



John Locke on Personal Identity: Memory, Consciousness and Concernment

Nargish Afroza 

Department of Communication, Philosophy and Politics, University of Beira Interior, Covilhã, Portugal
Email: nafroza.ma@gmail.com, d2464@ubi.pt

How to cite this paper: Afroza, N. (2023) John Locke on Personal Identity: Memory, Consciousness and Concernment. *Open Access Library Journal*, 10: e10949.
<https://doi.org/10.4236/oalib.1110949>

Received: November 1, 2023

Accepted: December 26, 2023

Published: December 29, 2023

Copyright © 2023 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

These questions come to our minds whenever we turn to the discussion of “Personal Identity”, “Are you the same person, you were a year ago”, or “Are you the same person now as we were working together last night”? “How do we persist over time” and “Is there a life after death?” Many philosophers have advanced diverse theories to try and answer questions like these. In 1690 famous empiricist John Locke’s famous work “An Essay Concerning Human Understanding” (Locke 1690) [1] presented a theory of personal identity which was the beginning of the modern discussion of these issues. According to Locke, the identity of a person is preserved with the identity of their consciousness, which means, one’s personal identity extends only so far as their own consciousness. Thus, he advocates that personal identity is a matter of psychological continuity and that it only “Depends on consciousness, not on substance”. More explicitly stated, a present person is identical to a past one only insofar as she or he remembers, or it is possible for her or him to remember herself or himself to thinking and acting in the past. But Locke’s theory has been scrutinized, debated, and rejected by his contemporaries and modern philosophers for many centuries, many of whom concluded that consciousness and memory is a necessary condition of personal identity and many of whom rejects the notion of memory to identify a person’s identity. Two historically significant objections were filed in the eighteenth century, one by Thomas Reid and the other by Joseph Butler. Both were criticized that if Locke’s memory theory of personal identity is possible then numerical identity is not possible. Thomas Reid in his work “Essays on the Intellectual Power of Man” (Reid 1785) [2] objects that Locke’s theory of personal identity lacks “Transitive Relation”. Also, Butler’s influential dissertation “Of Personal Identity,” appended to “The Analogy of Religion” in 1736 (Butler 1875) [3] objects that Locke’s theory of personal identity is “*Circular*” (added italics). So, besides their criticisms and objections, I want to rebuild in my paper the position of Locke’s theory of personal identity with “consciousness” and “con-

cernment”.

Subject Areas

Philosophy

Keywords

John Locke, Thomas Reid, Joseph Butler, Personal Identity, Self, Connected Consciousness, Memory, Psychological Continuity, Concernment

1. Introduction

In this paper, I am going to see how the 17th century English philosopher John Locke explains personal identity in his famous treatise “An Essay Concerning Human Understanding” (Locke 1690) [1]. One of the most important and controversial philosophical aspects of Locke’s treatment of personal identity is where his interpretation established itself as an important theory in modern philosophy. In recent philosophy, various commentators and scholars have declared that Locke’s theory of identity is a famously controversial doctrine.

John Locke clarified the problem of personal identity by specifying, how can a person who existed some time ago at $t + 1$ (time = t) continue to exist as the same person at time $t + 2$? He tries to specify the necessary and sufficient conditions for the re-identification of individuals in solving this problem; That is, it specifies the conditions under which a person can be said to be the same person at $t + 2$ as he was at some previous time $t + 1$. Locke characterizes our mode of personal identity as the experience accumulated in the mind by a psychological process, which makes a person now the same person he was then is the present consciousness of the ability to remember that past. For Locke, personal identity depends only on the acquisition of certain mental relations, which at the same time emphasizes that personal identity does not depend on any physical relations. The continuity of a person is nothing but continuity of mind (or soul) and continuity of body.

However, Locke’s criterion of personal identity in terms of memory and the connectedness of consciousness attracted considerable critical attention, of course, with two particular noteworthy arguments being advanced against it. The first argument associated with Joseph Butler finds that Locke’s account is *circular*, in as much as “consciousness of personal identity is presupposed and therefore cannot constitute personal identity” (Butler 1875, p. 358) [3]. Just exactly what Butler means here and just exactly how vulnerable Locke’s particular account may be to his charge remain disputed matters of historical interpretations. But one seemingly natural way of fleshing out an accusation of circularity against Locke involves the thought that, in so far as the very concept of memory relies upon a person at time $t + 2$ recalling the experience of that same person at some earlier time $t = 1$, it relies upon a notion of personal identity and so cannot

be used to define it. And the second argument associated with Thomas Reid finds that Locke's account is vulnerable to seeming disconnected between its implicit transitivity and the foible of memory. He seems to think that metaphysically speaking, memory is necessary but not sufficient for the creation of personal identity, which Locke acknowledged. For him (Reid), personal identity is simple and unanalyzable, because memory does not analyze one as the same person over time. Rather, memories allow immediate and direct knowledge of one's own past. However, I examine in my paper Reid's famous example of the "brave officer". Also in this context, the famous eighteenth-century skeptical philosopher David Hume's theory is also very famous because he clarified in his book "A Treatise of Human Nature" (Hume 1739) [4] that memory does not only express personal identity. Hume also finds Locke's argument is irrelevant, analyzing the theory that we must also admit necessary causes.

Suffice it to say that more or less satisfactory solutions I wrote in my paper have been proposed for the problems identified by Joseph Butler and Thomas Reid (see Noonan (Noonan 2019) [5], Galen Strawson (Strawson 2011) [6], Don Garrett (Garrett 2003) [7], Edmund Law (Law 2011) [8] ...), perhaps most notably in the latter case, in terms of psychological continuity, rather than simple connectedness of memory.

2. John Locke's Memory-Based Theory and Unity of Consciousness of Personal Identity

Locke's theory of personal identity is usually presented as a theory of mental succession "because of its emphasis on a psychological criterion" (Weinberg 2011, p. 398) [9]. In such a view, personal identity is determined and formed by memory. The recent philosopher Nimbalkar states in her abstract, "John Locke holds that personal identity is a matter of psychological continuity. He considered personal identity (or the self) to be founded on consciousness (viz. memory), and not on the substance of either the soul or the body." (Nimbalkar 2011, p. 268) [10]. Because Locke asserts that "...as far as this consciousness can be extended backward to any past Action or Thought, so far reaches the Identity of that Person; it is the same self now it was then; and 'tis by the same self with this present one that now reflects on it, that that Action was done" (Locke 1690, E2.27.9) [1]. In that case, for instance, if A remembers doing X, then A is the same person who did X. In the second view, personal identity depends on the continuity of consciousness. If B has an unbroken consciousness with A, then continuity is defined in different ways, the same as A. Of course, there are other ways to present Locke's theory as psychological. But Locke seems to have been taken as the ideal interpretation by these two contemporary philosophers Butler and Reid.

His argument is that consciousness is "inseparable from thinking", that when we perceive or deem or we know that we do it, and it is by this consciousness that we each consider ourselves, as an inseparable thought. It is through this reflected consciousness that our different sensations, perceptions, thoughts, and

desires belong to ourselves at any given time, and Locke thinks that the same principle is responsible for the identification of the self at different times.

However, in this context, it is important to note that John Locke is distinct between two kinds of identity, these are 1. Identity of organism and 2. Identity of mass of matter. Locke's scholar Lowe (Lowe 1995) [11] states about these distinctions on Locke's concept that:

“[an] example of a parcel of matter would be a lamp of gold or a piece of chalk. The general terms ‘gold’ and ‘chalk’ are known by linguists as mass terms, because they denote kinds of stuff rather than kinds of individual beings. However, given any such mass term it is possible to construct a corresponding sortal term with the aid of certain all-purpose nouns like ‘piece’, ‘lump’ and, indeed, ‘parcel’. Thus, we have to hand such (complex) sortal terms as ‘lump of gold’ and ‘piece of chalk’ which, like all genuine sortal terms, have both criteria of application and criteria of identity associated with their use” (Lowe 1995, p. 97) [11].

First, Locke is against physical organism continuity/identity of organism: Locke made it clear that he recognized two kinds of identity, one of things that change over time, such as animals and plants, and the other kind of identity of things that do not change over time, such as the mass of matter, so Locke analyzed in his book that the fixedness of a human organism or living thing can never be its identity.

Since Locke thinks that organisms survive, and yet it is undeniable that the mass of matter that forms at one time does not survive, Locke must look elsewhere for the survival of the organism. It is the life function of an organized body and is maintained by a gradually changing but properly disposed body of biological functions that an organism survives over time. Locke states that the same animal organism “is the same continued Life communicated to different particles of matter, as they happen successively to be united to that organized living Body” (Locke 1690, E2.27.8) [1]. So, all organisms and things exist based on diachronic identity because changes in animal organisms never specify synchronic identity, organisms always exist partly for one time and partly for another.

Second, living creatures do not depend on the mass of matter: Because it is not possible for living things to maintain the mass of matter. If the mass of matter is added or removed, it will lose its identity, so the mass of matter will always remain the same over time. That's why, Locke states, “...in the state of living creatures, their identity depends not on a mass of the same particles, but on something else. For in them the variations of great parcels of matter alter not the identity: an oak growing from a plant to a great tree, and then lopped, is still the same oak; and a colt grown up to a horse, sometimes fat, sometimes lean, is all the while the same horse: though, in both these cases, there may be a manifest change of the parts; so that truly they are not either of them the same masses of matter, though they be truly one of them the same oak, and the other the same horse, the reason whereof is, that, in these two cases—a mass of matter and a

living body-identity is not applied to the same thing” (Locke 1690, E2.27.3) [1].

Locke rejects the substance-based theory of personal identity, but it (personal identity) depends only on consciousness. He denies that to be the same person is to be or to have, one persisting immaterialism, spiritual soul-substance. Against this view, he does not deny that there are soul-substances but argues that their identity does not matter. If there is a soul-substance, then perhaps these could be reincarnated: the current president of America, as anyone knows, could be used as the soul of Socrates; But if Socrates has no consciousness from any of his actions or thoughts, no direct consciousness of those experiences like his, then he is not the same person as Socrates, again, if the same soul-substance carries two alternate sets. In the case of conscious thought, there will be two different persons with one soul (here I follow J. L Mackie’s example of Queensborough and Socrates (Mackie 1976, p. 174) [12]. Locke uses collateral reasoning to elucidate the inconsistency of the identity of both the living human body and the spiritual soul-substance. That’s why his theory of a person’s identity is only based on “consciousness”.

“Locke clearly has a ‘consciousness-based theory’ of personal identity and is sometimes credited with originating the ‘psychological continuity theory’ of personal identity. There is, however, an obvious sense in which Locke does not hold a *continuous* consciousness theory. For instance, as he states, consciousness is often ‘interrupted by forgetfulness.’ Also, Locke’s thought experiments in 2.27 (e.g., Dayman/Nightman, Socrates-waking/Socrates-sleeping) do not make any sense unless consciousness can be non-continuous or gappy. A better label for Locke’s theory is a ‘Connected Consciousness Theory’” (Kaufman 2016, p. 244) [13].

If consciousness can indeed be transferred from one soul to another, then a person can remain constant, despite the change in the soul to which his consciousness is attached i.e., consciousness remains the same, according to Locke, he remains the same person.

After all, Locke claims that even if a person has the same soul, he can fail to be the same person (see *Essay* 2.27.14, 23, and 24) [1]. In the passage “Day Man and Night-Man”, Locke asks the reader to imagine two distinct immutable consciousnesses operating in the same body, one continually by day, the other by night, so he suggests that day-man and night-man are individuals like Socrates and Plato (Locke 1690, E2.27.23) [1]. Locke then clarifies that this is the case even though day and night persons share the same soul.

“For granting that the thinking Substance in Man must be necessarily supposed immaterial, it’s evident, that immaterial thinking thing may sometimes part with its past consciousness, and be restored to it again, as appears in the forgetfulness Men often have of their past Actions, and the Mind many times recovers the memory of a past consciousness, which it had lost for twenty Years together. Make these intervals of Memory and

Forgetfulness take their turns regularly by Day and Night, and you have two Persons with the same immaterial Spirit, as much as in the former instance two Persons with the same Body. So that *self* is not determined by Identity or Diversity of Substance... but only by identity of consciousness” (Locke 1690, E2.27.23) [1].

Locke argues that these two bodies will contain the same person, just as you can have the same person in two different clothes; And this still holds if there are two distinct inert substances. As in spiritual metabolism, the same consciousness is transmitted from one soul-substance to another. Similarly, it can be transmitted from one body to another:

“For should the soul of a Prince, carrying with it the consciousness of the Prince’s past life, enter and inform the body of a Cobbler as soon as deserted by his own soul, everyone sees, he would be the same Person with the Prince, accountable only for the Prince’s actions: But who would say it was the same Man?” (Locke 1690, E2.27.15) [1].

In the event that the prince’s soul with all his royal thoughts transmigrates into the cobbler’s body, the cobbler’s soul is gone. But still the prince doesn’t think of himself as a cobbler, he still thinks of himself as a prince, only his body has changed. According to Locke, the difference between man and person makes it possible for the same person to appear in a different body in the resurrection and still be the same person. Locke says,

“I know that in the ordinary way of speaking, the same Person and the same Man stand for one and the same thing. And indeed, everyone will always have a liberty to speak, as he pleases, and to apply what articulate Sounds to what *Ideas* he thinks fit and change them as often as he pleases. But yet when we will enquire, what makes the same *Spirit, Man, or Person*, in our Minds; and having resolved with ourselves what we mean by them, it will not be hard to determine, in either of them, or the like, when it is the *same*, and when not” (Locke 1690, E2.27.15) [1].

That is, here it becomes clear that, for Locke, whether one is the same person or not depends on whether the person possesses the same body. So, people who have consciousness of past and present actions belong to the same person. In this case, Locke compared the discussion of personality to a forensic term.

Many commentators in these examples have created a lot of puzzles that have led to a lot of criticism of Locke’s theory.

1) The Jekyll-and-Hyde type the same body and the same soul are associated with two separate unities of consciousness, two mental histories not linked by mutual awareness.

2) In Socrates and the president of America we may have the same soul but different bodies and disjoint consciousnesses.

3) In the prince and the cobbler, we have the same soul and the same conscious-

ness successively in two different bodies.

And so on.

Locke claims that in all these different combinations it is reasonable to say that we are the same person *where and only where* (my italics) we have the same consciousness; The uniformity of the living body is not necessary or sufficient for the formation of the same person, nor is the uniformity of spiritual substance. However, now I am going to describe the position of Locke's personal identity.

3. Locke's Position on Personal Identity

In "Of Identity and Diversity" Locke clarifies the concept of personal identity, if two persons have the same thoughts and actions then they appear to be the same person but two men. Because both of them are displaying the same thought power in the society. But if neither of those two transfer their thinking power to the other, they will never be known as the same person because then their thinking power will not be the same.

With some fundamental points of identity, we have to know firstly what Locke means by the terminology of "person". Locke claims that persons are agents and "person" is a "...Forensic Term appropriating Actions and their Merit; and so belongs only to intelligent Agents capable of a Law, and Happiness and Misery" (Locke 1690, E2.27.26) [1].

So he points out that the person is not only a thinking intelligent being by which he considers his actions and results and considers himself as the object of the same thoughts in different places and times, but he/she has to take responsibility for the actions that he/she is taking. By which he will later suffer the consequences of the action may be punishment or reward.

Locke's readers tend to focus on the problem of personal identity over the course of time. It is tenable since Locke elaborates on how we claim our existence of personal identity.

"Personal identity: This being premised, to find wherein personal identity consists, we must consider what a person stands for which, I think, is a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places; which it does only by that consciousness which is inseparable from thinking..." (Locke 1690, E2.27.9) [1].

Locke has always tried to make it clear that we can never be constantly conscious of every act in our personal identity. Moreover, even though we constantly try to be conscious, forgetfulness creates a barrier to our consciousness. So he makes it clear in this case, we can remember those ideas and events that give us a clear awareness of those ideas. Therefore, we can only be aware of our personal identity based on conscious mental ideas and actions, and it is not possible to reveal personal identity. This is why Dan Kaufman's statement that Locke's

theory of personal identity is a “connected consciousness” (Kaufman 2016, p. 244) [13] theory applies.

But according to Locke, individuals cannot always clearly define their connected consciousness, i.e. their actions lack awareness. Hence, in considering such a theory, the individual declares the present and future individuals to be one and the same in order to clarify his existence. However, according to Locke, a present person will allow the possibility of being the same person in the future if he is able to be aware of present and past actions and future actions such as preconceptions. Locke cryptically states that in the context of interpretation,

“Consciousness makes personal identity. But it is further inquired whether it be the same identical substance. These few would think they had reason to doubt of, if these perceptions, with their consciousness, always remained present in the mind, whereby the same thinking thing would be always consciously present, and, as would be thought, evidently the same to itself” (Locke 1690, E2.27.10) [1].

According to this view, we can say that a person can only extend the consciousness of past actions in retaining his identity but not the future, which prevents him from retaining his future identity. So, in the account of *personhood* as constituted by reason, reflection, and conscious consideration of self at different times and places, Locke concludes that the temporal scope of a person is determined by the consciousness of the past,

“[s]ince consciousness always accompanies thinking, and ’tis that, that makes everyone 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 backward to any past Action or Thought, so far reaches the Identity of that Person; it is the same self now it was then; and ’tis by the same self with this present one that now reflects on it, that that Action was done” (Locke 1690, E2.27.9) [1].

Each person brings with his consciousness a special kind of “mental character” that is conscious of his own parts, thoughts, and actions. Locke calls this special kind of character “concernment”. Let’s illustrate the point with an example, in this case, I am currently mentally connected with every part of my body. For example, my hand is a part of the body that I am conscious of, *i.e.*, it is a source of sensation and can cause pleasure or pain. So, when I see that my left hand has come in contact with the fire, I think about what the result may be, immediately I remove the hand from contact with the fire because I am “concerned” about this. But if I see my friend David’s left hand in contact with the fire and it’s about to burn, I will have all kinds of mental characters such as emotional action-reactions. But since I will not feel any pain about what happens to David’s left hand, I am not concerned about the fate of David’s left hand in

Locke's sense. But if my left hand is amputated, I will no longer be conscious of that hand, *i.e.*, I will not have to think about that part of the body as more than a remote part of the object. If separation from my consciousness renders a part of the body that was once mine and no longer part of me, this is probably true of any part of me that would be mine but not something that I am currently conscious of (Kaufman 2016, p. 246) [13].

We can say that a person's consciousness is what carries anxiety in him and through it he can be mentally connected to his various activities. So that he can become familiar with his every action which may bring him happiness or sorrow as a result. Consciousness therefore extends mentally to every part of our body which determines our every action. We are also conscious of the work done by us and hence we suffer the results of that work which may be punishment or reward. So it is always recognized that the one who does the work is responsible for it and no one else. He states in that context that "[M]oral good and evil. Good and evil, as have been shown... nothing but pleasure or pain, or that which occasions or procures pleasure or pain to us" (Locke 1690, E2.27.5) [1].

Each of us is concerned with ourselves, that is, conscious. Locke described that we are "a thinking intelligent Being" (Locke 1690, E2.27.9) [1], then he seems that the person can be known as the same person only for the things that are thought, because each of our thoughts is related to its succession, so perceptions never have an independent and uninterrupted existence. Because we have an endless process of moving from one thought to another and from one thought to another. So we can keep our memory conscious of the similarity of our different perceptual thoughts and thus identify a person as the same person.

However, Locke defines a person as a "*forensic term*" (added italics) where he makes it clear that a person cannot be known as the same person with Nestor if that person is united with some of Nestor's physical substance or with Nestor's consciousness, so that person will be known as the same person with Nestor only when he is aware of any of Nestor's actions. Locke states,

"Would no more make him the same Person with *Nestor*, than if some of the Particles of Matter, that were once a part of *Nestor*, were now a part of this Man, the same immaterial Substance without the same consciousness, no more making the same Person by being united to any Body, than the same Particle of Matter without consciousness, united to any Body, makes the same Person. But let him once find himself conscious of any of the Actions of *Nestor*, he then finds himself the same Person with *Nestor*" (Locke 1690, E2.27.14) [1].

That is, in Locke's view, if any part "A" of a person without consciousness is joined to a part "B" of another person which was part of "A", then "B" cannot be called the same person as "A". But if "B" makes "B" conscious of any act of "A", then "A" and "B" are considered to be the same person. Another example of Locke's reasoning in this context is the case of Dayman and Nightman (Locke 1690, E2.27.23) [1], where he says that, although two persons possess the same

thoughts, because they do not have “one” and “clear” consciousness, there is no “concern” for the actions of either of them or for their part.

Thus, Locke refutes the “substance-based theory” through the above example with explanations. Because he prefers only connected consciousness and concernment to establish his theory of personal identity, for him, we cannot remember the actions (all past events) done by us at times. So, this connected consciousness (that is, when it notices the resemblance of particular events in the past to the present) is only capable when it is “concerned”. Since Locke associates personal identity closely with consciousness, he always has vision problems whenever there is a barrier to consciousness that does not consciously distinguish between individuals. Locke considers several such cases:

Amnesia:

“...suppose I wholly lose the memory of some parts of my Life, beyond a possibility of retrieving them, so that perhaps I shall never be conscious of them again; yet am I not the same Person, that did those Actions, had those Thoughts, that I was once conscious of, though I have now forgot them?” (Locke 1690, E2.27.20) [1].

Drunkenness:

“But is not a Man Drunk and Sober the same Person, why else is he punished for the Fact he commits when drunk, though he is never afterward conscious of it? Just as much the same Person, as a Man that walks, and does other things in his sleep, is the same Person, and is answerable for any mischief he shall do in it” (Locke 1690, E2.27.22) [1].

Forgetfulness:

“But that which seems to make the difficulty is this, that this consciousness, being interrupted always by forgetfulness, there being no moment of our Lives wherein we have the whole train of our past Actions before our Eyes in one view: But even the best memories losing sight of one part while they are viewing another... and in a sound sleep, having no Thoughts at all... I say, in all these cases, our consciousness being interrupted, and we are losing sight of our past selves, doubts are raised about whether we are the same thinking thing...” (Locke 1690, E2.27.10) [1].

Thus, an individual, as an ethical doer, he/she is responsible for all past or present actions performed by his/her, of which he/she is conscious. Also, the person has to be conscious of why he/she will be punished for the bad actions he/she has committed or if that consciousness is lacking the pain of punishment will be just pain, needless pain. So Locke quoted about our personal identity or self depends on consciousness, not on substance, he states,

“[s]elf is that conscious thinking thing, whatever substance made up of, (whether spiritual or material, simple or compounded, it matters not) —which is sensible or conscious of pleasure and pain, capable of happiness or misery,

and so is concerned for itself, as far as that consciousness extends... That with the consciousness of this present thinking thing can join itself, makes the same person, and is oneself with it, and with nothing else; and so attributes to itself, and owns all the actions of that thing, as its own, as far as that consciousness reaches, and no further; as everyone who reflects will perceive” (Locke 1690, E2.27.17) [1].

But Locke made it clear that his theory of personal identity was based on a moral and spiritual perspective in order to set it apart from all opposing theories. Where it is said that:

“Many have interpreted Locke to mean by ‘consciousness’ either having the same memories or having one or more mental states by which the one who is aware of those mental states as her own can identify herself as a previous person. Therefore, consciousness is thought to be a purely psychological criterion. Locke also seems to see his theory of personal identity as complementing his moral and theological views. ‘Person’, Locke says, ‘is a Forensic Term appropriating Actions and their Merit’, for which ‘at the Great Day’ we are judged by God and receive eternal reward or punishment (II. xxvii.26, I.iv.5)” (Weinberg 2011, p. 398) [9].

So, he explains the theory of personal identity by saying that God is the only continuously existent by which any objective thing is known. There are many events that it is not possible for a person to remember, in which case God is the agent of our actions, and determines the consequences, and only through God can we retrieve memories and determine our actions. Thus, Locke envisions God as a spiritual entity in describing personal identity where the individual determines rewards and punishments by analyzing God’s judgments for his/her actions.

Now I turn to analyse why Butler objects that Locke’s personal identity is merely *circular*.

4. Bishop Butler’s *Circularity* Objection against Locke’s Memory Theory of Personal Identity and Dismiss

The eighteenth centuries famous religious philosopher Bishop Joseph Butler argues against Locke’s memory-based theory of personal identity in his famous work “Analogy of Religion” (Butler 1875) [3], instead that Locke’s theory a merely *circular* (added italics).

Butler writes in criticism of Locke,

“And one should really think it self-evident, that consciousness of personal identity presupposes, and cannot, therefore, constitute, personal identity any more than knowledge, in any other case, can constitute truth, which it presupposes” (Butler 1875, p. 358) [3].

Although, in this quote, the recent scholar Galen Strawson (Strawson 2011)

[6] identifies that Joseph Butler made a “Wonderful Mistake”, not Locke, Butler incorrectly identifies consciousness with memory. According to Strawson, Butler misunderstands Locke’s words that he never explained consciousness as personal identity, he only referred to the mental act by which we identify our personal identity.

However, here I examine Butler’s allegation that Locke’s treatment of personal identity is *circular*. Butler does not accept Locke’s distinction between the concepts of person and thought-substance, therefore, any clear explanation given of the identity of a person concerns the identity of the thought-substance or soul. According to Butler, there are a number of reasons to demonstrate that there is no need for conscious thought substance to form a personal identity. Thus, if a person is only a thought substance, then Locke must be wrong that we can have the same person but a different substance, *i.e.*, according to Butler, consciousness cannot be required for personal identity. He states,

“But though consciousness of what is past does thus ascertain our personal identity to ourselves, yet to say, that it makes personal identity, or is necessary to our being the same persons is to say, that a person has not existed a single moment, nor done one action, but what he can remember; indeed, none but what he reflects upon” (Butler 1875, p. 358) [3].

He claims that in Locke’s argument, consciousness provides misleading evidence that it constitutes personal identity, or it is proof of the existence of the individual. But consciousness does not hold our existence of an identity, but it provides our source of knowledge and forms our collection of experiences. That’s why, Butler argued that consciousness makes a presupposed identity, it cannot provide the truth of the forming person’s existence. So the consciousness of presupposes identity produces circularity. It can be said that Locke’s conception of personal identity is just a fallacy of circularity. Memory only produces our future decisions, but it does not produce our identity. Then, he accused that Locke made a “*wonderful mistake*” (added italics). He states,

“This wonderful mistake may possibly have arisen from them; that to be endowed with consciousness, is inseparable from the idea of a person, or intelligent being. For, this might be expressed inaccurately thus, that consciousness makes personality: and from hence it might be concluded to make personal identity. But though present consciousness of what we at present do and feel is necessary to our being the persons we now are; yet present consciousness of past actions, or feelings, is not necessary to our being the same persons who performed those actions, or had those feeling” (Butler 1875, p. 358) [3].

According to Butler, Locke assumes the relation of consciousness to identity which is not correct. For to accept such an explanation, we would always be misinterpreting personal identity and accepting consciousness as individuality. Let’s also say that all the activities that we do now will be considered as our per-

sonal identity of the present, but we will not identify the actions of the past as our personality because the awareness of the present is much more than the past, so in this case, we consider the present more aware than the past. Moreover, if we accept Locke's theory, then we cannot call our actions and experiences our own, so in this case, we will call what we remember as our own, which is not correct, because the same "I" twenty years ago and the same "I" now will create a doubt in this case. Butler says in this regard,

"And as the two former comparisons not only give us the ideas of similitude and equality; but also show us, that two triangles are alike, and twice two and four are equal: so the latter comparison not only gives us the idea of personal identity, but also shows us the identity of ourselves in those two moments; the present, suppose, and that immediately past; or the present, and that a month, a year, or twenty years past. Or in other words, by reflecting upon that which is myself now, and that which was myself twenty years ago, I discern they are not two, but one and the same self" (Butler 1875, p. 358) [3].

He continues to state against Locke's memory theory,

"And suppose this being endued with limited powers of knowledge and memory, there is no more difficulty in conceiving it to have a power of knowing itself to be the same living being which it was some time ago, of remembering some of its actions, sufferings, and enjoyments, and forgetting others, than in conceiving it to know, or remember, or forget, anything else" (Butler 1875, p. 363) [3].

Since it is argued that I need to know that I was the same person in order to have a verifiable memory of past actions. So, memory cannot make me the same person. To answer this objection below some defenders of the Lockean position has introduced the concept of a more general kind of actual memory.

However, in addition to Butler, the eighteenth-century philosopher David Hume rejected memory theory in his analysis of Locke's theory of personal identity. As Hume pointed out in his famous work "A Treatise of Human Nature" (Hume 1739) [4], memory acts only as an auxiliary factor in the analysis of personal identity theories but can never be a necessary factor. So he mentions three essential principles in analyzing this theory which are contiguity, resemblance and causality. We provide our identity based on these principles, although Hume calls these principles an imaginary connection principle in our mind that we mistakenly claim to be the same over time.

But in Locke's defense, we can say that Hume's theory is never tenable. Although Hume explains the theory of personal identity by admitting necessary causes without memory, if we analyze Locke's theory, we can see that the necessary causes that Hume is talking about are also created by our consciousness and memories. Because we cannot fully explain necessary causes other than memory. So we can never exclude consciousness and memory in describing Locke's theory

of personal identity. In this argument, the famous philosopher Terence Penelhum's (Penelhum 1955) [14] argument is quite tenable against Hume's concept of personal identity. He states,

“To consider first the general problem with which Hume deals: the problem of personal identity can be roughly described as that of trying to justify a practice which seems at first sight to be strange, and even paradoxical. This is the practice of talking about people as single beings in spite of the fact that they are constantly changing, and over a period of time may have changed completely. It almost seems a contradiction to say that John Smith at two and John Smith at fifty-two are the same person, because they are so different” (Penelhum 1955, p. 571) [14].

The recent scholar Harold Noonan argues in his work “*Personal Identity*” (Noonan 2019) [5] on Butler's Circularity objection,

“This objection does not apply only to putative cases of reincarnation, where a present-day defender of the psychological-continuity criterion of personal identity might claim that their criterion is anyway not satisfied (*i.e.* that the latter person merely seems to but does not actually have, *genuine* quasi-memories of the earlier person's experiences). It applies also in cases which the defender of the psychological-continuity criterion of personal identity must regard as providing undeniable examples of personal identity” (Noonan 2019, p. 12) [5].

In fact, Butler misinterprets Locke's theory here, where Locke never made it clear that consciousness is the identity of the individual, in fact he made it clear that we can easily hold in memory all the actions we perform consciously, and in the future we interpret these conscious actions as personal identity. I can so to say, that if a person is aware of his own actions and thoughts which are his own identity *i.e.*, a thought or action related to A and B is done by a person, then A and B are considered to be the same person. Thus, it is fair to say in the context of this discussion that Butler does not seem to understand Locke's theory of personal identity very well.

In this context, many did not readily support Butler's form of circularity objective, which I referred to through the statements of scholars such as Harold Noonan and Strawson. Because Locke's theory of personal identity, they argue, depends on conscious action, it can only be disrupted by disordered memory or unconscious action, but is never criticized as vicious by circular guilt. But in this case, some questions remain, if that person forgets his old actions, who will be responsible for those actions? In this case he clarified the divine power and said that he makes it clear to us what is morally right and what is morally wrong. So if the person is somehow responsible for the action himself then he will suffer the consequences. And based on this theory, Locke makes it clear that God will punish or reward him for such actions. So in this case we can clearly understand that according to Locke if the person forgets what he has done then the supreme

divine power we have is always aware of all our actions and based on this he morally determines the results of a person's actions. Locke states that:

“Divine law the measure of sin and duty: First, the divine law, whereby that law which God has set to the actions of men, —whether promulgated to them by the light of nature, or the voice of revelation. That God has given a rule whereby men should govern themselves; I think there is nobody so brutish as to deny. He has a right to do it; we are his creatures: he has goodness and wisdom to direct our actions to that which is best: and he has the power to enforce it by rewards and punishments of infinite weight and duration in another life; for nobody can take us out of his hands. This is the only true touchstone of moral rectitude; and, by comparing them to this law, it is that men judge the most considerable moral good or evil of their actions; that is, whether, as duties or sins, they are like to procure them happiness or misery from the hands of the *Almighty*” (Locke 1690, E2.27.8) [1].

The recent scholar Edmund Law marked that personal identity understood in this Lockean way is relevant to both divine and human judgment:

“This distinct consciousness of our past actions, from whence arise all the ideas of merit and demerit, will most undoubtedly be regarded with the strictest exactness in *foro divino*; and indeed has its due weight in *foro humano*, whenever it can be with certainty determined: wherever this appears to be wanting, all judicial proceedings are at an end” (Law 2011, p. 237) [7].

Hence “*sorrow without demerit*” (my italics) to punish a person for an action. What he cannot be conscious of will not be considered reasonable. Therefore, God would not punish a person for an act that is not conscious of the action, so the person cannot simply ask for forgiveness. After all, no one should be expected to feel remorse for another person's actions and without being conscious of an action as another person's action. Locke's theory, in my logical view, is that the relation between persons and the responsibility for their actions is that there must be a “*conscious concernment*” (my italics) about that person's actions by which God is responsible to punish or reward that person properly.

Now I am turning to the objection of famous eighteenth-century philosopher Thomas Reid on Locke's concept of personal identity.

5. Thomas Reid's *Transitivity* Objection against Locke's Memory Theory of Personal Identity and Dismiss

The term transitive relation means $A = B = C$ e.g., if $A = B$ and $B = C$, then $A = C$, as Thomas Reid says, but Locke's theory of personal identity lacks *transitive relation*. In this case, the famous philosopher of the late 20th century Derek Parfit's argument is very notable in his famous book “Reasons and Persons”, where he refers to the psychological connection to explain “personal identity”, but this connection must be strongly connected with every experience of the person.

Such psychological connection and continuity is referred to as Relation-R. But this Relation-R does not actually relate to all of the person's experiences, so it creates barriers to the person's identity. Because I am psychologically strongly connected with the experience of a day or two ago, but I am not *strongly connected* with the events of twenty years ago. In this case Parfit refers to the experience of elders where he says that they remember the very old days and explain to others the events of the past done by him. But such *psychological connection* is weakly connected to their experiences, so it can never be Relation-R. In this case it would be a transitive relation where we refer to the past self as the present self even though we cannot remember everything. Parfit rejects Locke's argument to clarify such an analysis. He explained the *transitive relation* for the identity of persons,

“A relation *F* is *transitive* if it is true that, if *X* is F-related to *Y*, and *Y* is F-related to *Z*, *X* and *Z* must be F-related. Personal identity is a transitive relation. If Bertie was one and the same person as the philosopher Russell, and Russell was one and the same person as the author of *Why I Am Not a Christian*, this author and Bertie must be one and the same person... Strong connectedness is *not* a transitive relation” (Parfit 1984, p. 206) [15].

It can be said from Parfit's conceptions that:

“Because we can imagine that a future person has all kinds of experiences that can be explained in terms of the person's brain. But it does not claim that all these experiences and their interpretations are by one person alone. So personal identity is not important, in this case, ‘what matters’ is fundamentally important to the ‘Relation-R’ in any cause. Because in this case, ‘where one person is R-related to two other people’ (Parfit, 1984) [15]. But ‘Relation-R’ never provides personal identity” (Afroza 2023, p. 25) [16].

Perhaps the most famous objection to Locke's theory of personal identity is that it not only fails to defend transitivity but is positively committed to denying transitivity in many cases. But it is clear (to many) that Locke's theory of personal identity is not transitive as it is constituted by consciousness. Reid illustrates the point with an example of “brave officer”.

“Imagine a brave officer who was flogged as a boy for stealing, captured an enemy standard during his first campaign in middle age, and in old age is made a general. Suppose that at the time he captures the standard, he could remember being flogged and that as an elderly general he could no longer remember the flogging but could still remember capturing the standard. Locke's theory seems to have the implication that while the aged general is the same person as the middle-aged officer, and the middle-aged officer is the same person as the boy, the aged general is not the same person as the boy. But identity is what logicians call a transitive relation: if *A* is identical to *B* and *B* is identical to *C*, then *A* is identical to *C*. So, Locke's account

seems to be reduced to absurdity. As philosophers inclined toward a Lockean view of personal identity have pointed out, however, this objection can easily be surmounted by relaxing the requirement that there be direct links of memory and the like between each stage of a person's life. If the aged general can remember being the younger officer, and the younger officer can remember being the boy, then there is at least an indirect connection between the consciousness of the aged general and that of the boy that suffices for continuity of consciousness and, thereby, for personal identity. A revised Lockean theory could therefore hold that A is the same person as B just in case either there is a direct continuity of memories, personality traits, etc., between A and B or, if there is no such direct continuity, then A and B are indirectly linked by intermediate stages that are directly continuous" (Feser 2007, pp. 68-69) [17].

In the above example, three times are mentioned, suppose, t1, t2, and t3. A brave officer in t2 realizes that he has been flogged as a young boy in t1. And in t3, an old general is conscious of accepting standards as a brave officer in t2. However, the old general of t3 is not conscious (and cannot be) of flogging in t1. According to Reid, Locke's theory implies that the young boy in t1 is the same person as the brave officer in t2 and the brave officer in t2 is the same person as the old general in t3. However, since no consciousness extends from t3 to flog at t1, the Old General is not the same person as the young boy. Thus, Locke's theory of personal identity violates transitivity, which is opposed to traditional personal identity theory.

Thomas Reid criticized Locke's theory of personal identity in order to establish a metaphysical hypothesis, known as memory theory, from the conceptual connection between memory and personal identity. According to Locke's theory, memory constitutes personal identity, because this memory is necessary and sufficient for the union of individuals, i.e., memory alone claims to be the same person again. In fact, Reid holds that it is impossible to account for personal identity in terms other than the self. Memory is only a first-person proof of personal identity, as I know I attended my graduation reunion because I can remember being there. According to Reid, memory not only recalls previously experienced events but in memory, the mind is not directed to previously experienced concepts or concepts of prior experience.

So, for this critique, I mention that, it is worthy Strawson (Strawson 2011) [6] clarified how to respond to the Reid's criticism above referred to by correctly pointing out that "an illustration of [the theory's] fundamental and forensic point, the commonsense point (it's commonsense relative to the story of the Day of Judgment) that human beings [*sic*] won't on the Day of Judgment be responsible for all the things they have done in their lives, but only for those that they're still Conscious of and so still Concerned in" (Strawson 2011, p. 59) [6]. Strawson says that "consciousness" constituting the core of Locke's theory is not the same as mere memory but something always accompanying "concernment". Strawson

makes it quite clear that Locke's conception of consciousness is not like memory, indicating that "consciousness", which Locke is an event in the present moment (hence, it is separate from memory), and not every memory necessarily accompanies "concern". The inconsistent allegations made by Reid on Locke's theory, will be refuted by Strawson's argument, since the Reid critique clearly interprets "consciousness" as "memory". Rather, personality "is differently constituted every day, on Locke's" (Strawson 2011, p. 55) [6], thus the situation presented by Reid's critique is consistent with (rather without contradicting) the meaning of Locke's "person", as opposed to what the Reid's criticism by intended.

An influential attempt to save Locke is found in a classic paper by Don Garrett (Garrett 2003) [7]. According to Don Garrett, if a person is conscious of his/her actions or past experience then he/she is obviously responsible for his/her actions. He says,

"Locke actually holds this general doctrine—*i.e.*, the doctrine that conscious memory of an earlier perception or action also extends the history of the person to whatever other perceptions or actions are implicated in the sameness of consciousness with that earlier perception or action, regardless of present ability to remember them" (Garrett 2003, p. 21) [7].

So, Don Garrett (Garrett 2003) [7] thinks that, as we have noted, according to Locke, the experience of action-responsibility involved in remembering one's own actions is irrefutable or indubitable. But since Locke holds that true memory is "*only one kind of representation of the past*" (added italics), representations of the false transference of consciousness through false phenomenal memory can be expected to have the same overwhelmingly "*phenomenological character*" (my italics). Locke holds that only those actions and thoughts are mine, which is my own consciousness and only I can remember, e.g., identity is inextricably linked to connected consciousness and memory. That's why Garrett states,

"And that it never is so, will by us, till we have clearer views of the Nature of thinking Substances, be best resolved into the Goodness of God, who as far as the Happiness or Misery of any of his sensible Creatures is concerned in it, will not be a fatal Error of theirs transfer from one to another, that consciousness, which draws Reward or Punishment with it" (Garrett 2003, p. 5) [7].

In that case, Udo Thiel states about Garret's comments,

"Garrett argues that there are passages in the Essay that 'strongly suggest' that Locke in fact holds such a view. This is evident, he says, from Locke's 'pointed limitation of the two cases he discusses in which personal identity fails despite the identity of thinking substance or man'" (Thiel 2011, p. 219) [18].

That's why Thiel clear that,

"Locke implies that personal identity reaches only as far as it can 'be ex-

tended' or can 'reach' by consciousness (e.g., Essay II.xxvii.9-10, 14, 23), he does not say that the extension by consciousness of a present person into the past is always limited to what the present person now remembers or even can now actually remember" (Thiel 2011, p. 219) [18].

Also, Edmund Law says in his famous article "Locke on Personal Identity: Consciousness and Concernment" to defend Locke's memory-based theory of personal identity. He states,

"Nor does it properly lie in a *power of causing* a return of the same idea; but rather in the *capacity of receiving* it, of readmitting the same consciousness concerning any past thought, action, or perception. Nor is it merely a *present representation* of any such act; but a representation of it as *our own*, which entitles us to it; one person may know or become conscious of the *deeds of another*, but this is not knowing that *he himself was the author* of those deeds, which is a contradiction; and to treat him as such upon that account only, would be inverting all rules of right and wrong; and could not therefore be practised by either God or Man, since no end could possibly be answered by such treatment, as observed above" (Law 2011, p. 248) [8].

Another point worth noting in defending Locke's idea is that if I accept Parfit's "Relation-R" then the moral of the individual cannot be considered because. "[A]s Parfit's example of teletransportation is quite relevant here when he explains, teletransportation is a vehicle that can create an exact replica of me that is mentally continuous with me. But if I make a commitment to someone in this case, will my counterpart be similarly committed? To whom shall the one with whom I am committed be particularly liable, to me or to my counterpart? I think Parfit's 'Relation-R' will struggle to answer many such questions" (Afroza 2023, p. 27). But Locke's account of the theory of personal identity does not raise any problems with such an ethical analysis.

Thus, a recurring criticism of Locke's account is that memory and consciousness are neither necessary nor sufficient for personal identity. However, this criticism is generally based on incorrect assumptions about Locke's theory. Locke maintains, however, that personal identity is not at the same time an object of consciousness formed by the latter, for the consciousness of sameness never constitutes sameness. Rather, the consciousness of thought and action constitutes its identity with the individual and with time, distinguishing the identity of the individual from that of man and soul.

Now I proceed to the conclusion of this discussion.

6. Conclusions

So, we can conclude that in the light of Butler and Reid's critique, it is safe to say that John Locke never equated consciousness and memory, always referring to "extended consciousness" as "concernment" and never saying that memory connections constitute personal identity. In this case, Strawson's book "Lock on

Personal Identity” (Strawson 2011) [6] simply came to our notice then. Strawson argues that the main error is that Locke uses the word “person” only in a general way, only as a word for an ideal permanent thing, such as “man” in reality, Locke uses “person” primarily as a “forensic” or legal term specifically designed for questions about praise and guilt, punishment, and reward. In these terms, your personal identity is largely a matter of your past activities for which you are still responsible because you are still “concerned” about them in the special sense of the word of Lock. So, I agree with Strawson that the forensic nature of the person and their divine judgment (“charged by their own consciousness”) is crucial to Locke’s personal identity theory. That is, Locke’s theory of personal identity correctly deals with “*forensic-consciousness-concernment-responsibility-divine judgment-and the allocation of punishment/reward*” (my italics), but it does so at the expense of an admirable theory of a person’s numerical identity over time.

Thus, Locke’s account of personal identity is embedded in a general account of identity. In this context recent scholar says,

“Personal identity for Locke is psychological continuity. But his theory is criticized by both Butler and Reid as a ‘wonderful mistake’ or ‘reduced to absurdity’. However, Locke’s theory has a profound influence in the field of education and the development of psychology” (Nimbalkar 2011, p. 274) [10].

As E.J. Lowe (Lowe 1995) [11] remarks in that context, “Locke seems to have been the first philosopher to address the problem of personal identity in anything like its modern form—indeed, it was he who was responsible for setting the terms of the modern debate, and his views on the issue remain highly influential” (Lowe 1995, p. 102) [11].

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] Locke, J. (1690) An Essay Concerning Human Understanding. The Pennsylvania State University Press, University Park.
http://www.philotextes.info/spip/IMG/pdf/essay_concerning_human_understanding.pdf
<https://doi.org/10.1093/oseo/instance.00018020>
- [2] Reid, T. (1785) Essays on the Intellectual Powers of Man. Cambridge University Press, Cambridge. <https://doi.org/10.1093/oseo/instance.00106533>
- [3] Butler, J. (1875) The Analogy of Religion, Natural and Revealed to the Constitution and Course of Nature. Nelson & Phillips, New York.
- [4] Hume, D. (1739) A Treatise of Human Nature. 2nd Edition, Oxford University Press, Oxford. <https://doi.org/10.1093/oseo/instance.00046221>
- [5] Noonan, H.W. (2019) Personal Identity. 3rd Edition, Routledge Taylor & Francis Group, London.
- [6] Strawson, G. (2011) Locke on Personal Identity: Consciousness and Concernment.

- Princeton University Press, Princeton.
- [7] Garrett, D. (2003) Locke on Personal Identity, Consciousness, and “Fatal Errors”. *Philosophical Topics*, **31**, 95-125. <https://doi.org/10.5840/philtopics2003311/214>
 - [8] Law, E. (2011) Appendix Two. A Defence of Mr. Locke’s Opinion Concerning Personal Identity. In: Strawson, G., Ed., *Locke on Personal Identity: Consciousness and Concernment*, Princeton University Press, Princeton, 233-252. <https://doi.org/10.1515/9781400851843-024>
 - [9] Weinberg, S. (2011) Locke on Personal Identity. *Philosophy Compass*, **6**, 398-407. <https://compass.onlinelibrary.wiley.com/doi/10.1111/j.1747-9991.2011.00402.x> <https://doi.org/10.1111/j.1747-9991.2011.00402.x>
 - [10] Nimbalkar, N. (2011) John Locke on Personal Identity. *Men’s Sana Monographs*, **9**, 268-275. <https://doi.org/10.4103/0973-1229.77443>
 - [11] Lowe, E.J. (1995) *Locke: On the Human Understanding*. Routledge, London.
 - [12] Mackie, J.L. (1976) *Philosophical Studies Problems from Locke*. Clarendon Press, Oxford. <https://doi.org/10.1093/0198750366.001.0001>
 - [13] Stuart, M. (2016) *A Companion to Locke*. John Wiley & Sons Ltd., Chichester. https://gkleung.com/wp-content/uploads/2021/08/A-Companion-to-Locke_compressed.pdf
 - [14] Penelhum, T. (1955) Hume on Personal Identity. *The Philosophical Review*, **64**, 571-589. <http://www.jstor.org/stable/2182635> <https://doi.org/10.2307/2182635>
 - [15] Parfit, D. (1984) *Reasons and Persons*. Clarendon Press, Oxford. <https://www.stafforini.com/docs/Parfit%20-%20Reasons%20and%20persons.pdf>
 - [16] Afroza, N. (2023) Derek Parfit on Personal Identity: Relation-R and Moral Commitments. *International Journal of Science Engineering and Technology*, **11**, 20-28. <https://doi.org/10.61463/ijset.icsshrt-2023.102>
 - [17] Feser, E. (2007) *Locke. Oneworld Thinkers*, Oxford.
 - [18] Thiel, U. (2011) *The Early Modern Subject: Self-Consciousness and Personal Identity from Descartes to Hume*. Oxford University Press, Oxford. <https://doi.org/10.1093/acprof:oso/9780199542499.001.0001>