

UNDERSTANDING LOCKE'S DISCUSSION OF PERSONAL IDENTITY

Kyle Driggers

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Philosophy in the College of Arts and Sciences.

Chapel Hill
2014

Approved by:

Alan Nelson

Thomas Hofweber

Mariska Leunissen

© 2014
Kyle Driggers
ALL RIGHTS RESERVED

ABSTRACT

R. Kyle Driggers: Understanding Locke's Discussion of Personal Identity
(Under the direction of Alan Nelson)

Almost all interpreters agree that Locke's goal in Book II, Chapter 27 of the *Essay* is to provide a metaphysical theory of personal identity. Though there are disagreements about the details of his theory, it is almost universally agreed that Locke does not argue for a theory of personal identity that can capture our common sense beliefs about the matter. I focus on two notable objections to Locke's theory that purport to show this. I argue that Locke has reasonable responses to these objections that have been overlooked by previous commentators, who have failed to appreciate Locke's claim that "person" is a forensic notion.

ACKNOWLEDGEMENTS

I would like to thank Alan Nelson for his diligent assistance with the writing of this thesis. Thanks are also due to Thomas Hofweber and Mariska Leunissen for their insightful comments. Finally, I owe special thanks to Jennie Minor, who gave me constant encouragement throughout the entire writing process.

TABLE OF CONTENTS

Section 1: Introduction.....	1
Section 2: Background: the standard picture of Locke’s “theory” of personal identity.....	2
Section 3: Two rival interpretations and their flaws.....	6
Section 3.1: Stuart’s “Simple Memory” Account.....	7
Section 3.2: Winkler’s Appropriation Account Entrenchment.....	11
Section 4: Getting clear on the cases and Locke’s responses.....	15
Section 4.1: The Brave Officer Case.....	16
Section 4.2: The False Memory Case.....	22
Section 5: Conclusion.....	25
WORKS CITED.....	27

1. Introduction

Book II, Chapter 27 of the *Essay* is widely understood as containing Locke's theory of personal identity. In that chapter, Locke famously emphasizes the role memory and consciousness play in the "sameness of a rational Being" (II.xxvii.9/335).¹ Almost all interpreters agree that Locke's goal in this chapter is to provide a *metaphysical* theory of personal identity.² However, interpreters disagree on the details of Locke's theory: some interpreters take Locke to hold the view that a person is identical to a past person if and only if the former remembers the actions and thoughts of the latter; some interpreters take Locke to hold the view that a person is identical to a past person if and only if the former "appropriates" the actions and thoughts of the latter.³

Regardless, it is almost universally agreed that Locke does not argue for a theory of personal identity that can capture our common sense beliefs about the matter. There are at least two notable objections to Locke's theory that purport to show that Locke's account is problematic. The first is Thomas Reid's Brave Officer Objection; the second is what I'll call the False Memory Objection.

¹ I will adopt the following standard scheme of referring to Locke's text: (II.xxvii.9/335) refers to the ninth section of the twenty-seventh chapter of the second book of Locke's *Essay*. The relevant passage occurs on page 335 of the Nidditch volume.

² Examples include Stuart, Winkler, Garrett ("Locke on Personal Identity, Consciousness, and 'fatal errors'."), Ayers (*Locke*, Volume II: Ontology), Chappell ("Locke on the Ontology of Matter, Living Things and Persons."), Strawson (*Locke on Personal Identity: Consciousness and Concernment*) and others. One notable exception is Newman ("Locke on Substance, Consciousness, and Personal Identity").

³ The view that memory is necessary and sufficient for personal identity is defended by Stuart, whose interpretation I discuss at length in 3.1. "Appropriation" is a technical term that Winkler uses in his introduction. The meaning of that term will be discussed in 3.2

In this essay, I argue that Locke has reasonable responses to these objections that have been overlooked by previous commentators. The resources for these responses in Locke's text have been overlooked because the cases presented by the objections are ill-formed in virtue of various ambiguities in key terms. I will show how to remove that ambiguity by breaking the cases presented by these objections into sub-cases and then showing how Locke can respond to the sub-cases one-by-one. My responses on Locke's behalf will help us get a better grip on his comments on personal identity.

2. Background: the standard picture of Locke's "theory" of personal identity

There is widespread disagreement about the details of Locke's account of personal identity, but it is widely understood that Locke rejects a few theses about personal identity before moving into his positive picture.

Locke's chapter on *Identity and Diversity* begins with an extended discussion of the identity conditions of non-persons: these include inanimate material bodies and living things (II.xxviii.1-8). Most of the details are irrelevant to our current discussion, except for Locke's distinction between *human beings* and *persons*. Locke tells us that much of the confusion about personal identity results from disputants not being clear on the kind of thing under discussion (II.xxvii.7/332). Our names for ourselves and our fellow humans ("Socrates," "I," etc.) are ambiguous between two ideas that depend on the perspective from which we consider ourselves and other humans.⁴

Locke reserves the name "human being" for when we take a third-person perspective on another human being and consider whether he is identical across time. The standard for his persistence is whether or not he continues to participate in the same life, whether his constitutive

⁴ I will use "idea" exclusively in Locke's sense for the entirety of this essay. Ideas are "whatsoever is the Object of the Understanding when a Man thinks" (I.i.8/47). Locke's extended discussion of the different sources and kinds of ideas begins in the second book of the *Essay*.

matter is vitally organized (*Ibid.*). Locke's discussion about personal identity is not about the identity of human beings.⁵

Locke reserves the name "person" for the kind of thing that we are interested in when we argue about personal identity. And much of the discussion over the nature of persons is an attack on the view that persons are thinking substances. The main purpose of Locke's famous thought experiments is to illustrate why "person" cannot correspond to an idea of a substance.

According to Locke, substance ideas *in general* are confused: we have only a "supposition" that there are substances in which various correlated qualities inhere (II.xxiii.2). Locke illustrates this point in his discussion of personal identity by considering cases where (1) the substance that purportedly constitutes a person is switched for another substance without a resultant change in our ideas (2) a single substance constitutes two different persons at different times without any resultant change in our ideas (II.xxvii.12/337). The details don't concern us here: the upshot is that confused ideas of substances can't correspond to propositions about personal identity because changes in substance don't result in corresponding changes in our ideas.

Locke's rejection of the theses that persons are either substances or human beings is intertwined with his positive story about that in which personal identity consists. He begins his discussion by an explicit definition of "person" as "a thinking, intelligent Being" that can think about itself at different times by "that consciousness, which is inseparable from thinking" (II.xxvii.9/335). Understanding Locke's talk of consciousness is the locus of debate for the various interpretations of Locke's account of personal identity. Locke tells us that,

5 Some interpreters hold that Locke's distinctions between persons, masses of matter, human beings, and particles suggest that was a relativist about identity. That is, the identity of a thing is relative to the sortal under which it is being considered. There are no sortal-independent identity conditions, nor are there identity conditions for "dummy sortals" like "thing." See the section on Stuart for further discussion. Also, see Chappell (1989).

For since consciousness always accompanies thinking, and makes every one to be, what he calls *self*; and thereby distinguishes himself from all other thinking things, in this alone consists *personal Identity*, *i.e.* the sameness of a rational Being: And as far as this consciousness can be extended backwards to any past Action or Thought, so far reaches the Identity of that *Person*...(II.xxvii.9/335)

Whatever consciousness turns out to be, it must play the crucial role in the sameness of persons. The first role is the individuating role: consciousness distinguishes between persons at any given moment, and is that by which persons are identical to other persons across time.

The notion of “extending consciousness” is typically understood as Locke’s oblique reference to memory and many interpreters of Locke’s comments take memory to play the central role in his account. Thus, it is insofar as I can remember the idea of some thought or action that I am the person that had that thought or performed that action. However, Locke tells us that we have consciousness of present actions and that we can extend consciousness into the future (II.xxvii.10/336). So it’s reasonable to think that while memory is an extension of consciousness, the past is not the only direction that consciousness can extend.

Consciousness apparently plays another role: the role of what Winkler calls appropriation. What we are conscious of partially determines what we are concerned for and what actions and thoughts we are willing to ascribe to ourselves (II.xxvii.16-17/341). Locke tells us that when we extend our consciousness via. memory to previous thoughts and actions we become “concerned” and “accountable” for those actions. We *appropriate* those actions to ourselves (II.xxvii.26/346). It is especially important that we appropriate certain actions because some of those actions will be righteous, and others sinful, and God will judge us accordingly (*Ibid.*).

Popular interpretations of Locke disagree on which is the primary role of consciousness in personal identity. In the next section, we will consider an interpretation that emphasizes the individuating role and specifically how memory unites persons across time. We will also consider an interpretation that emphasizes the appropriative role.

However, regardless of what the details of Locke's account of personal identity are, it is widely assumed that there are forceful objections to his account. I will consider the two most important objections: the Brave Officer Objection and the False Memory Objection.

Virtually all interpreters agree that Locke held that memory is in some sense necessary for personal identity.⁶ Thomas Reid's Brave Officer Objection purportedly shows that a person's remembering the actions and thoughts of a previous person is not *necessary* for the identity of the former to the latter.⁷ Reid asks us to consider the following case,

Suppose a brave officer to have been flogged when a boy at school, for robbing an orchard, to have taken a standard from the enemy in his first campaign, and to have been made a general in advanced life: Suppose also, which must be admitted to be possible, that when he took the standard, he was conscious of his having been flogged at school, and that when made a general he was conscious of his taking the standard, but had absolutely lost the consciousness of his flogging (Reid 217-8).

According to Reid, Locke must agree that the boy is identical to the brave officer, that the brave officer is identical to the general, but that the boy is not identical to the general. In other words,

⁶ See Stuart and Garrett. Garrett differs from Stuart by arguing that remembering that one remembers some thought or action is necessary for personal identity.

⁷ It should be noted, as Winkler does, that the idea behind this objection originates with Berkeley (206). I will consider Reid's version because it is forceful and clear.

Locke must abandon the transitivity of identity. And given that the transitivity of identity is a logical truth, it appears as if Locke holds an absurd theory.

Most interpreters also agree that Locke held that memory is sufficient for personal identity. The False Memory Objection purportedly shows that a person's remembering some thought or action is not sufficient for his being identical to the person that had that thought or performed that action. By extension, it shows that a person's *appropriating* some action or thought is not sufficient for his being identical to the person that had that thought or performed that action.

Consider the following case: Mary is sure that she remembers taking out the trash and thus appropriates the action of taking out the trash to herself. But, as a matter of fact, someone else took out the trash when Mary claimed she had. But since Mary remembers taking out the trash and thus appropriates that action, it must be the case that Mary is identical to the person that took out the trash. That is an absurd result.⁸

In the next section, I will argue that two popular interpretations of Locke's account of personal identity do not provide Locke with the resources necessary to respond to these objections.

3. Two rival interpretations and their flaws

In this section, I consider two divergent metaphysical interpretations of Locke's discussion of personal identity: Matthew Stuart's Simple Memory Account and Kenneth Winkler's Appropriation Account. I will consider each view in turn: first, I will explicate the view, briefly discuss why the text would seem to warrant such a view, present the objections to the view, and discuss why the view cannot adequately respond to these objections. Their lack of

⁸ See II.xxvii.13/338 for a discussion of consciousness-switching that can be construed as a worry about false memories. I won't use Locke's example because it unnecessarily complicates my presentation of the objection.

an adequate response does not entail that Locke did not hold the theory they attribute to him, but it does provide us with motivation to search for alternative interpretations, if we aim to be charitable to Locke.

3.1 Stuart's "Simple Memory" Account

According to Stuart, Locke adopts the 'simple memory' theory on which a person is identical to a previous person if and only if the former can remember the ideas of the thoughts and actions of the latter from the latter's first-person perspective (352). (This rules out remembering the actions of other persons from a third-person perspective that would be irrelevant to *personal* identity.) This theory emphasizes the extension of consciousness to past events. Remember that Locke explicitly says that,

...as far as this consciousness can be extended backwards to any past Action or Thought, so far reaches the Identity of that *Person*...(II.xxvii.9/335).⁹

Further text supports the interpretation that by "extending consciousness" Locke just means, "having the ability to recall the relevant action or thought":

For as far as any intelligent Being can repeat the *Idea* of any past Action with the same consciousness it had of it at first, and with the same consciousness it has of any present Action; so far it is the same *personal self*. (II.xxvii.10/336)

The ability to "repeat the *Idea* of any past Action" is just Locke's definition of memory (II.x.2/150).

Note that Stuart's interpretation ignores two relevant parts of the text that were mentioned earlier. First, he ignores the fact that consciousness can be extended into the future. In fact, Stuart considers it a *limitation* of Locke's theory that it cannot handle the survival of

⁹ All of the underlining in quoted text is mine.

persons into the future and is instead focused on the appropriation of past thoughts and actions (388). So, his interpretation is restricted to the identification of present and past persons.

Second, Stuart doesn't follow Locke's text explicitly: Locke seems to suggest that a person's *ability* to extend consciousness backwards to the thoughts and actions of another person is what is necessary and sufficient for his being that person. Stuart is not careful with Locke's appeal to the ability to remember and so is forced to say that Locke allows persons to have "gappy" existences such that my forgetting some action or thought entails that I am not the person that performed that action or had that thought (378).

Because Stuart is not careful with these details, his Locke is subject to the Brave Officer Objection. According to Stuart, remembering some action or thought of a previous person is necessary for being identical to that person. The Brave Officer Objection purportedly shows that this claim leads to the absurd result that Locke must deny the transitivity of identity.

Stuart is aware of this objection and the consequences it has for his interpretation. However, Stuart's response on Locke's behalf is unsatisfying: he argues that Locke should reject the transitivity of identity when the identity relation obtains between persons, even if that prevents Locke from capturing some of our intuitions about the case (379).¹⁰

According to Stuart, Locke maintains that the identity relation is always sortal-dependent. The general, officer, and boy fall under two distinct sortals that are relevant to considerations of personal identity: the "person" sortal and the "human being" sortal. Remember that for Locke, the conditions under which something counts as a human being are different than the conditions under which something counts as a person. Human beings persist like other biological organisms,

¹⁰ In fact, Stuart repeatedly emphasizes that Locke's theory is a failure insofar as its goal is to capture our ordinary intuitions about personal identity (See § 54 of his chapter "Persons").

by participating in the same life (II.xxvii.6/331). According to Stuart, persons persist by remembering their actions and thoughts.

Stuart contends that there is no sortal-independent identity relation. So when we determine whether two things are identical with one another, we must first fix which sortal we are considering them under. There are two sortal-dependent identity relations that are relevant for this case: the “is the same person as” relation and the “is the same human being as” relation (379). In Reid’s case, the “is the same human being as” relation holds between the general, officer, and boy just in case all three participate in the same life and independently of whether one remembers the actions of another. So, Locke can hold that the “is the same human being as” relation is transitive, perhaps capturing some of our intuitions about the case.

However, Stuart believes that Locke should respond to Reid’s objection by arguing that the “is the same person as” relation is not transitive: that the brave officer is the same person as the boy and the general is the same person as the officer does not entail that the boy is the same person as the officer. Reid’s case does not present an objection; it illustrates a central detail of Locke’s theory.

Unfortunately, Stuart’s response on Locke’s behalf seems *ad hoc*: the view that this relation is intransitive appears to be a stipulation motivated solely by an objection rather than a principled consequence of Locke’s theory. And, insofar as one finds Reid’s contention that the general and the boy are identical to be an intuitive feature of his case, one can worry that Stuart has saddled Locke with a counterintuitive theory of personal identity. So, we might have reason to worry about the charity of his interpretation. If Stuart’s response on Locke’s behalf is the correct one, it deserves a better motivation and an explanation as to why Locke’s commitment to the non-identity of the general and the boy is not counterintuitive.

Stuart's Locke is also subject to the False Memory Objection, since Stuart makes remembering an action or thought sufficient for being the person that performed that action or had that thought. Stuart is aware of this objection and spends considerable time explaining why Locke *cannot* give a reasonable response to it.

Stuart's primary focus is on a passage in which Locke apparently discusses false memories. Locke writes,

But that which we call the *same consciousness*, not being the same individual Act, why one intellectual Substance may not have represented to it, as done by it self, what it never did, and was perhaps done by some other Agent, why I say such a representation may not possibly be without reality of Matter of Fact, as well as several representations in Dreams are, which yet, whilst dreaming, we take for true, will be difficult to conclude from the Nature of things. And that it is never so, will by us, till we have clearer views of the Nature of thinking Substances, be best resolv'd into the Goodness of God who ... will not by a fatal Error of theirs transfer from one to another, that consciousness, which draws Reward or Punishment with it (II.xxvii.13/338).

Stuart suggests that there are two ways of reading this passage, both of which have Locke confused about the problem of false memories and thus have him making a mistake.

The first reading is that, despite what the text might suggest, Locke does not address the problem of false memories in general, restricting his focus to cases where consciousness are switched between substances. Locke's worry on this reading is that, if persons are substances, such a switch would result in the transfer of memories among persons and thus would lead to the wrongful punishment of one person for another one's actions. On this reading, the best Locke

can do is lamely assure us that God will prevent any such switching or will correct it before the Last Judgment (Stuart 374-5). The problem with this reading, of course, is that it does not provide a reasonable solution to the False Memory Objection because it only considers one kind of false memory.

The second reading, Stuart's preferred reading, is that Locke "shifted his attention" while considering the general problem of false memories and took his solution to the special cases of substance-switching to hold for all cases of false memories (374). Stuart hypothesizes that Locke confused the subclass of substance-switching cases with the class of false memory cases because he wrongly thinks that he doesn't have to handle cases of false memories that don't involve switching substances because Locke denied the possibility of false memories in these other cases. Stuart tells us that Locke's attempts to "reassure us" that there are no false memories "is misguided." As a matter of fact, says Stuart, we do have false memories and Locke must mistakenly say that we are identical to persons to which we are obviously not identical (374-6).

So, unlike in the Brave Officer case, Stuart doesn't attempt to give a reasonable response on Locke's behalf to the False Memory Objection. I will argue later that Stuart is misreading this passage. Locke's account of memory straightforwardly stipulates that there cannot be false genuine memories, for he defines memory as the revival of ideas that were previously had by the person who remembers them (II.x.1/193). We'll see below that Locke can give a reasonable account of apparent memories.

3.2 Winkler's Appropriation Account

According to Winkler's Appropriation Account, Locke held that it is my *appropriating* certain actions and thoughts to myself that is sufficient for my being the person that performed

those actions and had those thoughts. Winkler's account finds support in the following piece of text:

Person, as I take it, is the name for this *self*. [...] It is a Forensick Term appropriating Actions and their Merit; and so belongs only to intelligent Agents capable of a Law and Happiness and Misery. This personality extends it *self*[*sic*] beyond present Existence to what is past, only by consciousness, whereby it becomes concerned and accountable, owns and imputes to it *self* past Actions, just upon the same ground, and for the same reason, that it does the present.

(II.xxvii.26)

Winkler tells us that this bit of text along with others wherein Locke makes mention of appropriation and ownership of actions and thoughts suggests that memory plays only a secondary role in Locke's theory of personal identity to appropriation (II.xxvii.14/339, II.xxvii.16/341). That is, memory can provide me with thoughts and actions that I appropriate to myself, but it is the appropriation that plays the key role in personal identity.

According to Winkler, Locke held that the self is in some sense "subjectively constituted": when I represent some action or thought as belonging to my self, I am thereby identical to the person that had performed that action or had that thought (205). However, appropriation is not something that we have control over, according to Winkler. It is not as if I can decide which of my, say, memories I *want* to assign to my self. Instead, Winkler writes, "I accept what my consciousness reveals to me" (206).

Winkler claims that his Locke can give an adequate response to the Brave Officer Objection (207). In fact, he claims that the objection establishes that there is a logical "limit" on

the authority of the self over its own constitution. To see what he means, consider how he deals with the Brave Officer case:

The general and the young officer are the same, and it is therefore true of the general that he at one time recalled some of the thoughts or actions of the schoolboy. At the moment the general is unaware of those thoughts or actions. But they are nonetheless parts of his self because they were appropriated by the young officer, whose acts of appropriation the general cannot disown without calling his identity with the officer into question (207).

While the subjective constitution of the self is in many cases determined by memory, memory is not necessary for appropriation, so long as the relevant links of appropriation are present between selves. Winkler argues that the transitivity of identity *forces* the general to acknowledge that he is the schoolboy because he is identical with someone that appropriated the thoughts and actions of the schoolboy. So, I am identical with a person whose actions and thoughts I appropriate, and that appropriation can be determined by a law of logic. Apparently, Locke should accept the transitivity of identity and agree with Reid that the boy is identical to the general.

Winkler's response to the Brave Officer Objection is unacceptable for three reasons. First, it appears to be an addition to Locke's account by Winkler on Locke's behalf. No text in the *Essay* suggests that a law of logic could play the role that Winkler suggests it does.¹¹ Second, Locke's comments about appropriation suggest that I can only appropriate those ideas to myself that I can bring before my conscious awareness (II.xxvii.11/336-7, II.xxvii.14/339,

¹¹ In fact, Locke would be suspicious of the "force" that Winkler ascribes to laws of logic. His discussion of laws of logic and other maxims (IV.vii.4/592-3) suggests that the certainty of the self-evidence of "general propositions" like "If A=B and B=C, then A=C" is wholly dependent on the perceived agreement between ideas that correspond to instances of those general propositions. So, Locke would presumably reject the contention that the general is "forced" to appropriate the action of the child *because* of the "self-evidence" of a general proposition.

II.xxvii.17/341). And the general cannot do that, regardless of what logic commits him to saying. Finally, it's not clear whether considerations of logic could provide the general with the *concern* for the plight of the schoolboy that apparently comes with the appropriation of an action (II.xxvii.26/346). So, while Winkler's interpretation does purportedly allow Locke to maintain the transitivity of identity, it does so on grounds that Locke would not obviously accept.

In any case, Winkler's account is subject to the False Memory Objection, since memory is sufficient for appropriation. In our case, Mary remembers taking out the trash and so appropriates that action to her self. That is sufficient for her being identical to the person who took out the trash. But that is an absurd result.

Winkler doesn't explicitly consider the False Memory Objection because he focuses on Locke's worry about substance-switching that we mentioned earlier (222-3). However, he claims that worries about false memories can be allayed if we consider the fact that the appropriations of selves are criticizable from the perspective of the person that appropriates those thoughts and actions.

He tells us that the appropriations of a particular self,

...cannot be criticized from a viewpoint wholly external to the self. Any

legitimate criticism must be based on the self's own appropriations. It cannot be

said that some act or thought should have been appropriated if it never was (208).

Following this reasoning, it would be inappropriate for Tom, who actually took out the trash, to criticize Mary for appropriating that action to herself by insisting that her memory was faulty. At best, what Tom could do would be to get Mary to see that her own story about herself is inconsistent or otherwise flawed by appealing to other actions or thoughts that she appropriates.

Unfortunately, Winkler's response to this objection also seems hopelessly *ad hoc*. It is not an obvious consequence of the view that the self is subjectively constituted that that constitution is immune to external criticism. That is, Mary's self might be subjectively constituted but it seems reasonable that Tom could criticize her appropriations even though they are not his appropriations. For example, Tom could provide evidence that Mary's memory is often faulty. Furthermore, there is certainly no text from the *Essay* that independently supports the view that the appropriations of a self are not criticizable. Furthermore, one might find the view that Winkler ascribes to Locke counterintuitive: while appropriation plays a crucial role in the constitution of the self, it appears too strong to say appropriation is not criticizable by a third-party as Mary's case suggests. Thus, we have reason to reject his interpretation of Locke.

4. Getting clear on the cases and Locke's responses

We have just seen that two prominent interpretations of Locke's comments on personal identity fail to provide him with an adequate response to the Brave Officer and False Memory Objections. I think that failure results from both not charitably interpreting Locke's text and not being clear on what situations the Objections present to Locke. It's important that we are clear on what situations the cases present before we can show how Locke should respond to the cases. As we will see, Locke would handle different interpretations of these cases in different ways and understanding that fact is crucial for understanding why these cases present no significant threat to Locke's account. In this section, I break up the cases presented by the objections into different sub-cases and show how Locke can give an adequate response to each sub-case.

4.1 The Brave Officer Case

The Brave Officer Objection presents us with a case in which an officer is conscious of the flogging of a schoolboy and a general is conscious of the officer's taking a standard, but the general is not conscious of the flogging.

The first question we should ask is: from whose perspective are we considering this case? Recall that one of the recurring themes in Locke's comments on personal identity is the difference between the conditions under which human beings are identical and the conditions under which persons are identical. We could be considering the case from the point of view of someone that is not the general, officer, or boy. In that case, we could be trying to determine whether the general, officer, and boy are the same *human being*. That presents us with the following sub-case:

BO1: The general, officer, and boy are considered as human beings, not persons.

If **BO1** were the case that Reid had in mind, then Locke's response would be quick. The consciousness of each person is irrelevant to their being identical human beings. The boy and the officer participate in the same life, as do the officer and the general, so the boy is the same human being as the officer. Either Reid's contention about intransitivity of identity is confused or his intuition about the identity of the general and the boy originates from his confusion over what the names of the different persons apply to.

However, **BO1** does not appear to be the most charitable interpretation of Reid's case. After all, Reid is explicit that it is consciousness of certain thoughts and actions that unite the boy and officer and officer and general. So, we need to turn to another sub-case.

Reid uses Locke's language of "consciousness of" certain thoughts and actions thought his presentation of the Brave Officer Case. We can be reasonably clear about what Locke had in

mind with his usage of “consciousness of” certain thoughts and actions, but Reid’s case doesn’t tell us precisely *what* each of the persons is conscious of.

According to Locke, “Consciousness is the perception of what passes in a Man’s own mind” (II.i.19/115). To say that I am conscious of some thought or action is to report whatever I am currently thinking of, perceiving, and so on.

Locke sometimes talks of my consciousness of past and future actions. My consciousness can be “extended” into the past via. memory and into the future presumably through a variety of different mental operations like planning, predicting, prophesying, and so on (II.xxvii.10/336).¹² To say that I am conscious of some past event is just to say that I am currently remembering that past event. To say that I am conscious of some future event is just to say that I am currently imagining how that future event will play out, for example.

Again, Reid tells us that “when [the officer] took the standard, he was conscious of his having been flogged at school”, that the “when [the officer was] made a general he was conscious of his taking the standard”, but that the general has lost consciousness of the flogging at that time. All of those statements are in the present tense and suggest that the consciousness of the flogging and the taking of the standard were contemporaneous with the taking of the standard on the promotion of the officer to general. And when the general was being promoted, he was not conscious of his flogging. This interpretation has Reid meaning “is presently aware or thinking of” by “is conscious of.” That presents us with the following case:

BO2: The general can remember the flogging, but is not presently doing so.

If **BO2** is the correct characterization of Reid’s case, then he ascribes the following view to Locke: a person is identical to a previous person whenever the former remembers an action or thought of the latter.

¹² Locke is never explicit about what it would mean to extend consciousness into the future, so I am speculating here.

If this is the correct interpretation of Reid's case, then Reid has badly misunderstood Locke's comments on extending consciousness. Locke insists that it is a person's *ability* to remember some thought or some action that is necessary for his being the person that had that thought or performed that action. It is insofar as a person "can repeat the *Idea* of any past Action" of some other person that he is identical with that person. In this case, it's still possible that the general could remember the flogging, say, once the ceremony of his promotion is over. In that case, the general is still identical to the boy. So, on this interpretation of Reid's case, Reid has not shown that Locke must abandon the transitivity of identity.

Perhaps this is not the most charitable interpretation of Reid's case. After all, Reid says that the general has "absolutely lost consciousness" of the flogging. Perhaps what this means is that the general can no longer remember the flogging.

I will assume that Reid is being fair to Locke and that he adopts Locke's account of memory when presenting this objection. It will be useful to review what that account is.

For Locke, we get memories from the impressions that previous perceptions have left on our minds (II.x.4/151). The faculty of memory is the mind's ability to "revive again in our Minds those *Ideas*, which after imprinting have disappeared..." (II.x.2/149). That means that memories are not ideas that are literally stored in our minds but escape our conscious awareness (II.x.2/150). When I remember something, I am able to revive the perception I previously had, except in memory that perception has "an additional Perception annexed to [the original perception], that it has had [it] before" (*Ibid*).

So, for Locke, to lose the memory of some thought or action is to lose the ability to revive the relevant perception that the mind once had of that thought or action. In the Brave

Officer case, to say that the general can no longer remember the flogging is to say that the general has lost that ability.

However, to say that the general can no longer remember the flogging is itself ambiguous, because what the general *can* do is relative to certain restrictions. It's not clear what situation Reid's case presents because he doesn't tell us *why* the general can't remember the flogging or if there are any circumstances under which *could* remember the flogging. We can imagine a range of cases:

BO3.1: The general cannot currently remember the flogging, but could if his memory were appropriately jogged.

BO3.2: Because of the trauma of the flogging, the general cannot remember the flogging until he undergoes extensive psychotherapy.

BO3.3: The general cannot remember the flogging because he has succumbed to global amnesia.

There are possibly infinitely many cases that we could consider here. However, I think we can present the case at a level of generality so that we can make an important point about how Locke can respond to Reid's objection.

God's ability to cause the general to remember the flogging will play a key role in later discussion, so I will lump **BO3.1-BO3.3** into a single case, described at a high level of generality.

BO3: God can cause the general to remember the flogging, but some other set of circumstances prevents him from doing so.

Now, consider the range of cases in which the general is in a set of circumstances such that not even God could bring about his remembering the flogging. We might conceive of this case as the case wherein it is metaphysically impossible for the general to remember the flogging. This is the loosest restriction we can put on the case that's still relevant for this discussion.

BO4: The general cannot remember the flogging and even God can't make it such that he can.

To understand Locke's response to **BO3** and **BO4** and the false memory cases we will discuss in the next subsection, we need to take a brief step back and consider why Reid introduces these cases in the first place.

Reid, along with Stuart and Winkler, presuppose that Locke's goal in his discussion of personal identity is to give a metaphysical account of the conditions under which persons are identical across time. And, they suppose, whatever account he gives should be general: it should tell us what persons are identical no matter what times we consider. The point of the objections is to show that Locke's purported theory of personal identity fails because it produces false results in cases like **BO3** and **BO4**.

But it's not obvious from Locke's text that this is his interest in his discussion of personal identity. His arguments against substance accounts of personal identity might suggest to the reader that his goal is to supplant those accounts with an account that involves the appropriate consciousness or memory, and so on. But note that his arguments against these accounts are *never* metaphysical and only epistemological. The upshot of his thought experiments is always that we lack knowledge of personal identity insofar as it depends on the identity of substances (*e.g.*, II.xxvii.12/337-8).

Instead, Locke's goal seems to be to clarify our thinking about disputes about personal identity and that goal is achieved partially by criticizing the role of substances in our thinking and partially by fixing what ideas to which words like "person" apply. As was mentioned before, "person" is a forensic term for Locke: actions and thoughts are appropriated to persons so that they can be judged by God based on the merits of those actions (II.xxvii.22/344, II.xxvi.26/347).

Since God is just, he must only punish and reward persons for the actions *they* perform during their earthly lives.

Locke's suggestion is that consciousness of past thoughts and actions is what will constitute persons in the afterlife regardless of what that consciousness adheres to, be it a mental substance or a glorified body (II.xxvii.26/347). So, which actions a person can remember determines which actions that person will be judged for. Of course, there are cases like **BO3** in which a person temporarily loses consciousness of some past action while they inhabit their earthly body. But during the Last Judgment, Locke is confident that any such memory will be restored so that the person can be judged accordingly.

Locke writes,

...the Apostle tells us, that at the Great Day, when every one shall *receive according to his doings, the secrets of all Hearts shall be laid open*. The Sentence shall be justified by the consciousness all Persons shall have, that they *themselves* in what Bodies soever they appear, or what Substances soever that consciousness adheres to, are the *same*, that committed those Actions, and deserve that Punishment for them (II.xxvii.26/347)

If God justly punishes or rewards us for our earthly actions and those actions are only ascribable to us if we remember them, then it must be that God's ability to "lay the secrets of all Hearts open" includes his ability to revive the memories of those past actions in his subjects.

That means that Locke is only interested in the conditions under which actions are appropriated to persons under the ideal conditions of the Great Day. So, a person is identical to a previous person on the Great Day if he can remember the actions and thoughts of that person with assistance from God. So, if Reid's case is **BO3**, then Locke has a natural response to the

objection: it may be that the general cannot remember the flogging for any number of ordinary reasons, but those reasons are irrelevant if God can assist the general in remembering the flogging on the Great Day. In that case, the general is identical to the boy and transitivity is preserved.

But consider the case where even God cannot revive the memory in the general, **BO4**. Here, Locke must admit that the general is not identical to the boy and thus must deny the transitivity of identity. Reid has apparently found a counterexample to Locke's account.

Locke should accept this result, but it's not clear that it comes at any cost to his account. In fact, it is likely that there are no cases like **BO4**: God simply has the requisite power to enable the persons to remember his past actions and thoughts. Any cases in which God lacks this power are an affront to his omnipotence or omnibenevolence. The onus is on Reid to produce a case like **BO4** that is a legitimate challenge to Locke's account.

Thus, it appears, upon further reflection, that the Brave Officer Objection is not successful.

4.2 The False Memory Case

In the first section, we considered a case wherein Mary has the false memory of taking out the trash. Because Locke purportedly holds that memory is sufficient for personal identity, he apparently must hold that Mary is the person that took out the trash, an absurd result.

This False Memory case is under-described because it's not clear what a false memory amounts to. I believe there are at least three sub-cases of the False Memory Case because there are at least three different ways to understand what a false memory is.

Locke anticipates one sub-case in the discussion of false memory that Stuart discusses (II.xxvii.13/338). However, Stuart ignores the context in which that case is presented and so

presents Locke as having a lame response to it. Understanding that context is important for understanding Locke's response to the case.

Locke considers this case as part of his attack on the view that personal identity is dependent on substance identity. His primary question for those who hold such a view is whether "consciousness of past Actions," that is, memories, can be transferred from one substance to another. If they can, then it would appear that divine justice might be meted out unfairly, since consciousness of past actions "draws Reward or Punishment with it" (II.xxvii.13/338). That is, if one person's memory of committing a sin can be transferred to another person, then the latter person would presumably be culpable for that sin. We can use these comments to create our first sub-case:

FM1: Mary falsely remembers taking out the trash because Tom's memory of taking out the trash has been transferred to her constitutive substance.

Locke's response to this sub-case might seem unimpressive out of context: he writes that "the Goodness of God" is the only guarantee that this sort of memory-switching doesn't occur, or at least that every switch will be reversed before the Last Judgment (II.xxvii.13/338). But remember, this comment comes as part of an extended attack on substance views of personal identity. Those attacks always take the following form: we only have a confused idea of substance and changes of substance don't result in corresponding changes in our ideas. Given the similarity of the memory-switching case to the other substance switching cases (II.xxvii.15, II.xxvii.19, II.xxvii.23), it appears that Locke's goal in presenting this example is to show why substance views of personal identity are misguided. That's not to say that there couldn't be substances that we are annexed to, just that, if there are, our ideas of them are so confused that when we talk about personal identity, we aren't talking about substances. If substances do play a

role in personal identity, unbeknownst to us, then divine justice requires that any memory-switching that might have occurred be reversed before the Last Judgment.

Of course, the natural objection to this response is that it still allows for false memories while Mary is in her Earthly body. That is, in this strange scenario, Mary is still the person that took out the trash even if that memory will be transferred back to Tom before the Last Judgment. Locke still hasn't given us an adequate account of personal identity. However, as I argued earlier, Locke's goal is not to provide a metaphysical account of personal identity. He is only interested in how actions are appropriated to persons for the purposes of the Last Judgment. So, the objection is dialectically inappropriate.

Of course, ordinary cases of false memories don't apparently involve the switching of memories between substances. I take it that there are two other ways that a memory can be "false": either the person in question has a memory of something that never happened or the person has some non-memory idea of a thought or action that appears to them to be a memory. We then have two more sub-cases:

FM2: Mary genuinely remembers taking out the trash, but never took out the trash.

FM3: It appears to Mary as if she remembers taking out the trash, but she doesn't genuinely remember taking out the trash.

Stuart hypothesized that Locke mishandled cases of false memories because Locke confusedly rejected the possibility that someone could have a memory of something that never happened (374-6). However, Locke was not confused when he rejected the possibility of false memories outright: his definition of memory excludes their possibility.

Remember that memories are the impressions left by previous perceptions had by the mind (II.x.149) and that memory is the mind's power to "revive Perceptions, which it has once

had, with this additional Perception annexed to them, that it has had them before” (II.x.2/150).

That means that every genuine memory must have its source in some previous perception had by the mind: there cannot be genuine memories of perceptions that the mind never had. Strictly speaking, then, Locke’s definition of memory rules out cases like **FM2**.

To account for cases like **FM3**, Locke needs a story about how an idea can appear as if a memory to a person when that idea is not a genuine memory. The key to this story is the “annexed perception” that Locke refers to when describing the nature of remembered ideas. This annexed perception is what phenomenologically “marks” the remembered idea as a memory, rather than as, say, an imagination (Brown 17). Presumably, in cases like **FM3**, this “additional perception” is annexed to ideas that weren’t previous perceptions. In these cases, ideas that appear to be memories are just ideas from imagination or some faculty other than perception with this “additional perception” annexed to them. That is, they are ideas that appear as if they are memories but in fact are not the result of imprinting caused by previously perceived thoughts and actions.

But Locke never says that a person has successfully extended their consciousness in cases where an idea appears to be a memory. Consciousness only extends when “[an] intelligent being can repeat the *Idea* of any past Action with the same consciousness it had of it at first...” (II.xxvii.10/336). So, the case presented by **FM3** presents no threat to Locke’s account.

5. Conclusion

In this essay, I’ve defended Locke against two prominent objections, the Brave Officer Objection and False Memory Objection. I showed how previous interpreters were unable to respond on Locke’s behalf to these objections because they falsely interpreted Locke’s comments on personal identity. I emphasized in my response to the objections that Locke’s

primary goal in his chapter on personal identity is not to provide a metaphysical theory, but to develop the idea that “person” is a forensic term for appropriating actions to persons for the purposes of divine judgment. This fact coupled with careful clarification of the cases presented in the two objections shows that these objections do not have the weight that they were previously assumed to have.

WORKS CITED

- Ayers, Michael. *Locke*. Vol. II. London: Routledge, 1991. Print.
- Chappell, Vere. "Locke on the Ontology of Matter, Living Things and Persons." *Philosophical Studies* 60 (1990): 19-32. Print.
- Chappell, Vere. "Locke and Relative Identity." *History of Philosophy Quarterly* 6 (1989): 69–83. Print.
- Garrett, Don. "Locke on Personal Identity, Consciousness, and 'fatal Errors'." *Philosophical Topics* 31.1 (2003): 95-125. Print.
- Locke, John. *An Essay Concerning Human Understanding*. Ed. Peter H. Nidditch. Oxford: Oxford UP, 2011. Print.
- Newman, Lex. "Locke on Substance, Consciousness, and Personal Identity." MS. University of Utah.
- Reid, Thomas, Ronald E. Beanblossom, and Keith Lehrer. *Thomas Reid's Inquiry and Essays*. Indianapolis: Hackett Pub. 1983. Print.
- Strawson, Galen. *Locke on Personal Identity: Consciousness and Concernment*. Princeton, NJ: Princeton UP, 2011. Print.
- Stuart, Matthew. *Locke's Metaphysics*. Oxford: Oxford UP, 2013. Print.
- Winkler, Kenneth. "Locke on Personal Identity." *Journal of the History of Philosophy* 29.2 (1991): 201-26. Print.