

Portland State University

PDXScholar

Philosophy Faculty Publications and
Presentations

Philosophy

2008

Locke on Consciousness

Angela M. Coventry

Portland State University, coventry@pdx.edu

Uriah Kriegel

Rice University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/phl_fac



Part of the [Philosophy Commons](#)

Let us know how access to this document benefits you.

Citation Details

Coventry, A., & Kriegel, U. (2008). Locke on consciousness. *History of Philosophy Quarterly*, 25(3), 221-242.

This Post-Print is brought to you for free and open access. It has been accepted for inclusion in Philosophy Faculty Publications and Presentations by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

LOCKE ON CONSCIOUSNESS

Angela Coventry & Uriah Kriegel

Abstract. Locke's theory of consciousness is often appropriated as a forerunner of present-day Higher-Order Perception (HOP) theories, but not much is said about it beyond that. We offer an interpretation of Locke's account of consciousness that portrays it as crucially different from current-day HOP theory, both in detail and in spirit. In this paper, it is argued that there are good historical and philosophical reasons to attribute to Locke the view not that conscious states are accompanied by higher-order perceptions, but rather that conscious states constitute perceptions of themselves.

Many aspects of John Locke's philosophy of mind have received considerable attention in the past three centuries. His theory of personal identity in terms of memory; his alleged indirect realism about perception; the distinction between primary and secondary qualities; the attack on nativism; his psychologism about meaning; his empiricist views on concept acquisition, etc., have all been subject to extensive scrutiny and debate. But Locke's theory of consciousness stands out as remarkably unattended to among his contributions to the philosophy of mind.

There are two main reasons for this neglect. One is that consciousness itself has been out of favor with philosophers through most of the twentieth century, or at least the first three quarters thereof, dominated as they had been by behaviorist and functionalist ideas.¹ The second reason for the little interest in Locke's theory, however, lies with Locke himself. His remarks on consciousness are scattered and relatively scarce. There is no systematic discussion of the

phenomenon. Thus, there is no mention of the word “consciousness” or any of its cognates in any of the 69 chapter titles, and only a handful within the 1122 section titles, of the *Essay Concerning Human Understanding*.²

When it is mentioned in current-day philosophy, Locke’s theory of consciousness is often appropriated as a forerunner of Higher-Order Perception (HOP) theories, but not much is said about it beyond that. In this paper, we offer an alternative interpretation of Locke’s account of consciousness that portrays it as crucially different from current-day HOP theory, both in detail and in spirit.

1. Current-Day Higher-Order Perception Accounts of Consciousness

One promising approach to consciousness construes it as a monitoring device: the faculty of consciousness is a mental scanner directed at the inner goings-on in the subject. Both David Armstrong and William Lycan have defended versions of this view. In the course of arguing for a materialist theory of the mind, Armstrong defines consciousness as an internal self-scanning device or mechanism in the central nervous system, so that consciousness of our own mental state becomes simply the scanning of one part of our central nervous system by another part. It is part of the view that the scanning state and the state scanned must be distinct mental states; thus he writes that “it is impossible that the introspecting and the thing introspected should be one and the same mental state.”³ According to Lycan, the mental scanning constitutive of consciousness is carried out by means of attention mechanisms:

[C]onsciousness is the functioning of internal attention mechanisms directed upon lower-order psychological states and events... Attention mechanisms are devices that have the job of relaying and/or coordinating information about ongoing psychological events and processes.⁴

On this view, consciousness involves the capacity to form second-order mental states, i.e., mental states whose objects are other mental states. This thesis is often called the Higher-Order Monitoring (henceforth, HOM) theory.

A mental state is conscious, on this approach, just in case the subject is aware of it through a higher-order state. That is, a subject's state S_1 is conscious iff she entertains also a second-order state S_2 , such that S_1 is the object of S_2 .⁵ For instance, a perceptual experience of a cube is conscious when, and only when, it is appropriately accompanied by a second-order state representing it.

To be sure, not every mental state we represent ourselves as having is thereby conscious. Thus, if x is told by her therapist that she is subconsciously angry at her father for the latter's childhood mistreatment of x , and x consequently comes to believe that she does hold such anger towards her father, it does not immediately follow that x 's anger is now conscious. This is why HOM theorists suggest that S_2 must represent S_1 "in the appropriate manner." What the appropriate manner is should not overly concern us here.⁶ One important requirement, though, is that, normally, S_2 must be itself an unconscious mental state.⁷ We will have occasion to discuss the motivation for this requirement momentarily.

On this model of consciousness, then, consciousness is an extrinsic property of S_1 : there is nothing in S_1 itself that makes it conscious. What makes it conscious is an altogether different state, S_2 . So S_1 is conscious in virtue of its relational property of being appropriately represented by a separate state.

What sort of faculty is the monitoring faculty that constitutes consciousness? Here there is an important division between those who take it to produce higher-order judgments and those who take it to produce higher-order perception-like states.⁸ The former view is known as the

Higher-Order Thought theory and the latter as the Higher-Order Perception theory. Intuitively, a mental state is perception-like if it is produced by a faculty whose operation is analogous in important respects to the operation of the faculties of vision, hearing, etc. The central feature of these faculties is that they are grounded in the operation of the senses. If the monitoring faculty is similarly grounded in the operation of some sort of inner sense, then it is more perceptual than judgmental.

Both Armstrong and Lycan take Locke to be a forerunner of the sort of Higher-Order Perception (henceforth, HOP) theory they wish to defend. Armstrong claims that his proposal that consciousness is nothing but the “awareness of the state of our own mind” is to be understood in the “way that Locke ... conceived it, as like perception” (1980: 198-9). Lycan (Ibid) is even more explicit:

Locke put forward the theory of consciousness as “internal Sense” or “reflection”... On that theory, consciousness is a perception-like second-order representing of our own psychological states... Locke’s idea has been urged in our own time by philosophers such as D. M. Armstrong and by psychologists such as Bernard Baars; I have previously defended it too.⁹

Contemporary critics of the Higher-Order Perception theory likewise attribute it to Locke. Thus, Guven Guzeldere writes:

There is a strong intuition, which dates back several centuries, that consciousness is not, or does not consist in, an ordinary mental state or process itself, but it is, or it consists in, the awareness of such states and processes. Locke...epitomized this intuition in his celebrated statement:

“consciousness is the perception of what passes in a man’s own mind”...¹⁰

This interpretation may appear to carry some initial plausibility. Consider Locke’s explanation in the Essay Concerning Human Understanding of the two fountains of knowledge from which all ideas originate: sensation and reflection. The former are sensorily based impressions of external objects, which give rise to ideas such as of sweet or bitter, blue or red, hot or cold, hard or soft,

etc. The latter are impressions derived from reflection, which give rise to ideas of the internal operations of our own minds, such as thinking, doubting, reasoning, believing, and willing (ECHU 2.1.2-4). When describing how exactly we are aware of the operations of our minds, Locke famously upholds the existence of a quasi-perceptual inner sense:

The other Fountain, from which Experience furnisheth the Understanding with Ideas, is the Perception of the Operations of our own Minds within us... And though it be not Sense, as having nothing to do with external Objects; yet it is very like it, and might properly enough be call'd internal Sense (ECHU 2.1.4).

Moreover, Locke holds that consciousness is the employment of this inner sense – most clearly in the statement cited by Guzeldere, from the Essay 2.1.19, that “Consciousness is the perception of what passes in a Man’s own mind.”

It is understandable, against this background, that present-day HOP theorists should consider Locke a forerunner. But as we shall now argue, this is in fact dubious.

2. Higher-Order Perception: The Problem of Infinite Regress

In this section, we will argue that HOP theory cannot be held in conjunction with the thesis, to which Locke was firmly committed, that all mental states are conscious. If so, the only way Locke could maintain the HOP theory is if he was unaware that it cannot be held in conjunction with this thesis. But in the next section, we will argue that there are historical reasons to doubt that.

Locke says explicitly that mental states are necessarily conscious in numerous passages. He first makes this point when attacking René Descartes’s claim that the mind is always thinking, even in dreamless sleep. Descartes held that thought is the essence or principal attribute of minds, just as extension is the principal attribute of bodies. Now since a substance cannot be without its principal attribute, a mind without thought would be impossible. Locke grants that the

mind in a waking person is never without thought but insists that this doctrine will not convince those who claim to sleep without dreaming and that “every drowsy Nod shakes [the] Doctrine ... That the Soul is always thinking”; only God thinks constantly and never falls asleep (ECHU 2.1.10-3). The relevant point is that during the course of the argument, Locke states that “Whilst [the soul] thinks and perceives... it must necessarily be conscious of its own Perceptions” (ECHU 2.1.12) and that “[T]o be happy or miserable without being conscious of it, seems to me utterly inconsistent and impossible” (ECHU 2.1.11).

Many chapters later, in the context of defending his famous doctrine of personal identity, Locke states that the existence of unconscious mental states is actually impossible. When getting clear on the definition of person as a thinking, intelligent being, he writes that “It [is] impossible for any one to perceive, without perceiving, that he does perceive” (ECHU 2.27.9). Note well the modal force of these claims: “necessarily” and “impossible” appear in each quotation. Locke evidently thinks it absurd, almost unintelligible, that a mental state should exist without being conscious. This was not, of course, an unusual view before Freud. For most pre-Freudian thinkers, it was deemed a conceptual truth that all mental states are conscious in the sense that one is conscious of them, or is aware of being in them.

This view cannot be held together with HOP theory, however. Recall that according to HOP theory, a mental state S_1 is conscious just in case its occurrence is represented by a second-order quasi-perceptual mental state S_2 . But if all mental states are conscious, then S_2 would have to be a conscious state as well. The question then arises, however, of what it is that makes S_2 conscious. According to HOP theory, a mental state is conscious in virtue of being represented by a higher-order quasi-perceptual state. If so, S_2 would have to be conscious in virtue of being represented by a (quasi-perceptual) third-order state, S_3 . And of course, S_3 would also be

conscious, since all mental states are. Therefore, S_3 would also require the existence of a fourth mental state, S_4 , which represents it – and so on ad infinitum. In short, the conjunction of the HOP theory of consciousness and the thesis that all mental states are conscious leads to an infinite regress of higher-order states. This is a reductio of the conjunction, in two ways: first, the emerging picture fails to account for consciousness in terms of something other than consciousness; second, it is empirically implausible – and probably impossible – that we should harbor an infinity of mental states whenever we are conscious.

Since the conjunction leads to an infinite regress, one of the conjuncts must be rejected. Current-day HOP theorists reject the thesis that all mental states are conscious.¹¹ This is why they hold that S_2 is normally an unconscious mental state, as we noted in §1. This feature of their theory has brought upon them much criticism: how can it be that S_1 inherits its consciousness from something unconscious? How can consciousness be grounded in the unconscious? And how can our awareness of our conscious states be normally unconscious, when in fact we normally have first-person acquaintance with it?¹² These are serious problems, yet surely they are preferable to an infinite regress.

Denying that S_2 is (normally) conscious is not the only way to avoid the infinite regress, however. Consider how Franz Brentano dealt with the same infinite regress two and a half centuries after Locke. After discussing just this regress problem, Brentano concludes that our awareness of our conscious states cannot be grounded in a separate mental state, but must be instead grounded in these very states:

[Every conscious act] includes within it a consciousness of itself. Therefore, every [conscious] act, no matter how simple, has a double object, a primary and a secondary object. The simplest act, for example the act of hearing, has as its primary object the sound, and for its secondary object, itself, the mental phenomenon in which the sound is heard.¹³

For Brentano, S_1 is conscious in virtue of representing its very own occurrence. A significant difference between the HOP model of consciousness and a model along Brentano's lines is that the latter, unlike the former, construes consciousness as intrinsic to conscious states. Recall that according to the HOP model, a conscious state is conscious in virtue of its relational properties. But according to the Brentanian model, it is conscious in virtue of its own intentional properties, in particular, the property of being intentionally directed at itself. We may call a Brentano-style account of consciousness a Same-Order Perception account (henceforth, SOP).¹⁴

3. Was Locke a Same-Order Perception Theorist of Consciousness?

The question we should ask ourselves at this stage is whether there is any evidence whatsoever to suggest that Locke held a HOP account of consciousness rather than a Brentano-style SOP account. Statements such as “consciousness is the perception of what passes in a man's own mind” do not distinguish the two views. A person may perceive what passes in his or her own mind either (i) by virtue of entering a new and separate mental state, which would count as a higher-order perception, or (ii) by virtue of having a self-perceiving (so to speak) first-order mental state. Therefore, the textual evidence commonly adduced to attribute to Locke a HOP account of consciousness is insufficient.¹⁵ Nor do parsimony considerations: although SOP posits a complex and quite special intentional structure involving self-perception, it economizes on the number of mental states posited.¹⁶

In the remainder of this paper, we will make a case for interpreting Locke as holding a SOP rather than HOP view. We divide our case into three parts: historical, textual, and “theoretical.” In §4, we offer historical evidence to suggest that Locke is unlikely to have been unaware of the infinite regress that attends the combination of HOP and the thesis that all mental states are conscious. That should be taken as indirect evidence for the claim that Locke held

SOP. (As we noted above, it seems to us that there is little textual evidence that bears on the issue. The fact that the HOP theory of consciousness cannot be consistently held in conjunction with the view that all mental states are conscious does not by itself mean that Locke did not hold it in conjunction with that view. Locke would have refrained from holding this conjunction only if he was aware of the inconsistency of holding it. What we want to argue in that section is that there are good historical reasons to think that he could not have been unaware of it.) In §5, we will argue that what little textual evidence there is that bears on the issue actually suggests that Locke probably held SOP. In §6, we will argue that attributing to Locke a SOP theory of consciousness has certain theoretical payoffs in terms of illuminating other parts of his philosophy of mind. In particular, our understanding of his notion of idea, his account of perception and reality, and some of his views on self-knowledge, are all altered in interesting ways in the light of the SOP interpretation. The alterations are not only interesting, however, but also appear to render Locke's views more plausible and more clearly consistent. On this part of the case for interpreting Locke as a SOP theorist a virtually indefinite amount can be written. Here we will limit ourselves to relatively superficial indications of how we think the different stories might go.

At the end of the day, however, the main reason to attribute to Locke the SOP account is simply the fact that the HOP account is untenable in conjunction with Locke's view that all mental states are conscious. The principle of charity prescribes that we attribute to Locke the account that does not lead straightforwardly to infinite regress.¹⁷

It might be objected that a SOP interpretation could not account for Locke's central postulation of the inner sense. If Lockean mental states were self-perceiving, there would be no need for an internal sense to perceive them. In addressing this objection, it is important to keep in

mind the common distinction between consciousness and reflection (or as we would rather put it, between ordinary consciousness and reflective consciousness). The latter requires attentive awareness of one's concurrent mental states, the sort characteristic of deliberate introspective reflection through which the subject scrutinizes her internal goings-on in an explicit way. The former requires only a less imposing, not necessarily attentive, phenomenologically "lighter" awareness of one's concurrent mental states. This is a distinction Locke was certainly making.¹⁸ We are stressing this distinction because even if the postulation of inner sense plays no role in the explanation of (ordinary) consciousness, it may still be motivated by the need to explain reflection (reflective consciousness). The latter requires the mind's ability to transform inattentive awareness into attentive awareness. It therefore requires the operation of a mechanism that can turn attention resources inwards – from external objects to our internal representations of those objects. The inner sense is probably postulated by Locke to fit that explanatory bill. If so, Locke's postulation of the inner sense does not pose a special difficulty for the SOP interpretation.¹⁹

Another possible objection is that it is a false dichotomy to suppose that Locke must have been either a HOP or a SOP theorist. He may well have thought that some conscious episodes involve higher-order perception while others involve same-order perception. However, note that the SOP account does not claim that higher-order perception never occurs. It merely claims that such higher-order perception is not the mark of consciousness. To be sure, there are episodes in which a higher-order mental act that targets a lower-order one in the standard way do occur. These are the episodes that characterize the more reflective moments of our conscious life, the moments in which we deliberately and voluntarily turn our attention inward. But even in those episodes, the presence of higher-order perception simply cohabits with a same-order perception.

For while the second-order state is a higher-order quasi-perceptual representation of the first-order state, it is also at the same time a same-order quasi-perceptual representation of itself. So higher-order perception can certainly occur in conscious episodes, but it does not capture consciousness per se, whereas by contrast, same-order perception (or “self-perception”) is necessarily implicated in all conscious episodes and does capture consciousness per se. At least this is what the SOP account maintains. The fact that there are conscious episodes involving higher-order perception thus does not threaten the generality of the SOP account of consciousness.

4. Historical Evidence: Aristotle and Leibniz on the Infinite Regress Problem

We noted that, at the end of the day, the main reason to ascribe to Locke a SOP account of consciousness is the application of the principle of charity. But there is more than charity to suggest that Locke did not hold an account that led straightforwardly to an infinite regress. There are historical reasons to suspect that Locke could not have been unaware of the looming infinite regress. For the problem of infinite regress was well known at least since Aristotle, and is very much on Gottfried Leibniz’s agenda. That is, the problem of accounting for consciousness without falling into infinite regress was discussed both before Locke’s time, by one of the most influential pre-modern philosophers, and during it, by one of the most influential contemporaries of Locke.

The first discussion of the problem is to be found in Aristotle’s *De Anima*. According to Victor Caston, Aristotle held a SOP account of consciousness.²⁰ Moreover, Aristotle’s reason for holding the SOP account is precisely the regress problem. In the first paragraph of *De Anima* 3.2, Aristotle writes:²¹

Since we perceive that we see and hear, it is necessarily either by means of seeing that one perceives that one sees or by another perception. But...if the perception of vision is a different

perception, either this will proceed to infinity or some perception will be of itself; so that we ought to posit this in the first place.²²

Aristotle's conclusion is stated most clearly in the *Metaphysics* 12.9 (1074b35-6): “[conscious] knowing, perceiving, believing, and thinking are always of something else, but of themselves on the side.”

Interestingly, Aristotle's view that conscious states represent themselves is defended by a number of early modern philosophers, both on the European continent and in the British Isles. At least, this is how Genevieve Rodis-Lewis and Kathleen Wider interpret Descartes; how Steven Nadler interprets Arnauld; how Andrew Brook, Dieter Sturma, and Rocco Gennaro interpret Kant; how Udo Thiel interprets Cudworth; and how Keith Lehrer, Keith Hossack, and J. T. Buras interpret Reid.²³

Even more interestingly, Leibniz, whom we know had many disagreements with Locke, also discussed our regress problem. Unlike Aristotle – who dealt with the regress by denying HOP theory's assumption that a conscious state is represented by a separate state – Leibniz deals with it by rejecting the view that all mental states are conscious. Consider this passage in *New Essays on Human Understanding*, his response (incidentally) to Locke's *Essay Concerning Human Understanding*:

it is impossible that we should always reflect explicitly on all our thoughts; [otherwise] the mind would reflect on each reflection ad infinitum, without ever being able to move on to a new thought. For example, in being aware of some present feeling, I should have always to think that I think about that feeling, and further to think that I think of thinking about it, and so on ad infinitum. It must be that I stop reflecting on all these reflections, and that eventually some thought is allowed to occur without being thought about; otherwise I would dwell for ever on the same thing.²⁴

According to Gennaro, this is the background for Leibniz's (unprecedented) introduction of the notion of unconscious mental states, his *petites perceptions*.²⁵ The idea is that Leibniz cleverly notices the problem of the infinite regress, that for every conscious state there would have to be a higher-order conscious state and so on *ad infinitum*, and so made room for the thesis that there are some perceptions of which we are not conscious. On Gennaro's reading, the first part of this quotation (that is, "it is impossible that we reflect explicitly upon all of our thoughts") should be taken to mean that what makes our mental states conscious cannot always be higher-order conscious thoughts, and so instead must sometimes be higher-order unconscious thoughts. It seems reasonable to suppose that Leibniz was saying, Gennaro maintains, that what makes a perception a conscious perception is the presence of a higher-order perception directed at it, but that the higher-order perception need not itself be conscious; otherwise, an infinite regress would follow and we would "dwell for ever on the same thing," never "able to move on to a new thought." Leibniz, then, really was a HOP theorist of consciousness in the modern sense – although (somewhat ironically) he is never cited by the theory's current proponents.

We should note at this stage that Leibniz's *New Essays* was not published in Locke's lifetime. (Leibniz finished the *New Essays* in 1704, but refrained from publishing it upon the death of Locke the same year, stating in a letter to Remond dated 14th March 1714 that he disliked "to publish refutations of dead authors, although they might appear during their lifetime and be communicated to the authors themselves.") The *New Essays* was finally published in 1765, after Leibniz's death. Still, the fact that Leibniz did entertain the regress problem suggests, in two ways, that the problem would be unlikely to go unnoticed by Locke. Firstly, it supports the suspicion that the issue was in some sense "in the air" when Locke was developing his system. Secondly, it suggests that it would be natural for the regress problem to occur to a

philosopher who invested much thought on the nature of thinking and mental life, as Locke did. To be sure, this is by no means conclusive or overwhelming evidence for thinking that Locke was aware of the regress problem. But it is some evidence.

The upshot is that Locke is very likely to have been aware of the straightforward infinite regress that attends the conjunction of HOP theory and the thesis that all mental states are conscious. Even if he had not himself thought of it, there were probably opportunities for him to happen upon it in the writings of the previous masters and perhaps even his contemporaries. It is impossible, of course, to rule out entirely the possibility that Locke remained unaware that the combination of HOP theory and the thesis that all mental states are conscious is untenable, but our contention is that this interpretation should be considered a last resort.

We should keep in mind that if we can interpret Locke to fall into the Aristotle-Brentano line of tradition, we should – if only for reasons of interpretative charity. As noted above, it would be preferable to avoid ascribing to Locke a view which leads quite straightforwardly to an infinite regress. This seems especially pressing given that Locke is keen to avoid an infinite regress in other areas of his philosophy. We have in mind here Locke's philosophy of religion, in particular his argument for the existence of God. Knowledge of God's existence is classified by him as demonstrative. Demonstrative knowledge contrasts with intuitive and sensitive knowledge. Sensitive knowledge is based on our certainty that the production of ideas in the mind via the senses is by objects that exist outside of us. Intuitive knowledge concerns the mind's immediate perception of an agreement or disagreement of ideas, e.g., when the mind perceives that white is not black or that a circle is not a triangle (ECHU 4.2.1). This is the irresistible, indubitable and most certain degree of knowledge, likened to bright sunshine that "forces it self immediately to be perceived, as soon as ever the Mind turns its view that way"

(ECHU 4.2.1). Demonstrative knowledge occurs when the mind perceives the agreement or disagreement of ideas by means of intermediate ideas called proofs. This sort of knowledge is not as clear and bright as intuitive knowledge, taking “pains and attention” (ECHU 4.2.4). Examples are mathematical proofs where one cannot grasp the entire series of steps in a single intellectual apprehension but must trace through each step in turn. To find a proof of God’s existence, Locke thinks that we need look no further than ourselves and our knowledge of our existence. Since I know that I exist as a knowing being, and since nothing can be made to exist except by something else which both exists and has powers at least equal to those created beings, it follows that there must have existed an all-knowing, all-powerful being, God (ECHU 4.10. 3-6). The idea here is that every knowing being proceeds from some other knowing being; but there cannot be an infinite regress of knowing beings; so there must have been a first knowing being, God.

Given Locke’s concern to avoid a regress in arguing for God’s existence, it is highly plausible to suppose that he would probably be anxious to avoid a regress in other areas of his philosophy, and that he was generally sensitive to threats of infinite regress. This makes it almost imperative to adopt an interpretation of his theory of consciousness that does not lead straightforwardly to an infinite regress.

5. Textual Evidence: Locke’s Discussion of Consciousness as the Perception of What Passes in a Man’s Own Mind

There is a passage in which we may find material to suggest that Locke took one’s awareness of S_1 to be part of S_1 , rather than part of a separate state S_2 . This is, in fact, the famous passage in the Essay where he lays down that consciousness is the perception of what passes in a man’s

own mind. The point of the passage is to counter the view that our mind thinks constantly even though we may not always perceive it:

If they say, The Man thinks always, but is not always conscious of it; they may as well say, His body is extended, without having parts. For 'tis altogether as intelligible to say, that a body is extended without parts, as that anything thinks without being conscious of it, or perceiving, that it does so. They who talk thus, may, with as much reason... say, That a Man is always hungry, but that it does not always feel it: Whereas hunger **consists in** that very sensation, as thinking **consists in** being conscious that one thinks... Consciousness is the perception of what passes in a man's own mind (ECHU 2.1.19; boldface added for emphasis).

The evidence here is twofold.

First, Locke claims that thinking consists in being conscious that one thinks. The locution “consists in” suggests a logical or conceptual relationship between a thought and the awareness of it. It suggests that there is an internal connection between S_1 and one's awareness thereof. In saying that thinking consists in being conscious that one thinks, Locke is effectively claiming that there is no logical gap between a conscious state and the subject's awareness of being in that state. Moreover, this choice of words appears to rule out there being a temporal lag between the conscious state and the awareness of it.²⁶ Such a temporal lag is possible, indeed virtually inevitable, within a HOP framework, but (obviously) impossible within a SOP account. If consciousness required the occurrence of an extra mental act to constitute the perception of what passes in a person's own mind, then assuming that such an occurrence is not instantaneous but involves some temporal duration, the occurrence would lag temporally behind the occurrence of the perceived mental state.

Second, Locke evidently thinks that the notion of there being S_1 without an awareness of S_1 is analogous to the notion that there is an extended body without spatial part. Both are incoherent, indeed unintelligible. It is not that extended bodies happen to have the property of

having parts; rather, if they did not have parts, they would not be extended bodies. Likewise, it is not that conscious states happen to have the property of being accompanied by consciousness of them; rather, if they were not so accompanied, they would not be conscious states. This is important because it suggests a non-contingent relation between conscious states and one's perception of their occurrence. But the connection between any two numerically distinct events – in this case, between the occurrence of two separate mental states – is always contingent. When the two are numerically distinct, they are “distinct existences,” meaning that one could occur in the absence of the other. That is to say, HOP theory goes hand in hand with a contingent relation between a conscious state and the awareness of it. Inasmuch as the connection is non-contingent, or necessary, it cannot be a connection between two distinct existences. It must be a connection between two aspects or parts of a single event or state.

The problem with this evidence is that it is not extensive. The impressions of one passage can always be explained as an aberration. At the same time, there is also material in Locke's discussion of the relationship between consciousness and selfhood, in the context of his well-known discussion of personal identity, that recommends this interpretation.

Locke declared personal identity to be a “forensic” notion, implying responsibility for actions and needed for courts of law and the Day of Judgement (ECHU 2.27.26). In order to deal out the appropriate legal and divine punishment, one must be considered the same person, so if A is the same person as B, then B is responsible for the actions of A. It turns out to be very important for Locke to hold that we are always conscious of our perceptions in his theory of personal identity.²⁷ This is because personal identity is grounded in consciousness and in memory. A person's identity extends as far back in the past as s/he can be said to be conscious of any past action, thought, or event as being his or hers. As a person is conscious of a past thought

or action being their own past thought or action, so that person is conscious of themselves as extending beyond the present into the past. This consciousness of the past is memory, which for Locke is essential to identify persons through time, and memory is always recollection of a person's own doings (such as thoughts, sayings, and actions) in the past. Thus Locke states: "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 as it has on any present action, so far as it is the same personal self" (ECHU 2.27.10).

In the context of discussing these matters, Locke says that a person can only know herself or himself as the same thinking thing in different times and places by consciousness, "which is inseparable from thinking ... When we see, hear, smell, taste, feel, meditate, or will any thing, we know that we do. Thus it is always as to our present Sensation and Perceptions." (ECHU 2.27.9) This passage, in particular the choice of the word "inseparable" to describe the relationship between thinking and one's consciousness thereof, offers further evidence in favor of the SOP interpretation that Locke thought of the relationship between a conscious state and one's awareness of it as non-contingent. Locke makes clear that perceptions are always accompanied "with their consciousness" and that there can be no separation between our perceptions and our consciousness of them, for we literally cannot "think without being conscious of it" (ECHU 2.27.9-13).

6. Theoretical Evidence: Locke on Ideas, Perception, and Self-Knowledge

Locke apologizes at the beginning of the Essay for his frequent use of the term "idea," defined as "that term which ... serves best to stand for whatsoever is the object of the understanding when a man thinks" (ECHU 1.1.8). It is commonplace in expositions of Locke's system in the Essay to point out that Locke's usage of the term "idea" is sloppy. Locke is guilty, we are told, of using

the term ambiguously, referring sometimes to the act of experiencing or perceiving and at other times to the object (or quality) being experienced or perceived. Thus, Richard Aaron states that Locke uses “the term idea ... ambiguously” and that he “ought to have analyzed [it] with greater care”.²⁸ In the same vein, J. B. Mabbot complains that Locke never clears up “obvious problems” to do with his definition of idea and that “it would have helped if he had distinguished clearly between the mental acts of “thinking” (perceiving, imagining etc.) and the ideas with which these acts are concerned” (1973: 73). More recently, Jonathan Bennett declares that Locke’s use of “idea” is a “terminological nuisance” and spends much time untangling the many and various ways in which Locke uses the term “idea” in his work.²⁹

Yet this would be a curiously naïve mistake for Locke to commit, and with such systematicity no less. Consider Book 3 of the Essay which is devoted to language, making it plain that “neither the inveterateness of the Mischief, nor the prevalency of the Fashion, shall be any Excuse for this, who will not take Care about the meanings of their own Words, and will suffer the Significancy of their Expressions to be enquired into” (ECHU 10). Locke thought that if we could only be clear and accurate in our use of language, if we had precise, determinate ideas as the referents for words, which “stand for ... the ideas in the mind of him that uses them” (ECHU 3.2.2), disputes will be minimized and controversies put to an end. He identifies six linguistic abuses, including employing words without clear and distinct ideas, the inconsistent application of words, purposeful obscurity, by giving old words new meanings or introducing new or ambiguous terms without defining them, and assuming that others know what they mean by their words when it is really not at all clear what they mean (ECHU 3.10.1-17). All of these abuses would be committed by Locke himself in his use of “idea” if the standard criticism is correct. Locke goes on to highlight remedies for such linguistic abuses, including not using a

word without a clear idea of what it means, assigning words the same meaning that others assign them, defining one's terms if there is a chance of ambiguity, and finally always use words consistently (ECHU 3.11.1-27). If one "varies the signification of any term ... there it concerns the Writer to explain his meaning, and shew in what sense he there uses that term" (ECHU 3.11.27). Given this evident preoccupation with using clear and consistent terminology and insistence that authors must let the readers know if there is any variation on the meaning of the term, to accuse Locke of using a central piece of terminology ambiguously throughout the Essay should most certainly be avoided as much as possible.

One of the important results of the SOP interpretation developed above is that it vindicates Locke on this matter. Locke is not using the key term "idea" ambiguously if on his view ideas are in fact both acts and objects. And the SOP view of consciousness entails that a (conscious) perceptual act is indeed always its own object, and is thus both an act and an object. The perceptual act may have other objects as well, but it is always also intentionally directed at itself. Thus the charge that Locke is naïvely using "idea" ambiguously may simply rest on a failure to appreciate Locke's conception of the intentional structure of consciousness.

Another element cast by SOP in new light is Locke's account of perception and reality. The standard interpretation has traditionally portrayed it as a paradigmatic form of indirect realism or representative theory, according to which we do not perceive the external world directly but rather indirectly. What we perceive directly, on this view, are ideas or perceptions. These in turn represent objects in the real world, which are thus perceived indirectly. Bennett thinks of it as the "veil of perception doctrine," according to which the objective world of real things is "placed beyond our reach on the other side of the veil of perception".³⁰ The standard interpretation draws on clear passages in which Locke appears to claim that the direct objects of

perceptions are ideas or internal representations of external objects. Recently, however, some commentators have argued for a direct-realist reading of Locke, pointing to some passages suggesting that what we perceive directly are external objects. E. J. Lowe, for example, in his book *Locke on Human Understanding*, defends an interpretation of Locke as holding a version of direct realism which affirms that we do perceive real physical objects existing independently of us but denies that we perceive these physical objects only indirectly by virtue of perceiving some mental objects of some sort such as ideas.³¹

Again, both claims – that we perceive directly external objects and that we perceive directly internal representations – can be true simultaneously if one embraces a SOP account of perceptual consciousness. On the SOP account, a perception always represents both an external object and itself. Moreover, likely it represents itself precisely qua representation of the external object in question. Thus, suppose a person (consciously) perceives a tree. The perception is a representation of both the tree and itself. Inasmuch as it is a representation of the tree, its immediate or direct object is the external tree. But inasmuch as it is a representation of itself, its immediate or direct object is a mere representation of a tree, and the tree itself becomes an indirect object. If so, the perception of the tree has, in effect, two separate immediate objects – objects whose perception is unmediated or “direct.” One is the external object, the other is a mere representation of the external object. Locke suggests as much when accounting for our knowledge of the existence of external objects, where he concludes with certainty that our sensory ideas are directly produced by external objects that “really exist without us”: these external objects affect our sensory faculties and “actually produce that idea which we then perceive” (ECHU 4.11.9). There need not be competition between direct perception of the outer world and direct perception of an inner representation thereof, then. Both can cohabit in the

reflexive intentional structure of (conscious) perception.³² The SOP interpretation thus resolves the apparent tension between those passages in Locke (and readings of Locke) that suggest indirect realism and those that suggest direct realism.

Let us conclude this section with a final instance of the SOP interpretation casting new light on a theme from Locke. One Lockean thesis that has received relatively little attention is the claim that we have intuitive knowledge of ourselves. We have already had occasion to observe (in §4) that intuitive knowledge is knowledge that is not evidentially based – that is, not knowledge acquired by inference from particular ideas – and carries with it the highest degree of certainty.³³ With respect to self-knowledge, Locke writes that “The Knowledge of our own Being, we have by intuition” (ECHU 4.11.1); and “Experience convinces us that we have Intuitive Knowledge of our own Existence and an infallible Perception that we are” (ECHU 4.9.3). He thinks that “we perceive it so plainly ... that it neither needs nor is capable of any proof” (ECHU 4.9.3). This thesis is hard to reconcile with Locke’s empiricism, as Pappas (1998) shows. An empiricist would have to hold that knowledge of the self is mediated by inference from particular ideas, and more generally that no knowledge is intuitive. So Locke’s account of self-knowledge does not seem to be empiricist in spirit. The difficulty in reconciling this view with empiricism may partly explain the relative neglect of Locke’s views on self-knowledge and its intuitiveness in the secondary literature, as it threatens the integrity of the standard empiricist reading of Locke.

Here again the SOP interpretation may be helpful in mediating a workable reconciliation. If all of one’s perceptions are intentionally directed at themselves, then every piece of knowledge one has contains at least implicit reference to oneself as subject of these perceptions. Locke certainly seems to think that the existence of the self is implied in every state of consciousness.

Every element of our experience, every idea of which we are conscious, is a certificate of our own existence, as the subject of that experience:

For nothing can be more evident to us than our own existence. I think, I reason, I feel pleasure and pain: can any of these be more evident to me than my own existence? If I doubt of all other things, that very doubt makes me perceive my own existence, and will not suffer me to doubt of that. For if I know I feel pain, it is evident I have as certain perception of my own existence, as of the pain I feel: or if I know I doubt, I have as certain perception of the existence of the thing doubting, as of that thought which I call doubt (ECHU 4.9.3).

Thus, Locke concludes that in “every act of sensation, reasoning or thinking, we are conscious to ourselves of our own being” (ECHU 4.9.3). Although knowledge of specific features of the self might still require reflection and evidential inference, knowledge of the very existence of oneself as perceptual subject “comes for free,” so to speak, with every perception, and is thus intuitive. This is explained neatly as a by product of the fact that every mental act is partly about itself, and moreover may be about itself qua the subject’s.^{34,35}

7. Conclusion

In this paper, we have argued that there are good reasons to regard Locke’s theory of consciousness as a forerunner of the Same-Order Perception theory of consciousness, according to which a mental state is conscious when and only when it is a quasi-perception of itself. To summarize, the argument proceeds as follows. There is textual evidence aplenty that Locke believed that conscious states are perceived by their subject. This means that a conscious state is the object of a perceptual or quasi-perceptual mental state. That (quasi-)perceptual state must be either (a) part of the conscious state itself or (b) a separate mental state. However, (b) leads to infinite regress, and Locke was unlikely to be unaware of that. By contrast, (a) leads to a variety

of welcome results in the study of Locke's philosophy of mind. Therefore, it is likely that Locke held the view that conscious states are (quasi-)perceptually directed at themselves.³⁶

Endnotes

¹ Functionalism is often considered a huge liberalization over behaviorism, and rightly so, but after all it ventured to explain mental life in terms of the behavior of the brain rather the behavior of the body; and consciousness was always suspected to be a super-behavioral phenomenon.

² All references to Locke are taken from *An Essay Concerning Human Understanding*, P. H. Nidditch (ed.), Oxford: Oxford University Press, 1975. References cite ECHU followed by book, chapter and paragraph number.

³ Armstrong, 1968, *A Materialist Theory of the Mind*, New York: Humanities Press, p. 324. See also Armstrong's "The Nature of Mind" in N. Block (ed.), 1980, *Readings in Philosophy of Psychology Volume 1*, Cambridge MA: Harvard University Press, p. 199.

⁴ Lycan, 1990, "Consciousness as Internal Monitoring", in *Philosophical Perspectives* 9: 1-14. Page numbers refer to the reprint in N. J. Block, O. Flanagan, and G. Guzeldere (eds.), 1997, *The Nature of Consciousness: Philosophical Debates*, Cambridge MAS: MIT/Bradford, p. 755.

⁵ Or is at least a part, or an element, of its object. This subtlety will not concern us here.

⁶ One obvious requirement is that S_2 occur more or less simultaneously with S_1 . If S_2 occurs a years after S_1 , S_1 does not retrospectively become conscious. Another requirement, stressed by D. M. Rosenthal, is that S_2 be non-inferentially and non-observationally produced in reaction to S_1 . See Rosenthal, 1990, "A Theory of Consciousness", ZiF Technical Report 40, Bielfield, Germany, rpr. in Block et al. (1997).

⁷ Unless, of course, it is the object of a third-order state S_3 . According to Rosenthal (1990), this is precisely the case in states of introspective consciousness. See also Rosenthal, 1986, "Two Concepts of Consciousness", in *Philosophical Studies* 94: 329-359.

⁸ The former version of the HOM theory is defended by Rosenthal (1986, 1990; see also Rosenthal, 2002, “Explaining Consciousness”, in D. J. Chalmers (ed.), *Philosophy of Mind*, Oxford and New York: Oxford University Press) and P. Carruthers, 1996, *Language, Thought, and Consciousness*, Cambridge and New York: Cambridge University Press and, 2000, *Phenomenal Consciousness*, Cambridge: Cambridge University Press. The latter is defended by Armstrong (1968) and Lycan (1990). See also Lycan, 1996, *Consciousness and Experience*, Cambridge MA: MIT/Bradford Books.

⁹ The Internet Encyclopedia of Philosophy entry on HOP assumes that Armstrong and Lycan’s view of consciousness is rooted in Locke; see <http://www.iep.utm.edu/c/consc-hi.htm>. See also <http://www.iep.utm.edu/i/introspe.htm> and <http://plato.stanford.edu/entries/self-knowledge/>. See also A. Byrne, 1997, “Some Like it HOT: Consciousness and Higher-Order Thoughts”, *Philosophical Studies* 86, p. 104 and P. Carruthers, “HOP over FOR, HOT Theory”, in R.J. Gennaro (ed.), 2004, *Higher-Order Theories of Consciousness: an Anthology*. Philadelphia: John Benjamins, p. 118.

¹⁰ Guzeldere, 1996, “Is Consciousness the Perception of What Passes in One’s Own Mind?”, in T. Metzinger (ed.), *Conscious Experience*, Padborn: Schoeningh-Verlag, p. 789.

¹¹ See Lycan (1990.) Rosenthal (1986) is particularly explicit on these matters, though from the perspective of a HOT theorist.

¹² For criticism of the HOM theory along these lines, see G. Rey, 1988, “A Question about Consciousness”, in H. Otto and J. Tueidio (eds.), *Perspectives on Mind*, Norwell: Kluwer Academic Publishers, rpr. in Block et al. (1997), A. Goldman, 1993, “Consciousness, Folk Psychology, and Cognitive Science”, in *Consciousness and Cognition* 2: 364-383, rpr. in Block et al. (1997) and Kriegel, 2003a, “Consciousness as Intransitive Self-Consciousness: Two Views and an Argument”, in *Canadian Journal of Philosophy* 33: 103-132.

¹³ F. Brentano, 1874, *Psychology from Empirical Standpoint*, ed. O. Kraus, ed. of English edition L.L. McAlister, trans. A.C. Rancurello, D.B. Terrell, and L.L. McAlister, London: Routledge and Kegan Paul, 1973. The view is first introduced in Section 7 of chapter II (“Inner Consciousness”) in Book 2, which is entitled “A Presentation and the Presentation of that Presentation are Given in One and the Same Act.” In this section, Brentano canvasses his conception of conscious experiences as self-representational.

¹⁴ For a fuller discussion of Brentano's account of consciousness, and the role of the infinite regress argument in it, see Kriegel 2003a and, 2003b, "Consciousness, Higher-Order Content, and the Individuation of Vehicles", in *Synthese* 134: 477-504.

¹⁵ It is sufficient to prefer a HOP interpretation of Locke over a Higher-Order Thought interpretation, so when that is the salient contrast, it is clear why the relevant passages are taken to support a HOP interpretation. But when the contrast is between HOP and SOP interpretations of Locke, these passages carry no weight. We suspect that it has been customary to interpret Locke as a HOP theorist mostly because the salient contrast the interpreters had in mind was indeed that between HOP and Higher-Order Thought views.

¹⁶ We are somewhat skeptical about the weight that should be given to such considerations in this context, but in any event, we take these two considerations to cancel each other out so that neither interpretation is more parsimonious than the other.

¹⁷ This (major) component of the case for the SOP interpretation of Locke is already accomplished, or at least is meant to have been accomplished, through the discussion of §2.

¹⁸ See M. A. Kulstad, 1984, "Locke on Consciousness and Reflection", *Studia Leibnitiana* XVI/2 and 1991, *Leibniz on Apperception, Consciousness, and Reflection*. Munich: Philosophia, chapter 4.

¹⁹ We are indebted to Shoshana Smith for the handling of this objection to the SOP interpretation.

²⁰ Caston, 2002, "Aristotle on Consciousness", in *Mind* 111: 751-815. Caston in fact suggests that Brentano was directly influenced by Aristotle, which is highly plausible given Brentano's general interest in Aristotle's philosophy. Moreover, Caston notes that an interpretation of Aristotle along these lines is to be found in a dissertation on the unity of mental life in Aristotelian philosophy, written in 1873 under Brentano's supervision by one J. Herman Schell.

²¹ There is a long tradition of reading this passage in terms of perceptual modalities, or senses, rather than in terms of perceptual states, or activities. But that is not how Brentano reads it, and Caston (2002) argues in favor of Brentano's reading. According to Caston, the Greek word used by Aristotle, αἰσθησις, is ambiguous between sense and state. We are using here Caston's translation of the passage.

²² Aristotle, *De Anima*, trans. R. D. Hicks, Cambridge: Cambridge University Press, 1907.

²³ Rodis-Lewis, 1950, *Le probleme de l'inconscient et le cartesianisme*, Paris: Presses Universitaires de France; Wider, 1997, *The Bodily Nature of Consciousness*, Ithaca NY: Cornell University Press; Nadler, 1989, *Arnauld and*

the Cartesian Philosophy of Ideas, Princeton NJ: Princeton University Press; Brook, 1994, *Kant and the Mind*, Cambridge: Cambridge University Press (see also Brook's 2001, "Kant, Self-Awareness, and Self-Reference", in A. Brook and R. C. Devidi (eds.), *Self-Awareness and Self-Reference*, Dordrecht: John Benjamins and, 2006, "Kant: A Unified Representational Base for All Consciousness", in Kriegel and K. Williford (eds.), *Consciousness and Self-Reference*, Cambridge MA: MIT Press); Sturma, 1995, "Self-Consciousness and the Philosophy of Mind: A Kantian Reconsideration", in *Proceedings of the Eighth International Kant Congress*, Vol. 1, Milwaukee WI: Marquette University Press; Gennaro, 1996, *Consciousness and Self-Consciousness: A Defense of the Higher-Order Thought Theory of Consciousness*, Amsterdam: John Benjamins Publishing; Thiel, 1991, "Cudworth and Seventeenth-Century Theories of Consciousness", in S. Gaukroger (ed.), *The Uses of Antiquity: The Scientific Revolution and the Classical Tradition*, Dordrecht: Kluwer Academic Press; Lehrer, 1989, *Thomas Reid*, London: Routledge; Hossack, 2002, "Self-Knowledge and Consciousness", *Proceedings of the Aristotelian Society* 102: 163-181 and Buras, 2005, "The Nature of Sensations in Reid", *History of Philosophy Quarterly* 22.

²⁴ Leibniz, *New Essays on Human Understanding*, trans. and eds. P. Remnant and J. Bennett, Cambridge: Cambridge University Press, 1981, p. 118.

²⁵ Gennaro, 1999, "Leibniz on Consciousness and Self-consciousness", in Gennaro and C. Huenemann (eds.), *New Essays on the Rationalists*, New York and Oxford: Oxford University Press. It is worth noting that Leibniz was certainly not the first to propose such an idea: Aquinas, for example, had a similar view.

²⁶ We are indebted to Mark Kulstad for raising the issue of temporal lag and its implications for the choice between the HOP and SOP interpretations.

²⁷ We are indebted here to Mabbot's discussion of self-knowledge in *John Locke*, London: MacMillan, 1973.

²⁸ Aaron, 1955, *John Locke*, Second edition, Oxford: Clarendon Press, pp. 106-7.

²⁹ Bennett, 2001, *Learning from Six Philosophers*, Oxford: Oxford University Press, Volume 2, p. 25.

³⁰ Bennett, 1971, *Locke, Berkeley, Hume: Central Themes*, Oxford: Oxford University Press, p. 69.

³¹ Lowe, 1995, *Locke on Human Understanding*, London: Routledge, p. 61. The claim that Locke was a direct realist can also be found in several of J. Yolton's writing (e.g., Yolton, 1984, *Perceptual Acquaintance from Descartes to Reid*, Minneapolis: Minnesota University Press), as well as in unpublished manuscripts by Alston and Pappas (see, e.g., G. Pappas, Forthcoming, "Locke's Account of Perception", unpublished manuscript). But according to Pappas, the notion that Locke takes us to perceive directly both the external world and an internal representation of it is in

some ways more plausible than the competing readings of Locke as a straightforward direct or indirect realist (personal communication).

³² We keep mentioning “conscious” in parenthesis, because for Locke all perception is conscious, and hence it should go without saying that his take on conscious perception and his take on perception simpliciter are one and the same.

³³ For a study of intuitive knowledge and self-knowledge in Locke see G. Pappas, 1998, “Epistemology in the Empiricists”, in *History of Philosophy Quarterly* 15: 285-302.

³⁴ We should note that Pappas, in his discussion of this issue, claims precisely that intuitive self-knowledge extends only to one’s existence and perhaps “thingness,” while more substantive knowledge, e.g., of one’s character traits, is not intuitive (Pappas 1998: 293).

³⁵ This is, in effect, a position very close to R. Chisholm’s view of self-knowledge; see Chisholm, 1969, “On the Observability of the Self”, in *Philosophy and Phenomenological Research* 30: 7-2. See also Kriegel, 2004, “Consciousness and Self-consciousness”, in *The Monist* 87: 185-209.

³⁶ For comments on a draft of this paper, we are immensely indebted to Peter Anstey, Mark Kulstad, Paul Lodge, George Pappas and Chad Wiener. The paper also benefited from presenting the draft on three different occasions, at Rice University, University of Arizona, and Lewis & Clark College; we thank the audiences there.