believe that the entire universe was a single, constant and invariable entity; and, worse still, he would take this entity to be himself. Remember that Hume believes that impressions are not to be purely subjective mental events which can be distinguished from independently existing objects but to be the very objects themselves. For the imagination to conceive *all* perceptions, and this includes impressions as well as ideas, to be a unified, self-identical thing is, therefore, for it to conceive the entire world summed up into single thing. If Hume's account of personal identity were to succeed, I would have to regard all the impressions I have ever had, of mountains, plants, animals, churches, rivers, and the like, to be one object, and I must furthermore take that all-inclusive object to be myself.<sup>23</sup> But this is clearly absurd.

Now the problems I have just outlined stem from treating the self as if it were an object. But what exactly is it about the self that resists such treatment? One difference lies in certain unresolvable complications in fixing the extent

of the set of perceptions to be run together in constituting the belief in personal identity. These complications stem from the entanglement of self-reflexivity and temporality that is essential to self-conscious experience. In a nutshell, the problem is that while it is possible for there to be a bounded set of impressions which my imagination unites into an objective identity, this is not so where the identity in question is my own. The bundle of perceptions that constitutes myself must include all impressions and ideas I ever have had, ever will have, and, most crucially, all those that I am now having.

Let us reverse the course of time start with the future. When I form a belief in the identity of an object, at least according to Hume, I am merely affirming that a particular impression is the recurrence of another impression.<sup>24</sup> In this account there is no development of a concept of the object that bears within itself criteriological specifications by which the object may be reidentified in the future: there is simply the confusion of one perception for another, or, perhaps, for a set of other perceptions. This set will always be well defined, containing a finite number of impressions that one has received and bearing no implications for future impressions. Thus if I were to peer into the breast of another, I would presumably perceive a finite set of perceptions and my task in creating an imaginative union of these perceptions seems fairly easy to accomplish. True, as long as the person is still alive and conscious he or she will be continually adding new perceptions to the bundle, but I am not required to conjoin these additional perceptions to the original series; for the act of forming my belief freezes the continual flow of perceptions to be bundled together, and the content of my belief covers only those perceptions I have seen *up to that point*. There is no inherent implication in my ascribing identity to this particular set of perceptions that new perceptions will accrete to that same identity. The case seems wholly different, however, when it comes to my own identity, that is, to the subjective identity of my own mind. As long as I am alive, as long as I am conscious and capable of forming beliefs about myself or others, it would seem that any identity I ascribe to myself could not possibly halt right at the moment I form it—as if I were to repeat every few minutes, all this up till now was myself, all this up till now was myself.<sup>25</sup> Personal identity, rather, must be something inherently open to the future: if I know I *was* myself, then I must know I *am* myself, and if I know I *am* myself, then I must know I *will be* myself (if I am not struck dead). In Humean terms, if I imagine all my past perceptions to be part of the bundle, then I must also imagine that any future perceptions I may be fortunate enough to have will also become part of that bundle. The set of perceptions which must be united to form my idea of self must therefore be inherently open to the future.

Now let us skip for the moment the present and proceed directly to the past, for the past will lead us nicely to the present. Past perceptions are just that, past, i.e., nonexistent, irretrievably absent. Again, Hume's talk of looking into the breast of another almost seems designed to paper over this fact. In imagining oneself to survey someone else's perceptions there is a tendency to place what necessarily comes in temporal succession, side by side, so to speak, in a static display. With the perceptions arrayed in this manner, I may survey them at my leisure, placing past perceptions next to present ones and observing any relations of resemblance that might exist between them; I may even, under these conditions, go back and forth between the present and the past, or the recent past and the most remote at my leisure. But this is not at all the case when it comes to my own perceptions. Here the strictures of temporality are inviolate: a perception is a one-shot deal; once it is gone, it can never come back. The irretrievability of perceptions means that we cannot simply pluck a past perception from its oblivion and compare it to present ones.<sup>26</sup> This does not mean that we never recall or remember our past impressions; rather, it means that the act of recollecting past perceptions consists in the present production of memory-ideas that, however much they may resemble the past perceptions, are each entirely new and distinct existences. Given that memory operates in this manner, it becomes impossible to have a neatly bounded set of perceptions that we can inspect for relations of resemblance and then conflate into the idea of self. New perceptions that must themselves be included in what we would take to be our constant and invariable self are produced by the very survey that Hume supposes produces our belief in personal identity. If the self were an object it would be a slippery object indeed, one that, like the fabled particles of quantum mechanics, is changed by the very act of observation. The self-reflexivity of personal identity means that the very ideas by which the belief is formed must be included in the content of that belief.

Nelson Pike, however, believes that self-reflexivity is not really a problem for Hume. He starts with a defense against certain attacks made by MacNabb, who asks, somewhat incredulously: "How can a series of conscious states be aware of itself as a series?"<sup>27</sup> This conundrum, Pike thinks, does not pose a true threat, for Hume does not need a *series* that is conscious of itself as a series, but merely a series that contains at least some perceptions which "are aware of themselves as members of the series."<sup>28</sup> While MacNabb professes himself confounded by the question of what it would be for such self-aware perceptions to exist, Pike thinks he has found the answer: "A series of conscious states might *contain* an awareness which is *of* itself presented as a series," he contends, much in the same way that "a collection of pictures

might *contain* a picture which is of itself pictured as a collection."<sup>29</sup> It seems however, that Pike has chosen a particularly infelicitous analogy; for, though a gallery may well contain a picture that depicts the gallery itself (there are some actual examples of this in the Louvre as well as other museums), the picture can contain representations of every picture in the gallery but one, namely, itself. Pike's gallery analogy, therefore, backfires.

In any case, I believe the whole analogy to be somewhat beside the point. In Hume's account of the origin of the belief in personal identity there is no idea of self over and above those perceptions that form the content of that belief. Just as in the case of objective identity there is a conflation of perceptions but no *additional* perception that would function as the idea of that object, the belief in personal identity arises due to a confusion of perceptions and is neither instantiated in any single one of them, nor in some new perception. There is, then, no picture that is *of* the gallery itself. The problem therefore is not so much that the idea of self must itself be included in the set of ideas and impressions which are taken to constitute the self; the problem, rather, is that there are certain unavoidable psychological processes involved in the formation of that belief which produce perceptions that must themselves be included in the bundle to which the belief refers.<sup>30</sup> In trying to form beliefs about itself, the mind is always one step behind itself, like a dog trying to catch its tail, or a boy, his shadow.

But was Hume himself aware of these difficulties? And is this why he recanted in the Appendix? To answer these questions, let us follow Hume along his path up to the point where he loses himself in the labyrinth. In Book I, Hume writes:

every distinct perception, which enters into the composition of the mind, is a distinct existence, and is different, and distinguishable, and separable from every other perception, either contemporary or successive. But, as, notwithstanding this distinction and separability, we suppose the whole train of perceptions to be united by identity, a question naturally arises concerning this relation of identity; whether it be something that really binds our several perceptions together or only associates their ideas in the imagination. That is, in other words, whether in pronouncing concerning the identity of a person, we observe some real bond among his perceptions or only feel one amongst the ideas we form of them. (T 259)

The first half of this passage amounts to a paraphrase of the two propositions that Hume will find inconsistent by the time he writes the Appendix. These are:

(1) all our distinct perceptions are distinct existences;

and

(2) the mind never perceives any real connexion among distinct existences.

Given this, there can only be one answer to the question Hume poses in the second half of the passage. Thus he concludes:

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.

Here at least Hume is full of confidence. Is personal identity *real* or *imagined*? For it to be real there would have to be a "real" connection binding perceptions together that the mind can observe. If this were the case, if we did indeed observe a real connection amongst perceptions, then belief in personal identity would be epistemologically justified. But Hume has already denied this: all perceptions are distinct existences and there is nothing other than perceptions that could constitute the mind.<sup>31</sup> The mind therefore cannot be anything other than a collection of distinct existences and accordingly cannot *observe* itself to be anything other; though it may well *imagine* that it has a constant and invariable existence. So much so good. All that there is left for Hume to do is specify what *imaginative*, as opposed to *real*, relations between perceptions lead to the false ascription.

Now let us turn our attention to the Appendix. There Hume repeats many of the propositions that led him to the conclusion he reached in Book I, apparently still satisfied with the validity of the propositions themselves, if not the conclusion. Hume, however, follows up this reaffirmation with a gloomy note: all is well up to this point, he says, but he can go no farther for he has come upon an insoluble conundrum. He writes:

But having thus loosen'd all our particular perceptions, when I proceed to explain the principle 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. (T 635)

Or alternatively:

But all 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. (T 635-6)

It is directly after this that Hume mentions the two principles that he cannot “render consistent.” Now, of course, there is nothing whatsoever inconsistent about the propositions themselves; (2) is, in fact, the epistemological corollary of (1). The hunt was then on for some third principle that conflicted with the two, either taken singly or together. The natural candidate, suggested by Passmore and others, is what would be negation of the conclusion if the two principles were major and minor in a syllogism, i.e., “the mind perceives a real connexion among our perceptions.”<sup>32</sup> But this proposition is so much at odds with Hume’s overall philosophy that he denies it virtually once or twice each page of the entire *Treatise*. It does not, then, seem a likely candidate for the illusive *tertium quid*.

But no matter how unlikely it would seem, I do believe Passmore was on the right track: if Hume does not feel himself compelled to assent to the fact that the mind perceives real connections amongst perceptions, he does feel compelled to assent to something quite like it. The two passages quoted above lend support to this interpretation. The central question in Book I, besides *why* we attribute identity to the self, was *whether* we perceive a real bond between perceptions or “only associate their ideas in the imagination.” There the answer was clear. Hume rejected the first alternative and was left with the latter. In the two passages above, however, the question has shifted. Hume now seeks to explain (but cannot) “the principle of connexion, which binds [our perceptions] together, *and* makes us attribute to them a real simplicity and identity” (T 635, emphasis added). What was once an “either/or”—*either* there is a real bond between perceptions *or* we simply attribute a false identity to them—here becomes an “and”: there is something that makes us attribute identity to our perceptions *and* this very same thing binds our perceptions together. Here in the Appendix, Hume seems to be assuming exactly what he denied on virtually every page of main text, that there are real and perceivable connections between perceptions. And this is not the only passage in which he seems to be doing so. Even in Book I, he denies that there are any real connections between perceptions upon the grounds that “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 imagination, when we reflect upon them" (T 260, emphasis added). But this statement assumes that even if there is no true union between cause and effect there is, indeed, one between their ideas; are not ideas perceptions too, and if so, does this not amount to a "real" connection between perceptions? The apparent contradictions reach their climax in the Appendix where Hume writes: "No connexions among distinct existences are ever discoverable . . . we only *feel* a connexion or a determination of the thought, to pass from one object to another," and "when reflecting on the train of past perceptions, that compose the mind, the ideas of them *are felt* to be connected together" (T 635, emphasis added). What more do we need to have an impression of a connection between perceptions than to *feel* it? Moreover, for Hume, I cannot *feel* a connection unless there actually is a connection. To feel a connection is to have an impression of it; and, since impressions can neither lie nor deceive,<sup>33</sup> to have an impression of a connection is for there to be a connection. There is, then, much textual evidence that Hume does indeed come very close to saying that the human mind can perceive connections amongst impressions.

But however much Hume is forced toward this conclusion, he cannot accept it. And, I would assert, he cannot accept it, not simply because it contradicts his entire philosophical project as summed up the two supposedly "inconsistent" propositions, but also because, given his atomistic conception of perceptions, he cannot begin to understand the type of connection between perceptions that, by the time he came to write the Appendix, he realized he must posit. Remember that Hume denied the existence of "real" relations by resolving them into "ideal" relations: there is no real necessary connection, for example, between the impressions of cause and effect, only an "ideal" or associative one between the *ideas* of those impressions. This was how he explained the formation of beliefs in body and causation despite the fact that there were no impressions that could serve as an epistemological justification for those beliefs. When it comes to the belief in personal identity, however, Hume reaches a dead end in the labyrinth, for his explanatory apparatus wraps around itself: the ideal relations into which real relations are resolved themselves rely on something that not only explains the formation of the belief in personal identity, but would amount to, if not a constant and invariable self, then at least some kind of unity of consciousness. In other words, "ideal" relations, while not bearing the same characteristics as "real" relations do (or, would, if they existed), still presuppose "uniting principles" that allow perceptions to be associated in the first place. Association and easy transition can occur, for example, only between ideas that belong to the same mind.<sup>34</sup>

Hume's predicament is brought to the fore when he confesses that he cannot explain "the principles, that unite our successive perceptions *in our thought or consciousness*" (T 636, emphasis added). Now the first thing to be noted about this phrasing is that the question is not *whether* there are indeed principles that unite our successive perceptions but *how* these principles are to be explained. The second, as Wayne Waxman notes, is that the phrase "in our thought or consciousness" is ambiguous: it could refer either to principles that unite those perceptions that constitute the mind or to ones that merely make us *think* those perceptions are united because their corresponding *ideas* are associated in the imagination.<sup>35</sup> In other words, the connections could either be "real" or "ideal." But the ambiguity is itself obscured because of the peculiar character of the problem of personal identity. For whatever connections or relations between ideas of perceptions that make us *think* the perceptions are united, actually unite those ideas; true, the ideas preserve, in some sense, their distinctness—they are still associated ideas and not one idea—but the question here is not of the unity and constancy of a particular perception, nor even of the conflation of all perceptions into one constant and identical object, but rather of there being enough unity of consciousness to account for the association of ideas, that is, "ideal" connections. Hume, I believe, came to realize this. That is why he speaks of the "union" of ideas in imagination, claims we "feel" (i.e., have impressions of), connections between perceptions, and, most tellingly, says that there are "uniting principles" that both "bind [perceptions] together, and make us attribute to them real simplicity and identity" (T 635).

But however much Hume might recognize the need to account for the unity of consciousness as the basis of the associability of perceptions, his atomistic conception of perceptions cuts off such an explanation. For, even if Hume posited some kind of impression of a connection between perceptions, this, it itself, would not be enough to bind perceptions together. The impression of the connection would itself be a perception; and this perception, in turn, would have to be "connected" to the set of perceptions that are included under the belief in personal identity. There could, of course, be another impression of *this* connection, but the regress here is obvious. Impressions, ideas, and the association of ideas alone can never explain the unity of consciousness (i.e., "the principle of connexion, which binds [perceptions] together") necessary for the association of ideas in the first place.

The self-reflexivity of personal identity then reveals the contradiction at the heart of Hume's philosophy: he cannot explain what he uses to explain everything else, ideal relations or associations. And he cannot explain these because he cannot explain "the principles, that unite our successive

perceptions . . ." (T 636). It is these principles of connection, I believe, that are the best candidate for the *tertium quid*. After all, it is directly after mentioning these principles that Hume proceeds to cite the two principles he cannot render consistent; perhaps he means not that they are inconsistent with each other but that they are inconsistent with what he has just written about: the principles that unite thought. Now Hume cannot bring himself to give up either of his two empiricist principles; they form the backbone of his philosophy. But he realizes that they prevent him from explaining what unites our successive perceptions *and* "makes us attribute to them real simplicity and identity" (T 635). In other words, without abandoning his two principles—and their emphasis on the distinctness of perceptions—he cannot explain how it is that the human mind is united enough even to form the belief in personal identity. His explanatory apparatus—ideal relations—here enters into contradiction with his phenomenalist empiricism—the conception of perceptions as being so distinct and separable that they have no inherent relation to each other. This is what led Hume to such despair that he threw up his hands, confessed the difficulty too hard for his understanding, and hoped that another philosopher might lead the way out of the labyrinth and "discover some hypothesis, that will reconcile those contradictions" (T 636). That philosopher, I believe, was Immanuel Kant. Only he did not so much reconcile the contradictions—between phenomenalist empiricism and the conditions of the possibility of imaginative association—as use the latter to destroy the former.

## NOTES

1 David Hume, *Treatise of Human Nature*, ed. L. A. Selby-Bigge, 2nd ed., revised by P. H. Nidditch (Oxford: Clarendon Press, 1978), 633 (hereafter, referred to as "T").

2 This, of course, is a favorite question of Hume scholars and many answers have been proposed with no consensus in sight. I venture forth, nonetheless, into this territory because, though well-trod, it is quite fertile, and each explorer has returned with a deeper understanding of Hume's philosophy (as well as of certain philosophical issues that are still of lively concern); and though I have no pretension to delivering a definitive answer, I do hope that my efforts are similarly rewarded.

3 This may seem to be at odds with Hume's "naturalist" interpreters. But I readily admit that for Hume to say that such beliefs are, as I put it, "epistemologically unjustified," does not mean—as the "naturalist" interpreters have rightly reminded us—that he holds them to be untrue, unuseful, or even unnecessary for our survival. In fact, he says he is thankful that nature has not left it up to human reason to decide whether or not to

form such beliefs (see T 183). Hume does, however, deny that these beliefs, however fortunate we may be to have them, are firmly based on any rational, logical, or evidentiary grounds. I believe my dissension from the naturalist camp is, therefore, one of temperament and not of interpretation. I am just rationalist enough to be more disquieted by Hume's denial of rational justification for these beliefs than comforted by his naturalistic view that they may be justified on other than rational grounds, i.e., expedience, natural impulse, survival, etc.

4 Hume does mention two new "relations" in the section on personal identity. These are: "combination to some *common end* or purpose" (which explains why we say a ship remains the same ship even if all its planks have been replaced one by one) and "*sympathy of parts to their common end*" (T 257) (by which he means the reciprocal causal relations that are the hallmark of organic systems). But, once it comes to the actual explication of how the mind forms the belief in personal identity, these drop completely out of sight. Instead, Hume goes back to the old favorites, resemblance and causation (though not the *reciprocal* causation of the "sympathy of parts"). This may indicate a slight realization on Hume's part that personal identity does not operate exactly upon the same lines as does that of plants, animals, churches, rivers, and the like. The vast majority of mankind never manage to marshal any but the most meager portion of their perceptions toward one end or goal. What may be true of physical parts—bones, muscles, blood, and sinew—is not necessarily true of mental ones—perceptions: my heart, lungs, and legs may all act in unison toward sustaining my life, but where is the sympathy or common end of a series of perceptions, one now the impression of my foot, another the idea of a sonata, a third the fantastic image of some impossible creature?

5 There are only two ways in which all perceptions do resemble each other: one, they are all experienced as part of a temporal succession; two, they are all *my* perceptions (at least all the ones I shall ever encounter). But these bare resemblances do not seem enough to provide the basis for the belief in personal identity, that is, unless one wants to equate the self with time.

6 See T 255 and *passim*.

7 Barry Stroud, *Hume* (London: Routledge and Kegan Paul, 1977), 137.

8 See D. G. C. MacNabb, *David Hume: His Theory of Knowledge and Morality* (New York: Hutchinson's University Library, 1951), 251 and John Passmore, *Hume's Intentions* (London: Duckworth, 1980), 82–3.

9 Terence Penelhum, "The Self of Book 1 and the Selves of Book 2," *Hume Studies* 18 (1992): 282.

10 *Ibid.*

11 Lawrence Ashley and Michael Stack, "Hume on the Self and its Identity," *Dialogue* 8 (1974): 241.

12 Don Garrett, "Hume on Personal Identity," *The Philosophical Review* 90 (1981): 340.

13 See Don Garrett, "Hume on Personal Identity."

14 A full discussion of the uses to which Hume puts this idea of self in Book II, while interesting of itself, lies beyond the scope of my present concern.

15 Wade Robison, "Hume on Personal Identity," *Journal of the History of Philosophy* 12 (1974): 181-93.

16 Don Garrett, "Hume's Self-doubts about Personal Identity," *The Philosophical Review* 90 (1981): 345. See also his *Cognition and Commitment in Hume's Philosophy* (New York: Oxford University Press, 1997), 169-70.

17 It may be objected here that the problem of the subject of mental activity is here just pushed back a step: Who or what after all does the surveying? The Garrettian retort to this would of course be that the act of surveying itself was nothing more than the occurrence of ideas of the perceptions surveyed.

18 Terence Penelhum, "Hume's Theory of the Self Revisited," *Dialogue* 14 (1975): 389-409.

19 Terence Penelhum, "Hume's Theory of the Self Revisited," 399. Emphasis added.

20 See Penelhum, "Hume's Theory of the Self Revisited," and David Pears, "Hume on Personal Identity," *Hume Studies* 19 (1993): 289-99.

21 Terence Penelhum, "Hume's Theory of the Self Revisited," 399.

22 Don Garrett notes that Hume "accounts for our tendency to think of ourselves as having a continuing identity through time by utilizing essentially the same mechanism of identity-ascription that gives rise to the belief in 'continu'd and distinct existences'" (*Cognition and Commitment in Hume's Philosophy*, 163-4). This he cites as a "source of pride" for Hume. My argument, however, is that the mechanisms are so essentially the same that they cannot possibly produce such disparate beliefs as those in body and personal identity.

23 This is related to the reason why, as Robert Fogelin notes, Hume thinks that affirming the simplicity of the soul (or self) is tantamount to adopting Spinozism, that is, the doctrine that thought (unextended perceptions) and matter (extended perceptions) are identical (*Hume's Skepticism in the Treatise of Human Nature* [London: Routledge and Kegan Paul, 1985], 97). If all the perceptions of the soul (or self) are believed to be bound together into a single thing, then no distinction between self and world can be made.

24 Philosophers may add to this some imagined material substance whose purpose is to paper over the interrupted nature of our impressions (which the "vulgar" simply ignore).

25 And each time the "this," "now," and "myself" would be different.

26 Hume realizes this when he retracts his assertion that one method of distinguishing memory-ideas from imagination-ideas is to compare the "order and form" of the ideas in question to that of the impressions that came before them; this method is unpracticable, he comes to conclude, since it is "impossible to recall the past impressions, in order to compare them with our present ideas, and see whether their arrangement be exactly similar" (T 85).

27 MacNabb, *David Hume: His Theory of Knowledge and Morality*, 251–2. MacNabb's question, though Kantian in tone, has resonance with the long history of "idealist" attacks on "empiricist" conceptions of the self that includes, among others, T. H. Green's criticism of Hume, Bradley's opposition to Bain, or even, depending on how far back one wishes to go, Plotinus's dim view of the Epicureans.

28 Nelson Pike "Hume's Bundle Theory of the Self: A Limited Defense," *American Philosophical Quarterly* 4 (1967): 163.

29 Ibid.

30 The distinction here is subtle but crucial. There is no *one* idea that is of the self and that must be self-referential. Instead, the act of reflection produces various ideas (recollections of past impressions and ideas), none of which is *the* idea of self, but all of which must be included in the set of perceptions covered by the belief in personal identity.

31 See T 635.

32 Passmore, *Hume's Intentions*, 82–3.

33 See T 190.

34 See Fogelin, *Hume's Skepticism in the Treatise of Human Knowledge*, 107–8. His concern there is similar to my concern about Hume's inability to account for the unity of mind. Fogelin, however, focuses on whether the mind is unified enough to make sense of Hume's anti-Berkeleyan contention that perceptions can be separated from minds, while my focus is on whether the mind is unified enough to account for the association of ideas.

35 Wayne Waxman, "Hume's Quandary Concerning Personal Identity," *Hume Studies* 18 (1992): 236.