Locke and Hume on Personal Identity: Moral and Religious Differences

RUTH BOEKER

Penultimate version. Please cite published version.


Abstract: Hume's theory of personal identity is developed in response to Locke's account of personal identity. Yet it is striking that Hume does not emphasize Locke's distinction between persons and human beings. It seems even more striking that Hume's account of self in Books 2 and 3 of the Treatise has less scope for distinguishing persons from human beings than his account in Book 1. This is puzzling, because Locke originally introduced the distinction in order to answer questions of moral accountability, and Hume's discussion of self in Book 2 provides the foundation of his moral theory in Book 3. In response to the puzzle, I show that Locke and Hume hold different moral and religious views and these differences are important to explain why their theories of personal identity differ.

1. Introduction

According to Hume, “personal identity . . . has become so great a question in philosophy, especially of late years in England” (T 1.4.6.15; SBN 259).1 It is clear that Hume's theory is developed in response to Locke's account of personal identity. Hume follows Locke in discussing, first, the identity of plants and animals before turning to personal identity (see Locke, Essay, 2.27; Hume, T 1.4.6.5-20 (SBN 253–62).2 According to Locke, personal identity consists in sameness of consciousness, and neither sameness of man3 nor substance are necessary or sufficient for personal identity. Hume in part agrees with Locke in so far that he accepts that memory helps us to form beliefs in personal identity, but he also criticizes Locke when he claims that “we can . . . extend our identity beyond memory” (T 1.4.6.20; SBN 262) by causation.4 However, it is striking that Hume does not put much emphasis on Locke's distinction between the terms “person,” “man,” and “substance,” let alone acknowledges Locke’s claim that “person” is a forensic term (see Locke, Essay 2.27.26). This distinction was of great importance for Locke, and he introduced it in order to address questions of moral

Address: School of Philosophy, Newman Building, University College Dublin, Belfield, Dublin 4, Ireland
E-mail: ruth.boeker@ucd.ie
accountability. As I will show, it is, at least prima facie, even more striking that Hume’s account of self in Books 2 and 3 leaves less scope for distinguishing persons from human beings than his Book 1 account. This is puzzling, because Locke originally introduced the distinction in order to answer questions of moral accountability, and Hume’s discussion of self in Book 2 provides the foundation of his moral theory in Book 3. This puzzle calls for an explanation. Is this just a careless oversight on Hume’s part? Or does he have philosophical reasons for not adopting Locke’s distinction? In response to the puzzle I will show that Locke’s account of persons and personal identity is based on very specific moral, religious, and metaphysical background assumptions, and because Hume, in Books 2 and 3, does not share Locke’s specific moral and religious views, Locke’s distinction between persons and human beings holds no attraction for him.

Before I explain in more detail how Locke’s and Hume’s moral, religious, and metaphysical background assumptions shape their different accounts of persons and personal identity, it is helpful to clarify in what, if any, sense Locke and Hume draw a distinction between persons and human beings. Locke initially introduces a conceptual distinction between the terms “person,” “man,” and “substance.” On the basis of this conceptual distinction, Locke argues that personal identity consists in sameness of consciousness and differs from the identity of human beings and substances. I take this to be a metaphysical claim. In present-day terminology, we can rephrase it as the claim that the persistence conditions for persons, human beings, and substances differ. It is worth noting that Locke’s chapter “Of Identity and Diversity” (Essay 2.27) focuses on specifying persistence conditions for persons and other kinds of beings, and to complete this task he does not need to take a stance on the question whether at a time persons and human beings are metaphysically the same or distinct, because as we will see in more detail below, for Locke the characteristic features that all persons share qua members of the kind person are relevant for specifying persistence conditions rather than the metaphysical boundaries of persons at a time. In the following I regard Locke as committed to a conceptual distinction between the terms “person” and “human being” (or “man”), and to the claim that the persistence conditions for persons and human beings differ.

When Hume turns from Book 1 to Book 2 of his Treatise, the focus of his discussion of self and personal identity shifts. For this reason it is helpful to consider his Book 1 account separately from his account in Books 2 and 3. This shift of focus is signalled in his claim that
“we must distinguish betwixt personal identity, as it regards our thought or imagination, and as it regards our passions or concern we take in ourselves” (T 1.4.6.5; SBN 253; see also T 1.4.6.19; SBN 261). The former is the focus in Book 1, and the latter in Books 2 and 3. In Book 1 Hume offers a close examination of the human mind and its operations. The self as discussed in Book 1 can be identified with the human mind, while as we will see in more detail, selves in Books 2 and 3 are embodied social creatures. Book 1 leaves scope for a conceptual distinction between persons (or minds) and human beings (or human animals), but there is not sufficiently clear textual support for reading Hume as drawing metaphysical distinctions between persons and human beings at a time and over time in Book 1. However, once Hume has shifted focus from the mental realm to the social realm in Book 2, his moral and social views presuppose—as we will see in section 7—that the boundaries of persons at a time and over time are the boundaries of ordinary human beings. This metaphysical presupposition in Book 2 of Hume’s Treatise is one strong indication of Hume’s divergence from Locke’s account of personal identity. The task now is to explain how the differences between Locke’s theory and Hume’s Book 2 account are tied to their different moral and religious background assumptions.

2. The significance of Locke’s moral views

Locke emphasizes that we have to distinguish the terms “person,” “man,” and “substance” (see Essay 2.27.7, 2.27.9–29). However, why is it so important for Locke to introduce the term “person” in addition to our ordinary notion of a human being, or “man,” to use Locke’s term? He believes that we need the distinct term of a person in order to make sense of questions of moral accountability (see Essay 2.27.15–20, 2.27.22, 2.27.25–26). According to Locke, only persons are proper bearers of moral accountability. To shed further light on Locke’s theory, it is helpful to consider more closely how Locke’s moral views shape his thinking about persons and personal identity.

Locke argues that before we can specify what personal identity over time consists in, we need to clarify what we mean by the term “person.” According to Locke, the immediate signification of general terms is an abstract idea (see Essay 3.3.12). In our present case the term “person” will signify a set of characteristic features that we associate with the term “person.” Locke also calls this set of characteristic features the nominal essence, and every member of
The kind *person* will have the person-characteristic features, or fall under the nominal essence of persons (see *Essay* 3.3.14–15).

Although Locke offers his initial definition of a person in 2.27.9, his characterization of a person in 2.27.26 helps to make important features of persons explicit:

This being premised to find wherein *personal Identity* consists, we must consider what *Person* stands for; which, I think, is a thinking intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing in different times and places (*Essay* 2.27.9)

*Person* . . . is a Forensick Term appropriating Actions and their Merit; and so belongs only to intelligent Agents capable of a Law, and Happiness and Misery. (*Essay* 2.27.26)

The characterization in 2.27.26 makes explicit that Locke regards persons as moral and legal beings; they are subjects of moral accountability. It is worth noting that moral and legal issues were closely intertwined for Locke:

*Morally Good and Evil* then, is only the Conformity or Disagreement of our voluntary Actions to some Law, whereby Good or Evil is drawn on us, from the Will and Power of the Law-maker; which Good and Evil, Pleasure or Pain, attending our observance, or breach of the Law, by the Decree of the Law-maker, is that we call Reward and Punishment. (*Essay* 2.27.5)

He understands morality in terms of law, distinguishes three types of law, namely, divine law, civil law, and the law of opinion and reputation, and argues that divine law “is the only true touchstone of moral Rectitude” (*Essay* 2.28.8).

Although it is not explicit in the chapter “Of Identity and Diversity,” Locke endorses a conception of morality that presupposes a superior divine lawmaker. According to Locke, sanctions are needed in order to establish binding moral laws and only a superior lawmaker, namely God, is in a position to create such laws. This means that Locke’s moral views are closely tied to the idea of a divine Last Judgement (see *Essay* 2.27.22, 2.27.26). However, it is worth noting that Locke conceives of the divine judgement not as merely external, but rather it is important that individuals can understand from the inside the justice of reward or punishment. This view finds expression in Locke’s claim that our own “Conscience accusing or excusing” us in the Great Day (*Essay* 2.27.22). What is involved in understanding the justice
of reward and punishment, according to Locke? First, we must be able to know that God exists and that he is a superior lawmaker who can enforce laws. Second, the divine laws to which our actions are meant to conform have to be made known to us, or we have to be able come to know them. Third, when we are held accountable for particular actions, we must be able to remember the actions as our own. It is worth adding that previous awareness is built into Locke’s notion of memory (see Essay 1.4.20, 2.10.2, 2.10.7). This means that in order to be able to remember a thought or action now, one must have been conscious of it at the original time and when we remember a past thought or action, we not only remember the thought or action, but also that we have had it before.

We can now see that Locke’s particular moral views and his particular thinking about moral accountability explain why he regards sameness of consciousness as necessary for personal identity over time. To further illustrate how controversial Locke’s thinking about moral accountability is, it is helpful to turn to his examples of sleep and drunkenness. With regard to sleep, Locke introduces a hypothetical example of daytime and night-time Socrates in Essay 2.27.19:

If the same Socrates waking and sleeping do not partake of the same consciousness, Socrates waking and sleeping is not the same Person. And to punish Socrates waking, for what sleeping Socrates thought, and waking Socrates was never conscious of, would be no more Right, than to punish one Twin for what his Brother-Twin did, whereof he knew nothing, because their outsides were so like, that they could not be distinguished; for such Twins have been seen. (Essay 2.27.19)

According to Locke, it is unjust to punish daytime Socrates for thoughts or actions of night-time Socrates, if daytime Socrates has no consciousness of, or access to, night-time Socrates’s thoughts or actions. This example illustrates why he regards sameness of consciousness as necessary for personal identity.

While many of Locke’s contemporaries were willing to accept that it is unjust to punish a sleepwalker, they were not willing to accept the implications with regard to drunkenness. As in the case of sleep, Locke argues that it is unjust to hold a person accountable for a crime, committed by the same human body while intoxicated, if he or she is unable to remember it afterwards. Locke’s view troubled his friend Molyneux, who writes to
Locke: “Drunkenness is itself a Crime, and therefore no one shall alledge it in excuse of any other Crime” (Letter no. 1685, in Correspondence, 4:767). However, Molyneux’s criticism does not move Locke to revise his view, but rather in response to Molyneux he asks Molyneux to consider a case where a drunkard gets a fever:

For I ask you, if a man by intemperate drinking should get a fever, and in the frenzy of his disease (which lasted not perhaps above an hour) committed some crime, would you punish him for it? If you would not think this just, how can you think it just to punish him for any fact committed in a drunken frenzy, without a fever? Both had the same criminal cause, drunkenness, and both committed without consciousness. (Letter no. 1693, Correspondence, 4:785-86)

Because there is no principled way to distinguish these two cases, Locke believes that in either case it is unjust to be held accountable for an action that was done without consciousness and that one is unable to remember.

Leibniz was another critic of Locke’s view. Leibniz, similarly as Molyneux, argues that sleepwalking and drunkenness are not analogous: “We punish drunkards because they could stay sober and may even retain some memory of the punishment while they are drunk. But a sleepwalker is less able to abstain from his nocturnal walk and from what he does during it” (New Essays 2.27.22, 243).\textsuperscript{17} Leibniz rejects Locke’s view that sameness of consciousness is necessary for personal identity and believes that “the testimony of others could fill in the gap in my recollection. I could even be punished on this testimony if I had done some deliberate wrong during an interval which this illness had made me forget a short time later” (New Essays 2.27.9, 236).

The important point for present purposes is that Locke’s thinking about moral accountability was and remains controversial. Rather than trying to evaluate Locke’s views in light of Molyneux’s and Leibniz’s criticisms, I want to draw attention to the implications that a different understanding of moral accountability has with respect to personal identity over time. If we assume, as Locke does, that persons are subjects of moral accountability, but agree with Molyneux’s or Leibniz’s understanding of moral accountability, then it will not be possible to give persistence conditions for persons in terms of sameness of consciousness, but rather, in order to maintain that a person now, who is unable to remember a criminal action, is the same
person as an individual who committed a crime while drunk, persistence will have to be explained in terms of bodily continuity, the continued existence of an immaterial substance, or some other condition.\textsuperscript{18}

To sum up the results, we have identified two important factors why Locke links his moral conception of a person to an account of personal identity in terms of sameness of consciousness: First, Locke regards persons as moral subjects of accountability and, second, he thinks about moral accountability in a particular and controversial way. Note that I am not claiming that these factors offer a complete explanation. Indeed, they only establish that sameness of consciousness is necessary for personal identity. Additionally, there are metaphysical and religious background assumptions at work, as I will show in the following.

3. The significance of Locke’s religious and metaphysical views
I now want to draw attention to two additional factors that are important to understand why Locke emphasises the distinction between persons and human beings, and why he argues that personal identity consists in sameness of consciousness: one concerns Locke’s metaphysical agnosticism that we are not in a position to know whether thinking takes place in material or in immaterial substance; the other, his religious belief in the afterlife and a Last Judgement.

Let me say more about the former. One important project of Locke’s \textit{Essay}, and Book 4 in particular, is to draw attention to the limitations of human understanding and to restrict the boundaries of knowledge to those matters that we can know with certainty. It does not follow, according to him, that inquiries that fall outside the scope of knowledge are pointless, but rather in such cases we can form probable beliefs or rely on faith. The important point being that we do not mistake probable beliefs or faith for knowledge. Locke argues that while we all experience that we have thoughts,\textsuperscript{19} we are not in a position to know whether these thoughts inhere in an underlying material or immaterial substance.\textsuperscript{20} His view that personal identity consists in sameness of consciousness holds irrespective of whether conscious states inhere in a material or an immaterial substance. Consequently, his view is not affected by the limitations of human knowledge and has broader scope than alternative theories that tie a person’s existence to a material body, an immaterial substance, or both. Any other view that ties personal identity to a particular metaphysical doctrine will collapse if the position is false.
Additionally, Locke’s religious beliefs shape his theory of personal identity. Locke was a Christian believer who argued that Scripture should be taken seriously. He distinguishes faith from knowledge, and although faith based on revelation is not absolutely certain, he argues that faith cannot properly be doubted and is therefore almost as certain as knowledge (see *Essay* 4.18.10). In light of Locke’s religious views, it becomes clear that it was of great importance for him to offer a theory of personal identity that can accommodate the possibility of life after bodily death, because the Bible reveals that there will be an afterlife (see *Essay* 4.18.7).21

Locke believes that we face “great Absurdities” (*Essay* 2.27.21) if we do not distinguish persons from human beings and fail to acknowledge that personal identity consists in sameness of consciousness. We will see that those who do not clearly distinguish persons from human beings face serious metaphysical difficulties to explain the possibility of the resurrection and the afterlife, or they are at risk to leave room for unjust reward and punishment. Since the term “human being” is not used univocally, let us follow Locke’s classifications in *Essay* 2.27.21. There he distinguishes three meanings that the term “man,” or “human being,” could have: It can stand for (i) an immaterial substance, (ii) a material human animal, or (iii) the union of a material body and an immaterial substance.

If we assume that human beings are immaterial substances, then the view is at risk to leave room for unjust reward and punishment, especially in the afterlife. Defenders of this view would not agree with Locke that personal identity consists in sameness of consciousness, and instead, they will maintain that personal identity consists in the continued existence of an immaterial substance. They have, however, the option to acknowledge, as Locke does, that a necessary condition for just accountability is that the individual to be rewarded or punished must be able to remember the relevant action as her own. I will show that either way, Locke believes that the view can lead to serious injustice. Due to our limited understanding of the ontological constitution of immaterial substances, we are not in a position to rule out the possibility of transfer of consciousness from one immaterial substance to another (see *Essay* 2.27.13). To illustrate how the view that we are currently considering can lead to injustice, consider one immaterial substance, call it $S$, and another immaterial substance, call it $S^*$. Let us assume further that at time $t_1$, $S$ committed a serious crime. At time $t_2$ all of $S$’s memories of the crime are transferred to $S^*$, and at all subsequent times $S$ is unable to remember the crime, while $S^*$ is able to be conscious of it. Is it an option to punish $S^*$ for it? While $S^*$ is conscious,
or at least, able to be conscious of the crime, punishing $S^*$ for it would violate the widely accepted condition that personal identity is a necessary condition for moral accountability. Since the crime was committed by $S$ and not $S^*$, this condition would be violated under the present proposal that personal identity consists in the continued existence of an immaterial substance. Is it possible to punish $S$ instead? This is not an option for Locke, because $S$ is not in a position to remember it and Locke regards it as unjust to punish a subject if they are not in a position to understand the justice of the punishment. However, not punishing any resurrected being for the crime would also be unjust. Hence Locke believes that immaterial substance views of personal identity can lead to serious injustice. If we replace the immaterial substance view of personal identity with Locke’s view in terms of sameness of consciousness this problem will be solved. Hence Locke has reasons to prefer his view.

If we assume instead that human beings are purely material human animals, then it follows that the human being will cease to exist at bodily death and the difficulty will be to explain why the same human being is recreated at the resurrection, rather than being newly created. Several solutions were proposed by Locke’s predecessors. One proposal was that resurrected beings must be composed of exactly the same material particles as human beings had at one or more times of their lives. However, which time(s) during ordinary life is or are relevant? Choosing only one time seems arbitrary and choosing more than one time will result in resurrected creatures that are rather different in size and shape from ordinary human beings. Moreover, there was the pressing worry that there simply will not be sufficient material particles in the afterlife to recreate all human beings. This problem arises because particles that were once part of one human being can become part of another human being at a later time, and it was lively debated as the thread of direct and indirect cannibalism at Locke’s time. Furthermore, it is unclear why numerical identity of particles is relevant for survival. Locke, on the contrary, has a convincing answer why sameness of consciousness is relevant for personal identity: it enables persons to understand the justice of reward and punishment.

Other attempts focused on explaining how the life of an organism can continue in the afterlife. One such proposal was appeal to the seed metaphor. The thought was that something analogous to a seed could preserve the life of a human organism between death and Resurrection. However, Locke points out that the analogy with seeds is not suitable to explain numerical identity and confounds it with membership in the same species (see Works, 4:319). Moreover, seeds could easily be destroyed, for example, by fire. An attempt to solve this latter
worry can be found in some Rabbinic sources that postulate an indestructible resurrection bone, which is said to have the size of an almond. However, this proposal seems ad hoc and lacks evidence. If there really was an indestructible bone, one would expect to find it, for instance, when a corpse is burnt to ashes. The only option to rescue this proposal would be by appeal to divine providence. However, then it is unclear why we additionally need a resurrection bone. The advantage of Locke’s theory is that it is more plausible, because it does not postulate mysterious entities that lack evidence.

The third option is to account for human beings as unions of material bodies and immaterial souls. Given this view, there are a number of different ways to explain the persistence of human beings. First, if persistence of the human being is explained in terms of the continued existence of the same soul and the same body, then this view will inherit the problems that arise for material views. Second, if persistence does not require that the same body continues to exist, but only that the soul is united to any body at all times, then the human being persists in virtue of the continued existence of the immaterial soul, and the view will inherit the problems that arise for immaterial views. It follows that this option is not attractive either.

To sum up, in this section I drew attention to two further factors that are relevant for understanding why Locke introduces a conceptual distinction between the terms “person” and “human being” (or “man”) and argues that personal identity consists in sameness of consciousness and thereby offers distinct persistence conditions for persons and human beings. First, Locke’s metaphysical agnosticism with regard to the materiality or immateriality of thinking substances helps to see why it was significant for Locke to offer a theory that is consistent with both materialism and immaterialism. Second, due to his religious beliefs, he aimed to offer a theory of personal identity that takes seriously the possibility of an afterlife and a Last Judgement. His theory escapes problems of injustice that arise for immaterial views and can better explain the possibility of the afterlife than material views.

4. Contrasting Locke’s and Hume’s moral and religious background assumptions
Before we turn to the details of Hume’s account of persons and personal identity, let me introduce Hume’s very different moral and religious, or better irreligious, background assumptions. It is interesting to contrast Locke’s and Hume’s views about persons and
personal identity, precisely because Hume opposes the moral and religious background assumptions that inform Locke’s theory.

Hume’s critical attitude towards religion finds clear expression in his essay “Of the Immortality of the Soul.”30 There he presents a series of metaphysical, moral, and physical arguments that question the possibility of an afterlife and the immortality of the soul.

As regards metaphysical arguments, Locke and Hume agree that the immortality of the soul cannot be proven a priori, and they share a certain metaphysical agnosticism, but some nuanced differences come to light when we turn to Hume’s physical arguments. Locke makes clear that we cannot know whether thinking takes place in a material or an immaterial substance, but he maintains that it is more likely that thinking substances are immaterial (see Essay 2.27.25).31 Hume, on the contrary, argues that human beings are analogous to non-human animals and in light of the great analogy, it is more likely that we are mortal material beings.32

We can find perhaps the most significant disagreements when we turn to Hume’s moral arguments. He argues that the idea of a divine Last Judgement does not make sense for a number of reasons. First, it is unjustified to ascribe intentions to God that are not grounded in present experiences. Second, it is unclear what a divine standard of reward and punishment could be. According to Hume, human standards for punishment are grounded in sentiments, but “[s]hall we suppose, that human sentiments have place in the deity?” (“Immortality of the Soul,” 594). Third, divine punishment lacks a purpose, because it takes place after ordinary life on Earth and therefore cannot lead to any improvements of life and society in this world. Hence, it “is inconsistent with our ideas of goodness and justice” (“Immortality of the Soul,” 594). Fourth, punishment should be proportional to the offence. Eternal divine punishment fails to meet this requirement, because eternal punishment for finite crimes is disproportional.33 Fifth, a divine Last Judgement presupposes that human beings can be divided into two different types, namely good and bad, but rather, according to Hume, goodness and badness come in degrees.34

Hume’s philosophy does not leave scope for a religious foundation of morality as we find it in Locke. Instead, the aim of his philosophical project, as he tells us in the Introduction to the Treatise, is “to explain the principles of human nature” and thereby “propose a compleat system of the sciences, built on a foundation almost entirely new, and the only one upon which they can stand with any security” (T Intro 6; SBN xvi). According to Hume, any science,
including mathematics, natural science, natural religion, morals, politics, has a relation to human nature. This means we cannot develop morality independent of a study of human nature, but rather Hume believes that it is important to begin with an examination of the principles and mechanisms that govern human nature. On this basis, let us turn to Hume’s account of self and personal identity.

5. Hume on self and personal identity

Hume introduces the topic of self and personal identity in Book 1 of the Treatise in a section titled “Of Personal Identity” (T 1.4; SBN 251–63), he raises serious doubts about his Book 1 account in the Appendix, and returns to discussions of self in Book 2 when he develops his theory of the passions, which provides the foundation of his account of morals in Book 3. While his Book 1 account and his critical remarks in the Appendix have received much attention in the literature, I will here only briefly comment on his Book 1 account and then turn to his account of self in Books 2 and 3, because in order to understand the differences between Locke and Hume, we will have to examine what role Hume’s self plays in relation to moral questions, such as questions of moral accountability. Since these questions are more or less absent in Book 1, Books 2 and 3 provide the more interesting sources for present purposes.

To begin, let me outline the questions that Hume addresses throughout the Treatise. They include the following:

(1) What, if any, kind of self do we find when we introspect? (T 1.4.6.1–4; SBN 251–53)
(2) How do we form beliefs in personal identity over time? (T 1.4.6.5–21; SBN 253–62)
(3) What makes me the person I am? (T 1.4.6.3–5, 1.4.6.19, 2.1.2–11; SBN 252–53, 261; SBN 275–324)
(4) What psychological mechanisms govern the self? (T 1.4.6.18–20, 2.1.2–6, 2.1.11; SBN 260–62, 277–94, 316–24)

There is a further question that Hume does not address explicitly, but there is some indirect evidence, at least in Book 2, that suggests that he takes a stance on it:

(5) What are the boundaries of selves at a time and over time?

One main question in Book 1 concerns the self that we find within ourselves via introspection. Hume denies that we can find a self that is simple and has perfect identity, by
which he means that it exists invariably and uninterruptedly over a period of time (see T 1.4.2.24, 1.4.2.26, 1.4.2.30–31; SBN 199–202). Instead he maintains that the self, or mind, that is given to us in experience is a bundle of changing perceptions\(^{36}\) (see T 1.4.6.3–4; SBN 252–53).

The second major question concerns how we can explain that we, nevertheless, form a belief in personal identity over time. Hume writes:

\[\text{[A] question naturally arises concerning this relation of identity; whether it be something that really binds our several perceptions together, or only associates their ideas in the imagination? That is, in other words, whether in pronouncing concerning the identity of a person, we observe some real bond among his perceptions, or only feel one among the ideas we form of them? This question we might easily decide, if we wou'd recollect what has been already prov'd at large, that the understanding never observes any real connexion among objects, and that even the union of cause and effect, when strictly examin'd, resolves itself into a customary association of ideas. For from thence it evidently follows, that identity is nothing really belonging to these different perceptions, and uniting them together; but is merely a quality, which we attribute to them, because of the union of their ideas in the imagination, when we reflect upon them. (T 1.4.6.16; SBN 259–60)}\]

This passage supports that Hume’s primary task in Book 1 is to offer a psychological explanation of how we form the belief in personal identity rather than a metaphysical account of the persistence conditions for persons. Having restricted the focus of discussion to psychological explanations, he argues that the only candidates to explain our belief are the three associative principles of the mind: resemblance, contiguity, and causation (see T 1.4.6.16; SBN 260).\(^{37}\) While all three principles are relevant to explain our belief in the identity of plants and animals, Hume claims that only resemblance and causation play a role when we form beliefs in personal identity, and contiguity “has little or no influence in the present case” (T 1.4.6.17; SBN 260).\(^{38}\)

At this stage it is interesting to return to the question in what sense Hume’s Book 1 account of personal identity leaves scope for a distinction between persons and human beings. In T 1.4.6, Hume uses the term “mind” interchangeably with “self” or “person” (see T 1.4.6.4,
1.4.6.15–16, 1.4.6.22; SBN 253, 259, 263). He develops his Book 1 account of personal identity in analogy to the identity that we ascribe to plants and animals (see T 1.4.6.5; SBN 253) and regards human beings as human animals: “An infant becomes a man, and is sometimes fat, sometimes lean, without any change in his identity” (T 1.4.6.12; SBN 257). This passage undoubtedly resembles Locke’s statement: “a Colt, grown up to a Horse, sometimes fat, sometimes lean, is all the while the same Horse” (Essay 2.27.3). This gives us strong evidence that Hume in Book 1 identifies human beings with human animals and persons with minds. This means his approach to human beings is third-personal while his approach to persons is first-personal.

On this basis we can say that Hume’s Book 1 account involves a conceptual distinction between persons and human beings, but it is important to note that it is not the same as the distinction that we find in Locke’s chapter “Of Identity and Diversity.” According to Locke, persons are subjects of accountability; they are agents who are “capable of a Law, and Happiness and Misery” (Essay 2.27.26). Since Hume has restricted the focus of his Book 1 account to the mind, or to “personal identity, as it regards our thought or imagination” (T 1.4.6.5; SBN 253), his Book 1 account is too restricted to explain issues that are at the heart of Locke’s account of persons. For Hume “personal identity, as it regards our thought or imagination” has to be distinguished from personal identity “as it regards our passions or concern we take in ourselves” (T 1.4.6.5; SBN 253). In order to address questions of moral agency and moral accountability that interest Locke, Hume has to extend the discussion beyond the mental realm and has to turn to personal identity in the latter sense.

Hume acknowledges that the Book 1 account of personal identity is incomplete (see T 1.4.6.5, 19; SBN 253, 261). It is primarily backwards directed and explains how we regard ourselves as identical with our past selves. Once we turn to the passions, we can account for our concerns for the future, and this makes it possible to develop an account of personal identity that extends into the future.39

Interpreters disagree whether his Book 2 account is consistent with his Book 1 account and merely supplements it, or whether he develops a significantly different account of self in Book 2.40 When Hume first returns to the topic of self in Book 2, he characterizes self as in Book 1, as “that succession of related ideas and impressions, of which we have intimate memory and consciousness” (T 2.1.2.2; SBN 277). Prima facie, this passage provides strong support for supplementation interpretations, but to develop a nuanced interpretation of
Hume’s theory, it is helpful to return to the five questions. In Book 2, Hume’s focus has shifted to questions (3) and (4), namely, the question of what makes a person the person she is, and questions concerning the psychological mechanisms that govern the self.

Lorenzo Greco argues in a recent article that it is helpful to adopt Marya Schechtman’s distinction between “reidentification” and “characterization” questions, and that Hume’s discussion of personal identity in Book 1 concerns the reidentification question, while in Book 2 he turns to the characterization question. In rough terms, this distinction captures the shift from Book 1 to Book 2. I am, however, reluctant to use Schechtman’s terminology, because she introduces the reidentification question as a metaphysical question concerning the necessary and sufficient conditions of personal identity. I do not think that there is strong enough textual support for interpreting Hume’s Book 1 account as metaphysical; at best, this is controversial. Instead, question (2) above is better supported by Hume’s text, and thus I take it to be his primary focus in Book 1. This becomes clear, for instance, when he asks “[w]hat then gives us so great a propension to ascribe identity to these successive perceptions, and to suppose ourselves possest of an invariable and uninterrupted existence thro’ the whole course of our lives?” (T 1.4.6.5; SBN 253). Here he clearly asks for a psychological explanation why we form beliefs in personal identity over time and sets the agenda for the rest of the chapter. Nevertheless, I believe that it is helpful to consider separately whether Hume endorses metaphysical claims about the boundaries of selves at a time and over time. For this reason, I have added question (5) as an additional question. Hume does not address this question explicitly. As argued above, I believe the textual support is not strong enough for ascribing metaphysical views about persons and personal identity to Hume in Book 1, but I will argue that his views in Books 2 and 3 imply that persons are embodied human beings and that personal identity over time is bound by the continued existence of a human body.

Since for present purposes, Hume’s account of self in Book 2 is of particular interest, I will turn immediately to Book 2, explain his theory of the construction of selves in Book 2 and why his views in Books 2 and 3 imply that selves are embodied and continue to have the same body over time.

6. Construction of selves in Book 2

In Book 2, Hume returns to the self in the context of an examination of the indirect passions and the mechanisms that govern them. Self is the object of the indirect passions pride and
humility, similarly as another person is the object of love and hatred. He argues that these four indirect passions operate by the same mechanism, which he often calls a “double relation of ideas and impressions” (see T 2.1.5–11, 2.2.2, 2.2.9, 2.2.11; SBN 285–324, 332–47, 381–89, 394–96). According to Hume, the indirect passions of pride and humility, love and hatred have a wide variety of causes and are immediately directed towards an object, which is self in the case of pride and humility, and another person in the case of love and hatred (see T 2.1–2; SBN 275–398). The causes include mental qualities such as wit, good humour, courage, justice, and their opposites, bodily qualities such as beauty or strength, and also external objects as houses, gardens, dogs, clothes, children (see T 2.1.2.5; SBN 278–79).

The double relation of ideas and impressions helps us to see that in the context of Book 2, a person or self is closely associated with a set of qualities, possessions, and other objects related to the person or self. These features play a constitutive role and make a person the person she is. Let us call these features the person-defining features. We characterize one person as witty, generous, a serious swimmer and good cook; another as selfish, beautiful, and an ambitious dancer; and someone else as having a strong sense of justice, being open-minded, a devoted scholar and talented piano player.44

How does this Book 2 self relate to the Book 1 view, according to which the self, as given in experience, is a bundle of rapidly changing perceptions?45 While these views need not be in tension, the focus in Book 2 has shifted. Instead of just considering any perceptions that pass through one’s mind, we now turn to the features that we most closely associate with a person and that we regard as characteristic of who she is; these features make a self proud (or humiliated) to be the person she is.

The interesting question is whether these person-defining features are stable or whether they are fluctuating and change as our other perceptions. Hume is well aware that our feelings can change quickly. While I may be proud about my achievements as a piano player my pride can vanish quickly and self-doubts can take its place. At this stage he draws attention to a secondary cause of pride and humility, namely, sympathy (see T 2.1.11.1; SBN 316).

Sympathy, according to Hume is a psychological mechanism whereby we tune into the feelings of others and make their feelings our feelings (see T 2.1.11, 2.2.5.21; SBN 316–24, 365).46 According to Hume, to strengthen our own feelings, it is important that other people share them and thereby increase the pleasure (or pain) that we receive from our person-defining features. For example, when other people take pleasure in my piano performances I receive
reassurance and my pride becomes more stable. Or, to turn to another example, assume that Rebecca is inclined to be generous. Hume would argue that it is important that other people take pleasure in her generous actions and the assurance, praise, and gratitude that she receives from other people will play a significant role in developing the stable character trait of generosity.

Although Book 2 selves are primarily constructed by a set of qualities and possessions, other people significantly influence and shape who we are. Through interaction with others, some of the person-defining features become stable, others vanish and may be replaced by new features. The features of a self in solitude are more likely to be fluctuating, while selves that interact with others in society are in a better position to develop stable characters.

7. Ontological boundaries of selves in Book 2

Having argued that selves in Book 2 are associated with a set of person-defining features, I do not want to suggest that selves in Book 2 can be ontologically identified with this set of features, rather I will show that there is direct and indirect textual evidence that the ontological boundaries of selves are the bodies of human beings. I will now introduce several arguments why selves in Book 2 have to be embodied at a time and continue to have the same body over time. This will mark an important difference between Locke and Hume, because, if it is correct that Book 2 selves are embodied and continue to have the same body, then it follows that Hume’s Book 2 account does not have much scope for ascribing different persistence conditions to persons and human beings. Subsequently, it is not important for Hume in the context of Book 2 to emphasize a conceptual distinction between persons and human beings as Locke does.

My first observation is that Hume treats self and another person analogously in his account of the indirect passions pride and humility, love and hatred. Pride or humility is directed towards self as love, and hatred is directed towards another person. Why is it important for Hume that selves or other people in Book 2 are embodied? Alternatively, he could still, as in Book 1, regard selves as bundles of perceptions, some of which lack spatial location. If this were his view in Books 2 and 3, then we would not be in a position to direct love (or hatred) towards particular individual persons such as Anne rather than Susan or vice
versa. Only if selves are embodied will we be in a position to distinguish Anne from Susan and
to direct our love towards one, both, or neither of them.

To further illustrate this point let us focus on generosity—one of the causes of pride
and love (see T 2.1.2.5; SBN 278–79). If generosity causes pride then the pride will be directed
towards self. If generosity causes love then it will be directed towards another person.
However, when we direct love towards another person, we do not merely direct it towards
some other person in the abstract, but rather it is directed towards particular individuals such
as Anne or Susan. In order to direct our love towards Anne rather than Susan, or towards
Susan rather than Anne, or towards both, there will have to be a means for distinguishing
Anne from Susan. We are able to do this if they are uniquely located in space. It follows that in
order to direct our love towards another person, the other person has to be embodied.
Because Hume regards self and another person as analogous, it follows that selves in Book 2
are embodied.

Second, according to Hume, Book 2 selves are social creatures. We seek the company
of others and “[a] perfect solitude is, perhaps, the greatest punishment we can suffer. Every
pleasure languishes when enjoy’d a-part from company, and every pain becomes more cruel
and intolerable.” (T 2.2.5.15; SBN 363). Social interaction with others would be severely
limited, if not impossible, were selves not embodied and would not continue to have the same
body. For example, it is not clear that we would be able to enjoy the pleasures of friendship.
How would disembodied selves meet, have dinner together, or enjoy each other’s company
during a walk through the countryside? Disembodied bundles of perception may be in a
position to share joint intellectual pleasures. One’s intellectual curiosity may inspire other
minds, but as already mentioned, it is not clear how we can appreciate, or in Hume’s terms,
direct our love towards the curiosity of a particular person; rather, if selves were disembodied,
it seems that the only option is that curiosity as such could inspire us, but not the curiosity of
particular individuals. In order to respond and interact with the curiosity of particular
individuals, we have to be in a position to uniquely locate their curiosity. While the arguments
so far focused primarily on embodiment at a time, let us consider further why it is important
that selves continue to have the same body. Social interaction takes place over time, and love
and hatred often develop over time. To effectively interact with each other and to take part in
complex actions or projects, we need to be able to recognize others and have assurance that
we continue to interact with them. Of course, it does not follow that the bodies of selves have
to be human organism, but merely that they need to have bodies that are uniquely located in space and can be recognized as theirs by others. Here I have primarily identified epistemological and pragmatic reasons why selves need to continue to have the same body, but we will see below that Hume’s account of sympathy and his view that a person can change her character offers additional reasons why the continued existence of a human body provides the boundaries of a person’s life.

Third, Hume’s account of sympathy further supports this view. According to Hume, sympathy plays an important role in the development of stable passions (see T 2.1.11.1; SBN 316). He maintains that the degree to which we can sympathize with others is proportional to the strength of the relations that we bear to others. This means that we would not be able “to feel sympathy in its full perfection” (T 2.1.11.8; SBN 320) if selves were not related by contiguity. Thus, it is central to Hume’s philosophy in Books 2 and 3 and in particular to his moral philosophy, which is based on sympathy, that selves in Books 2 and 3 are connected by contiguity, and thus embodied at a time and continue to have their bodies over time.

Once we take seriously that selves are social creatures and that sympathy is a very important principle of human interaction (see T 3.3.1.7, 10–11, 3.3.6.1; SBN 575–79, 618), we can see that it is central to Hume’s philosophy in Books 2 and 3 that selves are embodied and continue to have their bodies.

8. Hume’s embodied selves and moral responsibility

We have already seen that Hume endorses irreligious views and provocatively questions the plausibility of ideas of an afterlife and a Last Judgement. Hence, he does not share one of the major religious motivations that made it significant for Locke to offer different persistence conditions for persons and human beings. Due to this religious difference it makes good sense that Hume does not offer different persistence conditions for persons and human beings. While I believe that the religious differences are a very important factor in explaining the differences between Locke and Hume, I want to return to the moral background assumptions, because—as we have seen—Locke’s particular thinking about moral accountability motivated him to introduce a conceptual distinction between persons and human beings and to argue that sameness of consciousness is necessary for personal identity. Let us examine whether Hume’s understanding of moral accountability diverges from Locke’s.
Before we look at the details, it is helpful to reflect more generally on Locke’s and Hume’s moral views, respectively. Locke stands in the Natural Law tradition. According to him, divine law “is the only true touchstone of moral Rectitude” (Essay 2.28.8). This statement shows that divine law plays a fundamental role in his moral theory and that the moral rightness or wrongness of voluntary actions is to be assessed according to their “Conformity or Disagreement . . . to some Law” (Essay 2.28.5). Given the prominent place of laws and their enforcement by sanctions, we can see why Locke’s view focuses on reward or punishment for individual actions.

By the time Hume wrote the *Treatise* other British moral philosophers, with whose work Hume was familiar, had distanced their views from Locke’s morals views and the views of other proponents of Natural Law theory. Anthony Ashley Cooper, the Third Earl of Shaftesbury, for example, distanced his moral views from Locke’s and the search for the greater good, virtue, and character development are central topics in his philosophical works. Francis Hutcheson questions moral views that ground morality solely in relation to the law of a superior. According to Hutcheson, such theories are problematic, because “all Laws operate only by Sanctions of Rewards, or Punishments, which determine us to Obedience by Motives of Self-Interest.” His main thesis is that we are by human nature not only self-interested but also benevolent and have a special moral faculty, which he calls a moral sense.

Hume’s moral philosophy builds on the works of Shaftesbury, Hutcheson, and other British moral philosophers. He shares their interest in examining human nature, regards us by nature both self-interested and benevolent, and develops a detailed catalogue of artificial and natural virtues in Book 3 of the *Treatise*. The natural virtues correspond to the benevolent, the artificial virtues to the self-interested part of human nature, and the artificial virtues are based on social conventions and needed additionally to create a functioning society. As we have already seen, due to Hume’s irreligious views, his theory leaves no room for divine law; instead Hume examines in detail how social and sympathetic interaction with others shapes our character and leads us to regard each other as virtuous or vicious.

On this basis, let us turn to Hume’s remarks about moral responsibility:

Actions are by their very nature temporary and perishing; and where they proceed not from some cause in the characters and disposition of the person, who perform’d them, they infix not themselves upon him, and can neither
rebound to his honour, if good, nor infamy, if evil. The action itself may be blameable; it may be contrary to all the rules of morality and religion: *But the person is not responsible for it; and as it proceeded from nothing in him, that is durable or constant, and leaves nothing of that nature behind it, 'tis impossible he can, upon its account, become the object of punishment or vengeance.* (T 2.3.2.6; SBN 411, emphases added)

If any *action* be either virtuous or vicious, 'tis only as a sign of some quality or character. It must depend upon durable principles of the mind, which extend over the whole conduct, and enter into the personal character. Actions themselves, not proceeding from any constant principle, have no influence on love or hatred, pride or humility; and consequently are never considered in morality. (T 3.3.1.4; SBN 575)

These passages reveal important differences between Locke’s and Hume’s thinking about moral responsibility. While Locke’s theory focuses on moral accountability for individual actions and a person’s ability to be conscious of them, Hume argues that it does not make sense to hold individuals responsible for “temporary and perishing” actions, but rather, responsibility can only arise for actions done from stable and durable character traits. We do not have any textual support that Hume would agree with Locke’s view that sameness of consciousness is a necessary condition for moral accountability, but it is clear that for Hume, a necessary condition for moral accountability is that an action arises from stable character traits.

Since Hume shifts the focus towards stable character traits, it can be suggested that one possibility for Hume to adopt a Lockean distinction between persons and human beings is to regard persons as bundles of stable character traits. To be clear, Hume does not adopt such a view. Nevertheless, the interesting question is whether there is a principled reason why he does not, especially since such a view was in the air and discussed by Shaftesbury. 56 Shaftesbury, in his discussion of the self, emphasizes the ancient Delphic inscription “Know thyself!” and argues that self-knowledge is fundamental for any other inquiry and that through the guidance of self-knowledge we should aim to develop a stable character.57

For it is not certainly by virtue of our face merely that we are ourselves. It is not we who change when our complexion of shape changes. But there is that
which, being wholly metamorphosed and converted, we are thereby in reality
transformed and lost.

Should an intimate friend of ours, who has endured many sicknesses and run many ill
adventures while he travelled through the remotest parts of the East and the hottest
countries of the South, return to us so altered in his whole outward figure that, till we
had for a time conversed with him, we could not know him again to be the same
person, the matter would not seem so very strange nor would our concern on this
account be very great. But should a like face and figure of a friend return to us with
thought and humours of a strange and foreign turn, with passion, affections and
opinions wholly different from anything we had formerly known, we should say, in
earnest and with the greatest amazement and concern, that this was another creature
and not the friend whom we once knew familiarly. Nor should we in reality attempt
any renewal of acquaintance or correspondence with such a person, though perhaps he
might preserve in his memory the faint marks or tokens of former transactions which
had passed between us. (*Soliloquy*, 3.1, 127)

Here Shaftesbury goes so far to argue that a person cannot survive radical changes of
calendar. Let us return to Hume. How would Hume react to the example of a traveller who
has radically changed his or her character during a journey? In several passages, Hume
maintains that changes of character are possible and none of them support the view that
changes of character can result in loss of personal identity. In Book 1 of the *Treatise*, he
introduces the analogy with a republic or commonwealth to argue that “in like manner the
same person may vary his character and disposition, as well as his impressions and ideas,
without losing his identity” (*T* 1.4.6.19; SBN 261). I do not want to put much weight on this
passage alone, because for present purposes it is best to rely on Books 2 and 3. Fortunately, we
find additional support in Book 2. Hume makes clear that the indirect passions of love or
hatred cannot be caused by momentary and fluctuating actions, but only by actions that arise
from a stable intention or character. At the same time he acknowledges that it is possible that a
person changes her intentions by “repentance or a change of life” (*T* 2.2.3.4; SBN 349), and in
such a case the indirect passions that others express towards this person will also be altered.
He reiterates this point in the following passage;
Men are less blam’d for such evil actions, as they perform hastily and unpremeditately, than for such as proceed from thought and deliberation. For what reason? but because a hasty temper, tho’ a constant cause in the mind, operates only by intervals, and infects not the whole character. Again, repentance wipes off every crime, especially if attended with an evident reformation of life and manners. (T 2.3.2.7; SBN 412)

Since Hume speaks of “an evident reformation of life and manners” I take it that he wants not only to accommodate the possibility of changes of character, but also the possibility of radical changes of character. Thus he will not identify persons with stable bundles of character traits. To understand why, it is helpful to recall once more that sympathy plays a central role in Hume’s theory. The proposal that persons are stable bundles of character traits cannot easily be reconciled with Hume’s account of sympathy, because for Hume interactions with other people are not only essential in the formation of stable character traits but also in the correction of character traits that are disapproved of in one’s social community. While we aim to develop stable character traits through sympathy, sympathy presupposes the possibility of character change; otherwise it could not operate effectively.61

9. Conclusion: Explaining the differences
I hope to have shown that a careful examination of Locke’s and Hume’s underlying moral and religious background assumptions provides a powerful explanation of the puzzle as to why Hume does not follow Locke in offering different persistence conditions for persons and human beings, and why Locke’s conceptual distinction between persons and human beings is absent in Hume’s Book 2 account. I argued that Locke’s account of persons and personal identity is motivated by his aim to make sense of questions of moral accountability, by his metaphysical agnosticism concerning the materiality and immateriality of thinking substance, and by his religious beliefs in the possibility of an afterlife and a Last Judgement. I also showed that Locke thinks about moral accountability in a particular and controversial way. In light of these underlying moral and religious views and his metaphysical agnosticism, we can understand why Locke emphasizes the distinction between the terms “person,” “man,” and “substance” and argues that the persistence conditions for persons and human beings differ. While Hume shares Locke’s metaphysical agnosticism, he raises serious doubts concerning an
afterlife and a Last Judgement. Since he rejects Locke’s religious views, he could at most agree with Locke that sameness of consciousness is necessary for moral accountability. However, since he also challenges Locke’s thinking about moral accountability, which was tied to divine law, we do not have any evidence that he would agree with Locke on this point, and instead he emphasizes that accountability is only appropriate if actions arise from a stable and durable character. Moreover, Hume’s theory is meant to leave room for the possibility of character change. According to Hume, the characters of selves are shaped through sympathetic interaction with others, and in order to make sympathetic interaction possible and effective, it is important that selves are embodied and continue to have the same body. Hence, his theory in Book 2 does not have much scope for metaphysically distinct persistence conditions for persons and human beings, and subsequently it is neither important nor helpful to introduce a conceptual distinction between persons and human beings as Locke does. In Books 2 and 3, Hume regards persons as embodied social creatures and our sympathetic interaction with others is deeply embedded in his moral thinking. Hume rejects any religious foundation of morality as we find it in Locke, and we can see why notions of virtue and character development play a more important role in his moral philosophy than in Locke’s if we consider Hume’s moral philosophy in the context of other British moral philosophers who influenced his views.

NOTES

Earlier versions of this paper were presented at the Joint Session at the University of Exeter in July 2013, the University of St Andrews in August 2013, the Dutch Seminar in Early Modern Philosophy at Radboud University in February 2015, the University of Sydney in April 2015, Deakin University in April 2015, and the University of Otago in September 2015. I would like to thank all my audiences, the students in my graduate seminar at the University at Albany in the spring 2014, and Kenneth Winkler for very helpful discussions and feedback. I am also grateful to the anonymous referees for this journal for their constructive and helpful comments.

Soliloquy: Soliloquy or Advice to an Author, followed by Part and section number, followed by page number;
The Moralists: The Moralists: A Philosophical Rhapsody, followed by Part and section number, followed by page number;
Miscellany: Miscellaneous Reflections on the Preceding Treatises and Other Critical Subjects, followed by miscellaneous text number and chapter number, followed by page number.


3 I am here adopting Locke’s term that refers to human beings irrespective of sex or gender.

4 I do not want to suggest that Locke endorses a memory theory. I take it that Locke has good reasons to offer an account of personal identity in terms of consciousness rather than memory, because consciousness includes consciousness of the present and the future (see Locke, Essay, 2.27.10, 2.27.16–18, 2.27.24–26). However, memory, for Locke, is part of consciousness, and for this reason Locke and Hume agree that memory is relevant for personal identity. Thomas Reid used the terms “consciousness” and “memory” as exclusive and this explains why he criticized Locke for confounding memory with consciousness (see Reid, Essays on the Intellectual Powers of Man, 3.6, 277). Many scholars reject Reid’s criticism; they include Atherton, “Locke’s Theory of Personal Identity”; Matthews, “Descartes and Locke,” 28-30; Schechtman, Constitution of Selves, 105–12; Strawson, Locke on Personal Identity, chap. 9; Thiel, Early Modern Subject, 109, 122-26; Weinberg, “The Metaphysical Fact of Consciousness,” 405–12. For a recent defense of a memory interpretation of Locke, see Stuart, Locke’s Metaphysics, chap. 8.

5 Textual support for this view can be found in Essay 2.27.7: “It being one thing to be the same Substance, another the same Man, and a third the same Person, if Person, Man, and Substance, are three Names standing for three different Ideas; for such as is the Idea belonging to that Name, such must be the Identity.” See also Essay 2.27.15.

6 In the secondary literature the debate between defenders of relative identity interpretations and coincidence views has received much attention. Defenders of a relative identity interpretation tend to argue that persons and human beings are one and the same thing at a time, while defenders of a coincidence view argue that they are two distinct things. For a detailed and recent defense of a relative identity interpretation, see Stuart, Locke’s Metaphysics, chap. 7. Defenders of a coincidence interpretation include Chappell, “Locke and Relative Identity”; “Locke on the Ontology of Matter”; Kaufman, “Locke on Individuation.” I will bracket this dispute in this paper, because I do not regard it as central to Locke’s main task of specifying persistence conditions.

7 A similar view can be found in Alanen, “Personal Identity,” esp. 6.

8 See Boeker, “Moral Dimension”; Law, Defence of Mr. Locke’s Account; LoLordo, Locke’s Moral Man; Spector, “Grounds of Moral Agency”; Strawson, Locke on Personal Identity, esp. chaps. 1–3.
For further discussion of the relation between Locke’s 2.27.9 and 2.27.26 characterizations of persons see Boeker, “Moral Dimension.”


See Locke, *Essay* 2.28.5-8, and “Of Ethic in General.”

See Locke, “Of Ethic in General.”

See Locke, *Essay* 2.27.15–20, 2.27.22–23, 2.27.25–26; Locke, Letter no. 1693 in *Correspondence* 4:785–86.

For further discussion, see Garrett, “Locke on Personal Identity,” 100.


References to New Essays are to Leibniz, *New Essays on Human Understanding*, ed. Remnant and Bennett, hereafter “New Essays,” followed by Book, chapter, and section number, followed by page number.

For a related discussion, see McIntyre, “Personal Identity and the Passions.”

I here follow Locke and use the term “thought” in a broad sense to include perceptions and other mental states of which we have awareness.

In his correspondence with Stillingfleet, he tells us that we can know that thinking inheres in a substance, because thoughts cannot subsist by themselves, but we cannot know whether the substance is material or immaterial. See Locke, *Works of John Locke*, 4:37, hereafter cited as “Works.”

For further discussion of Locke’s interest in theology, see Nuovo, Introduction to *Writings on Religion*.


See also Burthogge’s letter to Locke, dated 19 September 1699, Letter no. 2617A in Locke, *Correspondence*, 6:685-86; Anstey, *John Locke and Natural Philosophy*, chap. 10, esp. 200–203, Anstey and Harris, “Locke and Botany.”


29 Boyle, endorses such a view and argues that the bodies to which the same soul is united can be very different in shape and size at different times. See Boyle, “Some Physico-Theological Considerations,” 205–206.

30 In Hume, *Essays Moral, Political, and Literary*, 590–98. See also Harris, *Hume*, 48–51; Russell, *Riddle of Hume’s Treatise*, chap. 17. According to Russell, irreligion is a dominant theme throughout Hume’s *Treatise*. “The key consideration here is the fact that Hume’s project in the *Treatise* is modeled or planned after Hobbes’s similar project in *The Elements of Law* and the first two parts of *Leviathan*” (248). Some Hume scholars maintain that Russell overemphasizes the dominant role that irreligion plays throughout the *Treatise* (see, for example, Harris, “Of Hobbes and Hume”; Millican, “Paul Russell”; Pyle, “Paul Russell.” For the purpose of my arguments in this paper it is sufficient that Hume does not endorse religious beliefs, and this is widely accepted among Hume scholars.

31 For further discussion see Winkler, “Locke on Personal Identity,” 211n12. Jolley, *Locke’s Touchy Subjects* maintains that Locke is more favourable to materialism than to immaterialism. I have doubts about Jolley’s thesis, because his discussion of immaterial views focuses on Cartesian views and does not give proper acknowledgment to non-Cartesian immaterial views such as the views of the Cambridge Platonists.


33 Actually Locke does not endorse the view that divine punishment is eternal. See Locke, “Resurrection et Quae Sequuntur,” in *Writings on Religion*, 232–37.


35 I use the term “perception” as Hume does to refer to any impression or idea (see T 1.1.1.1; SBN 1).


37 Hume does not elaborate on why he regards contiguity as irrelevant in T 1.4.6. However, we can find an explanation if we turn to T 1.4.5 (see Baier, *Progress of Sentiments*, 142; Garrett, *Cognition and Commitment*, chap. 8). There Hume argues that some perceptions lack spatial location (see T 1.4.5.10–16; SBN 235–40). On this basis, we can see why he regards contiguity as irrelevant with regard to our belief in personal identity in Book 1: Selves, as they are given in experience are bundles of perceptions and perceptions can lack spatial location. Hence it will not be possible to consider perceptions that lack spatial location as contiguous, at least in space.
McIntyre has done important work to emphasize that the forward looking aspect of Hume's Book 2 account supplements his Book 1 account. See McIntyre, “Hume and the Problem of Personal Identity” and “Personal Identity and the Passions.”

McIntyre is a major proponents of the view that the Book 2 account is consistent with the Book 1. See McIntyre, “Hume and the Problem of Personal Identity” and “Personal Identity and the Passions”. According to Annette Baier, it is important to acknowledge the overall narrative that Hume develops throughout the Treatise, and the Treatise has to be understood as a whole. The solitary view in Book 1 is supplemented by a social perspective in Book 2, where selves are portrayed as flesh and blood social creatures who interact with each other and seek the company of others. While Baier regards the Book 2 account as supplementing the Book 1 account, she also emphasizes that persons in Book 2 are embodied human beings (see Progress of Sentiments, chap. 6). Alanen, “Personal Identity,” argues that the political analogy of the self as a republic or commonwealth helps to understand how Hume’s Book 2 account complements his Book 1 account. Other interpreters who argued for the continuity between Book 1 and 2 include Capaldi, Hume’s Place in Moral Philosophy, 20–21, 168–84; Harris, “Compleat Chain of Reasoning.” Interpreters who emphasize the differences or gap between Hume’s Book 1 and Book 2 accounts include Penelhum, “Self of Book 1.”

See Greco, “Self as Narrative in Hume”; Schechtman, Constitution of Selves. I am not convinced that there is sufficient textual evidence, at least in the Treatise, to call Hume’s Book 2 account “narrative.” Greco, “Self as Narrative,” 711–14, and Pitson, Hume’s Philosophy of the Self, 92–96, both regard Hume’s Book 2 account of the self as narrative and primarily draw on textual evidence found in Hume’s EHU, section 3. Narrative theories of personal identity have been criticized, for instance, by Strawson, “Against Narrativity,” and Marya Schechtman has revised her earlier views in Staying Alive. I have doubts that the controversies that surround narrative theories of personal identity will help illuminate the interpretation of Hume’s Book 2 account. For this reason, I will not adopt the term “narrative.”

See Schechtman, Constitution of Selves, 1–2, 7–8.

To be fair, Greco acknowledges the interpretive disputes. See Greco, “Self as Narrative in Hume,” 702–706. Interpreters who question that Hume offers a metaphysical account of personal identity in Book 1 of the Treatise include Waxman, Hume’s Theory of Consciousness, chaps. 6–7. For further discussion, see Strawson, Evident Connexion, part 1.

I take this proposal to be in accordance with Ainslie’s view that there are person-defining features and that persons bear an “existential connection” to those features. See Ainslie, “Scepticism About Persons.”

For further discussion, see McIntyre, “Character: A Humean Account,” 200–205.

For further discussion of Hume’s account of sympathy, see Baier and Waldow, “Conversation between Annette Baier and Anik Waldow”; Taylor, Reflecting Subjects, chap. 2, “Sympathy, Self, and Others,” 188-205.

For further discussion, see Baier, Progress of Sentiments, chap. 6, esp. 136–42.

See Hume, T 2.1.11.5, 2.1.11.8, 2.2.4.2 (SBN 318, 320, 252). For further discussion, see Waldow, “Sympathy and the Mechanics of Character Change,” 233.

See Hume, T 3.3.1.10, 3.3.6.1 (SBN 577–78, 618).
See Haakonsen, *Natural Law and Moral Philosophy*, 51–58. While I regard the influence of the Natural Law tradition as important for understanding Locke’s thinking about the justification of morality, I do not want to suggest that Locke’s moral views fully match other traditional natural law theories. Some interpreters have argued that Locke’s appeal to natural law theory cannot easily be reconciled with his hedonism. See Aaron, *John Locke*, 250–69; von Leyden, *Introduction to Essays on the Law of Nature*, by John Locke, 71–78; Schneewind, “Locke’s Moral Philosophy,” 199–225. I am not convinced that there is a genuine tension, because he is a hedonist with regard to moral motivation and his appeal to divine (or natural) law concerns the justification or foundation of morality. Simply coming to understand by reason that we are bound by divine law does not mean that we are motivated to act in accordance with divine law. For this reason a hedonistic account of motivation can supplement a justificatory account of morality in terms of natural law. For further discussion of Locke’s account of motivation see Bolton, “Intellectual Virtue and Moral Law”; Weinberg, *Consciousness in Locke*, chap. 5.


Various interpreters have suggested that Hume’s moral philosophy can be seen as a virtue ethicist approach. See Hursthouse, “Virtue Ethics and Human Nature”; Schneewind, “Misfortunes of Virtue,” 50–54; Swanton, “Can Hume Be Read as a Virtue Ethicist?” However, Harris, “Hume on the Moral Obligation to Justice,” argues that Hume’s account of artificial virtues cannot be understood as virtue-ethical. Instead, Harris argues, it is better placed within natural law tradition.

See also Hume, EHU 8.30 (SBN 98–99).

In the following I do not take a stance on whether Shaftesbury’s considered view is that a person is to be identified with a stable character or set of stable character traits. Shaftesbury’s discussion of the self can be found throughout his *Characteristics*. Several of his works are written in dialogue form or contain inner dialogues, and notions of the self that are proposed at one stage are often challenged as the dialogue progresses. It is clear, however, that Shaftesbury was a critic of Locke’s theory of personal identity. For further details of his criticism of Locke see *Soliloquy*, 3.1, 127, *The Moralists*, 2.1, 253–54, *Miscellany* 4.1, 420–21 (all in *Characteristics*). See also Thiel, *Early Modern Subject*, 177–80; Winkler, “Personal Identity in Shaftesbury and Hume,” 6–9.

It is worth noting that the *Treatise* contains some passages that, at least prima facie seem to conflict with the possibility of character change. Hume writes: “it being almost impossible for the mind to change its character in any considerable article, or cure itself of a passionate or splenetic temper, when they are natural to it” (T 3.3.4.3; SBN 608). It is important to note that this passage concerns natural abilities. See also T 3.2.5.9 (SBN 521).

See Alanen, “Personal Identity,” for a good proposal how this passage bridges Hume’s discussion of the self in Books 1 and 2.
See also Hume, EHU 8.30 (SBN 98–99).

For further discussion, see Waldow, “Sympathy and the Mechanics of Character Change.”

WORKS CITED


Perronet, Vincent. A Second Vindication of Mr. Locke, Wherein his Sentiments relation to Personal Identity are clear'd up from some Mistakes of the Rev. Dr. Butler, in his Dissertation on that Subject. And the various Objections rais'd against Mr. Locke, by the Learned Author of An Enquiry into the Nature of the Human Soul, are consid'r'd. London: 1738.


Shaftesbury, Anthony Ashley Cooper, Third Earl of. *Several Letters Written by a Noble Lord to a Young Man at the University*. London: 1716.