Hume’s “Former Opinions”

Emily Marie Kelahan
Assistant Professor
Illinois Wesleyan University
USA
An Introduction to
ATINER's Conference Paper Series

ATINER started to publish this conference papers series in 2012. It includes only the papers submitted for publication after they were presented at one of the conferences organized by our Institute every year. The papers published in the series have not been refereed and are published as they were submitted by the author. The series serves two purposes. First, we want to disseminate the information as fast as possible. Second, by doing so, the authors can receive comments useful to revise their papers before they are considered for publication in one of ATINER's books, following our standard procedures of a blind review.

Dr. Gregory T. Papanikos
President
Athens Institute for Education and Research

This paper should be cited as follows:
Hume’s “Former Opinions”

Emily Marie Kelahan
Assistant Professor
Illinois Wesleyan University
USA

Abstract

David Hume advances an account of personal identity in Book I of his A Treatise of Human Nature and then retracts it in the Appendix to that work:

“But upon a more strict review of the section concerning personal identity, I find myself involved in such a labyrinth, that, I must confess I neither know how to correct my former opinions, nor how to render them consistent” (T App 10)\(^1\)

His explanation appears, perhaps at first, direct:

“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 would be no difficulty in the case” (T App 21)

However, this explanation is, at best, mysterious. As has been made abundantly clear in the vast literature on this topic, the two principles cited above are not inconsistent. This tells us that there must be a third claim, or set of claims, with which the two principles are inconsistent. A core assumption of the debate surrounding this mysterious text is that the “former opinions” Hume has in mind are philosophical views he advanced earlier in the Treatise, such as his rejection of the Cartesian view of the mind or his claim that the association of ideas in the mind can be fully explained by the principles of resemblance and cause and effect. This assumption is mistaken. I argue that the “former opinions” Hume cannot correct or render consistent are pre-theoretical opinions he formed through socialization and education long before self-consciously pursuing his science of human nature.

Keywords: Hume, personal identity, science of human nature

Acknowledgments: I am grateful to Illinois Wesleyan University for awarding me an Artistic and Scholarly Development Grant to support work on this article, and to audiences at the St. Xavier University Paul Hazard Colloquium Series in March 2012, the Mid-South Philosophy Conference at Rhodes College in February 2013, the Scottish Seminar in Early Modern Philosophy at the University of Aberdeen in May 2013, the Eastern Division Meeting of the American Philosophical Association in Baltimore, MD in December 2013, the Illinois State University Department of Philosophy Colloquium Series in December 2013, and the Athens Institute for Education and Research for their helpful feedback on previous versions of this paper. I am especially grateful to Alan Nelson, Ted Morris, Lewis Powell, and Timothy Yenter for their detailed feedback on previous versions of this paper.

Introduction

Hume was dissatisfied with his account of personal identity. He writes:

But upon a more strict review of the section concerning personal identity, I find myself involved in such a labyrinth, that, I must confess I neither know how to correct my former opinions, nor how to render them consistent” (T App 10).

“Former opinions” is almost universally taken to refer to philosophical opinions concerning Hume’s central theoretical apparatus, the Theory of Ideas, developed in Book I of his Treatise. However, “former opinions” is indeterminate in this passage.¹ I argue that Hume’s “former opinions” are best interpreted as pre-theoretical opinions he formed well before pursuing his science of human nature. Hume did not, I argue, intend to recant any philosophical opinions he developed in Book I. This paper has four parts. First, it briefly explicates Treatise 1.4.6, where Hume’s treatment of personal identity is advanced. Then, it develops an interpretation of Hume’s account of personal identity. Next, it defends a novel interpretation of Hume’s infamous Appendix passage. Finally, it offers a textual argument in support of the interpretation of Hume’s Appendix passage explicated in the previous section.

A Cursory Sketch of Treatise 1.4.6 and the Appendix

Hume rejects the dominant Cartesian theory of mind, according to which we are immediately conscious of a simple, individual self. He employs the Copy Principle (“that all our simple ideas in their first appearance are deriv’d from simple impressions, which are correspondent to them and which they exactly represent” (T 1.1.1.7)) to show that we do not have an impression of something simple and individual and so we cannot have such an idea of the mind:

It must be one impression, that gives rise to every real idea. But self or person is not any one impression, but that to which our several impressions and ideas are supposed to have reference. If any impression gives rise to the idea of self, that impression must continue invariably the same, thro’ the whole course of our lives; since self is suppos’d to exist after that manner. But there is no impression constant and invariable. Pain and pleasure, grief and joy, passions and sensations succeed each other, and never exist at the same time. It cannot, therefore, be from any of these impressions, of

¹I capitalize “Theory of Ideas” to distinguish Hume’s Theory of Ideas from the theory of ideas in general.
from any other, that the idea of self is deriv’d; and consequently there is no such idea. (T 1.4.6.2)

He argues that we have only particular perceptions that are in constant flux. What we call the self is nothing more than a constantly changing bundle of perceptions:

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 catch myself at any time without a perception...I may venture to affirm to the rest of mankind, that they are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement. Our eyes cannot turn in their sockets without varying our perceptions... 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; whatever natural propension we may have to imagine that simplicity and identity. The comparison of the theatre must not mislead us. 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. (T 1.4.6.3-4)

Hume is famously dissatisfied with this account. He identifies two principles he cannot render consistent:

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. (T App 21)

The two principles in the quotation above are quite obviously consistent. Because the two principles are consistent, and because Hume indicates that an inconsistency is lurking somewhere, it is customary for commentators to search for a missing or implicit third claim that is inconsistent with the two principles.
An Interpretation of *Treatise* 1.4.6

It is a popular, but mistaken, interpretation of *Treatise* 1.4.6 to see Hume as meaning to assign to each “self” a particular, fixed, discrete bundle of perceptions. I contend, without time to adequately argue for it here, that what Hume really maintains is that a different bundle of perceptions is uniquely associated with each reference to the “self.” Hume offers us a reductive explanation in terms of perceptions of what is happening whenever we take ourselves to be “the same” over time. As evidence in support of this hypothesis, I point to what Hume does soon after introducing this view of the self. He considers cases in order to show that identity attributions in general are associated with various sets of related perceptions, not the same perception or set of perceptions. Hume devotes nine paragraphs to this project (T 1.4.6.7-16).¹ I take this to be strong evidence in favor of the claim that Hume does not intend to explain or locate “the idea of the self,” but rather to give an analysis in terms of perceptions of different cases of verbal or non-verbal self-reference. Again, Hume is not attempting to develop an account that determines which perceptions in fact constitute a given individual, but rather to apply the framework of perceptions and relations between perceptions to cases of identity attribution so that we might have a better understanding of the cognitive process involved in making identity attributions. Hume considers at least seven cases of identity attribution. Interestingly, the cases he considers are not cases of personal identity attribution. Hume contends:

And here ‘tis evident, the same method of reasoning must be cont’d, which has so successfully explain’d the identity of plants, and animals, and ships, and houses, and of all the compounded and changeable productions either of art or of nature. (T 1.4.6.15)

This is interesting, but perhaps not surprising. Hume’s brand of empiricism commits him to seeing humans as part of the natural order, in no need of special explanation. Hume considers “any mass of matter,” a planet, a ship, an oak tree, a repetitious noise, a brick church rebuilt in stone, and a flowing river. Slightly different lessons may be gleaned from each, but all seven make the same point. Hume’s point is that we never really pick out just one thing with our identity attributions, that the “objects” to which we attribute identity are always changing, yet we call them “identical.” Our identity attributions to humans, Hume claims, work in precisely the same ways as our identity attributions to all other things, which is consistent with Hume’s naturalism (i.e., there is nothing ontologically or explanatorily special about human beings) and his empiricism. The point of considering all of these cases, I

¹Terrence Penelhum is among very few commentators who devote attention to any of the cases Hume examines, and Penelhum considers only two: the noise and the church. Mascarenhas assigns the cases an important role in Hume’s “mistake” with respect to personal identity, but does not discuss the cases in detail.
contend, is to show that ‘identity’ and ‘self’ do not stably co-occur with any particular perceptions or sets of perceptions.\(^1\) Furthermore, Hume does not intend them to.

Hume observes that we observe nothing that \textit{really} unites the perceptions to which we attribute identity. Rather, our attributions derive solely from the associative principles of resemblance and causation. He writes:

\begin{quote}
[A]ll the nice and subtile questions concerning personal identity can never possibly be decided, and are to be regarded rather as grammatical than as philosophical difficulties. Identity depends on the relations of ideas; and these relations produce identity, by means of that easy transition they occasion. But as the relations, and the easiness of the transition may diminish by insensible degrees, we have no just standard, by which we can decide any dispute concerning the time, when they acquire or lose a title to the name of identity. (T 1.4.6.21)
\end{quote}

The term “bundle theory,” as Hume’s account of personal identity is often called, is misleading, for it implies that Hume has a \textit{theory} of personal identity, and we are now in a position to see that he does not. Instead he has a theory that explains our identity \textit{attributions}; namely, the Theory of Ideas. This is why, I strongly suspect, Hume dedicates so much time to the discussion of various cases of identity attribution and to the associative mechanisms that explain those attributions. Hume’s story is not one about the nature of the self, but rather about the nature of our verbal and non-verbal references to what we call “the self.” Hume does not provide us with necessary and sufficient conditions for self-identity. On the contrary, “we have no just standard” of identity, so the best we can hope for is what Hume gives us: a theoretical apparatus capable of explaining various cases of identity attribution. Hume’s discussion of cases is intended to show that we attribute identity in disparate contexts; we apply no common standard of identity in all cases of identity attribution. Thus, it is reasonable to conclude that Hume here means that we are not in possession of a stable, objective standard of identity since we never seem to employ one. There is no metaphysical theory of the self in \textit{Treatise} 1.4.6, nor even an epistemological account of the conditions under which we are justified in attributing identity, but only an application of the Theory of Ideas to the subject of personal identity as a framework for understanding our behavior.

\textbf{The Appendix Revisited}

This interpretation eliminates many of the commitments usually attributed to Hume and cited as reasons for recanting his account of personal identity in

\(^{1}\)More likely, they stably signify a revival set, as Don Garrett has it.
the Appendix. Thus, I offer a new hypothesis explaining Hume’s apparent “recantation.” I propose that Hume’s second thoughts in the Appendix are brought on by reflection on his pre-theoretical opinions, those he formed before undertaking the Treatise. Once Hume completes his work, he attempts to achieve something like reflective equilibrium. He reflects on the claims of his Theory with respect to personal identity as compared to his pre-theoretical opinions on the same subject. He expresses something like existential, spiritual, or psychological dissatisfaction with what the analyses of his Theory reveal about personal identity. He does not conclude that he must have committed some grave error in constructing or applying his Theory. Rather, he concludes that his Theoretical and pre-theoretical understanding of personal identity are inconsistent. Neither his Theory nor any other theory applied to the same subject can entirely accommodate or explain away all of Hume’s pre-theoretical notions of personal identity and yet do the explanatory work such a theory is expected to do.

What Hume cannot render consistent is his pre-Theoretical understanding of the self with his Theory’s analyses of the self. Hume tells us that he “cannot discover any theory” that gives him “satisfaction on this head” (T App 20, my italics). I liken Hume’s position to that of a chemist engaged in somewhat existential reflection on the relationship between her work and her pre-theoretical beliefs. The chemist might think, “My theory says that I am just a collection of atoms, but there’s got to be more to me than that.” Like Hume, she may consider alternative theories that might better fit with both her principles as a scientist and her pre-theoretical notions of herself, but she may find “no satisfaction on this head.” Does this mean that the chemist finds fault with her theory qua theory? I think not. She will probably retain her theory because, though it does not satisfy all of her spiritual and emotional needs, it is a good theory. She might do precisely what Hume does in response to the crisis brought on by the analyses of one’s theory, and that is to humble oneself in the face of human reasoning. Now, Hume is different from the chemist in an important respect: feelings are not completely outside of the realm of theorizing for Hume; they play an integral role in Hume’s explanation of human belief. Hume’s Theory actually predicts his crisis. His feelings of dissatisfaction do, perhaps, lower his degree of belief in his theory, but this does not amount to a retraction of the Theory or the development of serious concern with it. Feelings might lower his degree of confidence in his Theory, but unless there are alternative theories in which he has a higher degree of confidence, there is not necessarily cause for retraction or allegations of theoretical inconsistency.¹

Hume’s explicit mention of inconsistent principles, might, understandably, lead one to think Hume’s second thoughts simply must be theoretical in nature. I agree that they are in part theoretical, but they are not, I claim, purely theoretical. They take as one of their objects Hume’s pre-theoretical beliefs, but they also take as objects his theoretical commitments. In the Appendix

¹I am grateful to Robert Adams for his helpful discussion of this point.
passage Hume stands outside of and attempts to reconcile two sides of himself: Hume the theorist, who is deeply committed to the principles of his Theory, and Hume the ordinary person, who remains in the grips of the deeply entrenched belief that human beings are special creatures whose nature cannot be captured with the same posits and connecting principles as everything else in the empirical universe. Hume the theorist is committed to the following two principles: “that all our distinct perceptions are distinct existences, and that the mind never perceives any real connexion among distinct existences.” Hume the ordinary person believes that his problem would disappear if the following circumstances obtained: “Did our perceptions either inhere in something simple and individual, or did the mind perceive some real connexion among them.” This sentence may be plausibly interpreted as suggesting that Hume’s second thoughts would vanish if there were some way to distinguish the self, or person, from all other empirical objects, but his theoretical principles will not allow it. Hume the theorist has constructed an apparatus that affords an explanation of all human cognition and behavior. However, Hume hasn’t quite shaken off all of the pre-theoretical baggage he acquired in his youth. According to this lingering part of him, humans are explanatorily special; it cannot be that there is simply one perception and then another. This is why the meta-theoretical Hume, the Hume who stands outside Hume the theorist and Hume the ordinary person writes, “For my part, I must plead the privilege of a sceptic, and confess, that this difficulty is too hard for my understanding” (T App 21). This Hume cannot see how to reconcile or render consistent the two other Humes (Hume the theorist and Hume the ordinary person), both of whom are deeply important to him.1

Hume’s problem in the Appendix is more existential, more personal, than theoretical in nature. Hume must reckon with the verdict of his science of human nature, which leaves him with a “is that all there is?” feeling. Hume’s Theory has it that all ascriptions of personal identity are completely explicable in terms of perceptions and relations between perceptions; just as all other phenomena in the science of human nature are. Thus, deep and personal claims such as “my mother is an excellent mother who has supported me through thick and thin,” or “Leslie is a committed public servant who has worked selflessly to better the lives of others,” or “I always try my very best,” or “Chris is a sinister and heartless monster” are completely explicable in terms of and reducible to perceptions and relations between perceptions. Our understanding of and behavior around murderers and torturers, as well as philanthropists and humanitarians, are explicable in terms of precisely the same machinery as are our understanding of and behavior around tables and chairs. There is nothing else to it. Though that verdict is hard to accept for reasons of personal, spiritual, or social significance, its theoretical standing cannot be diminished except by superior theories. Hume’s pre-theoretical opinions might lead to feelings of dissatisfaction with his Theory, but so long as he reasons that his Theory is the best available, then there is no serious philosophical problem

1I am grateful to Simon Blackburn for his helpful discussion of this point.
with it. We do not reject biology because it tells us that we are primates with particularly large frontal lobes, though we might be disappointed by that verdict. Only competing biological theories, not existential or emotional responses, one would hope, are capable of significantly shaking our confidence in the leading theory.

A Textual Argument

Hume’s account of personal identity comes just before Section 1.4.7, Conclusion of this book, which is commonly thought to raise somewhat existential doubts about the claims put forward in the preceding sections and about Hume’s justification for pursuing Books II and III. It is seldom noticed that the tenor of 1.4.7 is very similar to the tenor of the Appendix confession. When Hume comes to the paragraphs concerning personal identity, he departs significantly from the expressed purpose of the Appendix, which he writes is to “[remedy this defect] that some of my expressions have not been so well chosen, as to guard against the all mistakes in the readers” (T App 1). Hume is not, in the case of personal identity, clarifying his thoughts for his readers. Far from preventing his readers from entering a labyrinth, he confesses that he finds himself in a labyrinth. This might be taken to indicate a somber, existential shift in Hume’s thought. One might agree with me that Hume shifts tasks in the Appendix, moving from clarificatory remarks to more existential reflections, but still maintain that Treatise 1.4.7 is only very tenuously connected to the Appendix confession given the chronological order of the texts: the Appendix comes after 1.4.7, and so cannot be a contributing factor to the existential crisis Hume may or may not be having there. However, there is textual evidence in 1.4.7 that suggests that Hume was having existential second thoughts about his account of personal identity before the Appendix was written. In the opening paragraphs of 1.4.7, Hume identifies personal identity as one of the items causing him distress due to its flimsy roots in the imagination:

Nay, even to these objects, we cou’d never attribute any existence, but what was dependent on the senses; and must comprehend them entirely in that succession of perceptions, which constitutes our self or person. Nay farther, even with relation to that succession, we cou’d only admit of those perceptions, which are immediately present to our consciousness, nor cou’d those lively images, with which the memory presents us, be ever receiv’d as true pictures of past perceptions. The memory, senses, and understanding are, therefore, all of them founded on the imagination, or the vivacity of our ideas. (T 1.4.7.3)

This is an important textual discovery. It seems that the development of Hume’s second thoughts concerning his account of personal identity begin not
in the Appendix, but in 1.4.7. There is compelling textual, chronological support for the connection between my hypothesis and Treatise 1.4.7 in addition to the support provided by the similarity of task found in both the Appendix confession and 1.4.7. The Appendix confession and 1.4.7 are united by similarity of task in the following way. In both cases, Hume evaluates his Theory from outside the perspective of pure philosophical theorizing. He attempts to achieve something like reflective equilibrium between his pre-theoretical opinions and the results of his theorizing.

Turning our attention once more to Treatise 1.4.6, the dissatisfaction Hume expresses in the Appendix, like the dissatisfaction he expresses in 1.4.7, concerns not his Theory of Ideas or the results of its application; but rather his pre-theoretical conceptions as compared to his Theory’s analyses. Hume does not offer a metaphysical theory of the self that he later regrets. Hume’s second thoughts in the Appendix do not concern his Theory qua theory; they concern his pre-theoretical self-conception as compared to the results of the application of his Theory, which is a personal matter. Perhaps Hume (and I strongly suspect he is not alone in this) has deep and lingering anti-naturalist intuitions. He set out to explain all empirical phenomena using a single theoretical apparatus and succeeded only to find that some part of him, a very personal part, did not want everything to be explicable in terms of that apparatus. At one level of discourse, he is pleased that such personal-identity invoking notions as being a good parent or citizen, or a bad parent or citizen, are completely explicable in terms of the Theory. At another level of discourse, Hume is surrounded by pre-theoretical, anti-naturalist baggage; and is disappointed to find out that we are on a par with everything else in the empirical universe and that “the cat is on the mat” is explicable in terms of the same posits and connecting principles as “Suzie is a selfless and dedicated humanitarian.” Hume’s dissatisfaction with his Theory’s analysis of personal identity is personal dissatisfaction. His labyrinth is a labyrinth of levels of discourse. His former opinions are pre-theoretical opinions.

References


1I am grateful to Alan Nelson for his helpful discussion of this point.


