

Strawson on Locke's Theory of Personal Identity

Masaki Ichinose
(The University of Tokyo)

1. Consciousness and Concernment

The concept of "Person" that John Locke once proposed has become an absolute standard in this field of philosophy ever since. Actually, there would be no contemporary arguments concerning this concept, which do not take the least notice of Locke's theory, and if there is such, that must be insincere or lazy. However, in spite of this prevailing understanding of its importance, Locke's theory of person has not been so seriously examined as rightly deserved. This is probably, because of a historically famous criticism against Locke's theory through the logical law of transition as Berkeley and Reid once raised, or because of such doubt, which Joseph Butler once suggested, that Locke's theory sounds to be somehow circular. Locke bases personal identity (i.e. who he/she is) upon consciousness, but this argument seems to presuppose whose consciousness is focused upon. Additionally, we have to notice a historically fatal mistake or oversight, which has unbrokenly continued to be made as a pretty much public understanding. Namely, the concept of person in Locke's *An Essay concerning Human Understanding* and in *the Second Treatise* of Locke's *Two Treatises of Government* have been discussed separately for the reason that those two works of Locke belong to different fields or subjects respectively, although the concept of person is exactly the same concept that the same author uses in both. Considering the overwhelming importance of the concept of person for philosophical investigations, this strangely twisted circumstance, which we can easily find out once we rethink calmly, must be corrected and improved.

Galen Strawson's understanding of Locke's theory of personal identity developed in his book, *Locke on Personal Identity*, could be positively evaluated as one interesting attempt to create a stir in the traditional stream of interpretation on Locke. Strawson's contribution to this field lies not so much in proposing a new viewpoint as rightly emphasizing again the sound understanding of Locke's argument, which has not, quite strangely, planted its root in Locke studies despite the fact that many scholars have pointed out the understanding, and forcefully correcting the course of contemporary trends in

Locke studies.

First of all, it is worthy of special mention that Strawson clarified how to respond to the Berkeley-Reid criticism above referred to by correctly pointing out that “consciousness” constituting the core of Locke’s theory is not the same as mere memory but something always accompanying “concernment”. The Berkeley-Reid criticism can be summarized as follows; according to Locke’s theory of personal identity based upon the concept of consciousness, such a person *P* in his youth as remembers an event which *P* experienced in his childhood is supposed to be the identical person to the person *P* in his childhood. Further, if the person *P* in his old age remembers an event which *P* experienced in his youth, but does not remember an event which *P* experienced in his childhood, then the person *P* in his old age is supposed to be the identical person to the person *P* in his youth, but the person *P* in his old age is not supposed to be the identical person to the person *P* in his childhood. However, this argument violates the logical law of transition; hence Locke’s theory must be inconsistent.

To this criticism, Strawson makes it quite explicit that Locke’s notion of consciousness is not the same as memory by indicating that “consciousness”, which Locke focuses upon is a phenomenon in the present moment (hence, it is characteristically different from memory), and every memory does not necessarily accompany “concernment” that Locke attaches the importance to in arguing the issue of personal identity (Strawson 2011, 72-6). The inconsistency accusation that Berkeley and Reid leveled at Locke’s theory would be refuted by this Strawson’s argument, for the Berkeley-Reid criticism clearly presupposes an interpretation of regarding “consciousness” as exactly the same as “memory”. Rather, the personhood ‘is differently constituted every day, on Locke’s view’ (Strawson 2011, 55), thus the circumstance that the Berkeley-Reid criticism presented conforms to (rather than rebuts) the significance of the Lockean “person”, as opposed to what the Berkeley-Reid criticism initially intended.

Furthermore, Strawson attempts to adapt the basic elements of personhood, i.e. “consciousness + concernment”, to Locke’s another fundamental strategy to understand “person” as “forensic term”, by developing an interpretation of taking the notion of responsibility very seriously. In the light of the traditional tendency in Locke studies since the Berkeley-Reid criticism to discuss Locke’s theory of person from a cognitive point of view, Strawson’s emphasis of responsibility, although it must be simply right from the outset as far as we read Locke’s text honestly, should be positively evaluated in the

sense of returning the twisted trend to a proper course.

Additionally, it is quite interesting that Strawson clearly highlights “the Day of Judgment” as the theoretical background, which Locke’s theory of person or personal identity finally relies upon, and suggests a possibility for it to play the role of theoretically justifying Locke’s argument concerning consciousness. That is to say, we would be ultimately judged just or unjust on the day of judgement by God as to whether responsibility in each case was justly placed based on consciousness or not. Would I pry into Strawson’s psychological matters too much if I said that his emphasis of “the Day of Judgment” reflected his contemporary inclination of philosophical arguments?

2. The Radical Theory

Anyway, I entirely agree with the main direction of Strawson’s arguments, as that is exactly the proper course in understanding Locke’s theory. However, I cannot agree at all with his way of *developing* this direction. His pivotal point lies in the following; he takes Locke’s thesis, ‘consciousness makes personal identity’ (Locke 1975, 2.27.13), to constitute the core of Locke’s theory, and calls it “the radical theory” if the thesis is interpreted literally, then claims that Locke does not truly assert the radical theory, rather Locke presupposes the human subject of experience in the background of consciousness, which corresponds to (although not being based upon) immaterial substance. Why does he claim so? The ground he relies upon is that; if we interpret Locke to propose the radical theory literally, we cannot eliminate what Locke himself calls “a fatal error” (Locke 1975, 2.27.13), which means a possibility of injustice that “consciousness” of behaviours the agent concerned had nothing to do with might be transferred to the agent, namely that the agent might take responsibility for and suffer punishment for actions the agent did not actually perform, unless we can rely upon the Goodness of God not to let such accidents happen. Strawson calls this argument “the injustice claim”, and says,

No reading of Locke can be right unless it interprets the force of the radical claim in a way that renders it consistent with the injustice claim (Strawson 2011, 131).

In order to illustrate the problematic circumstance occurring here, Strawson argues that; if we accept Locke’s radical theory in a literal way, then we must admit the theoretical

possibility of the existence of a freakish creature, [Sx] (whom Strawson describes as ‘an immaterial-substance-based subject of experience’ (Strawson 2011, 127)), such as appears as a result of thousands of different consciousness’ cross-transferred to one human subject of experience, wherefore it clearly follows that irrelevant reward and punishment are unjustly given. However, Locke does not take such possibility into account. That is because Locke does not take the radical theory literally. Strawson draws such a conclusion (Strawson 2011, 127-30).

Strawson’s argument is so tricky and intriguing that I was attracted to it. However, the argument seems to include some basic misreading. First of all, [Sx] as Strawson formulates, can have “consciousness” of staying at the same time at different places, but this manifestly violates the law of non-contradiction. This situation is utterly opposed to Locke’s definition of “Person”, namely, ‘a thinking intelligent Being, that has reason and reflection’ (Locke 1975, 2.27.9). Reason clearly makes it quite difficult for us to refuse the law of non-contradiction, and the law belongs to a category of intuitive knowledge in the context of Lockean epistemology.

Second, it seems that Strawson’s arguments are deployed through understanding “consciousness” thoroughly from the ‘from-the-inside’ point of view to take it to be psychological phenomena crystal clear (only) to the agent himself/herself, hence making it possible to introduce a monstrous supposition like [Sx] without considering how consistent it is with external circumstances viewed ‘from-the-outside’. However, we must say that this line of argument sounds to be highly dubious, as long as Strawson’s interpretation of Locke’s theory of personal identity is primarily developed by focusing upon a close connection between the notion of “Person” and responsibility.

It is blindingly obvious that, even if we accept “consciousness” as the criterion according to which we judge whether to attribute responsibility to a particular agent or not, such the attribution of responsibility cannot be determined crystal clearly or uniquely at all. Remember issues of responsibility in the case of mental distraction, drunkenness, impulsive actions, ignorance, mistake, recklessness, negligence, or complicity. Just a glance at those legal, complicated, matters easily enables us to realize that matters on attribution of responsibility are so extremely hard to deal with that our society needs courts to sort those out and reach legal conclusions. “Consciousness” viewed ‘from-the-inside’, at least if applied alone, cannot be the solution at all. If it functions as a solution, why do people struggle through litigation? Should all be so simple? It seems to me that,

although it is an obviously inescapable fact in our society that problems on attribution of responsibility based upon the concept of person are so intricate, Strawson's argument on Locke's consciousness as the basis of responsibility simply skips this intricacy and tacitly presupposes the (actually unrealizable) possibility that the attribution of responsibility can be uniquely determined in terms of "consciousness" viewed 'from-the-inside'.

However, I want to say again, it is extremely obvious that we cannot determine the attribution of responsibility on human actions at all only from the first-person's point of view. Actually, if that were admitted, legal concepts like criminal negligence or a principal offender in the second degree would not make sense and the system of criminal law must suffer fundamental changes. If Locke's theory of personal identity led to such a consequence necessarily, his theory must sound to be extravagantly weird. That must be unbelievable.

3. Forensic Term

Why does Strawson regard "consciousness" viewed 'from-the-inside' as being capable of dealing with intricate circumstance about attributions of responsibility, as opposed to our common understanding on legal, perplexed, conditions in this issue that I mentioned above? It seems that's because Strawson's emphasis on Locke's expression of "Forensic Term" functions merely as a façade, resulting in being just a verbal claim. In order to confirm that Strawson's argument does not conform to Locke's ideas as well as our common sense, we only have to remember Locke's famous description about the drunkard. Locke says;

Is not a Man Drunk and Sober the same Person, why else is he punish'd for the Fact he commits when Drunk, though he be never afterwards conscious of it?.....the Drunkard perhaps be not conscious of what he did; yet Humane Judicatures justly punish him; because the Fact is proved against him, but want of consciousness cannot be proved for him (Locke 1975, 2.27.22).

Then, the ultimate place where this punishment is judged as just or not is put into "the Great Day" (ibid.). As is evident in this Locke's passage, "consciousness" mentioned here is the one from the first-person's point of view. Locke clearly develops an argument

that this kind of the first-personnel “consciousness” is neither almighty nor decisive in judging whether the agent should be punished or not. However, at the same time, Locke is supposed to keep what is called “the radical theory”, which implies that the concept of “consciousness” constitutes the essential core of personal identity, i.e. an identity of subject responsibility or punishment is inflicted to. If that is the case, we should naturally interpret that the concept of “consciousness” Locke highlights is not necessarily restricted to the first-personnel one. That must be supposed to be beyond that.

In any case, we should admit that it must be extremely far from being perfect, as it were, by definition, to determine personal identity by “consciousness” in the first-person’s sense and then to attribute responsibility, as the first-personnel “consciousness” crystal clear to the agent herself/himself is not necessarily the ultimate criterion of the attribution of responsibility in Locke’s context as well as in our actual matters of fact on the court proceedings. Rather, it seems to me, we should suppose that an attribution of responsibility has to be a kind of activity conducted by all the best people concerned can do, namely, an activity conducted from the third-personnel, ‘from-the-outside’ point of view, and always tentatively made by being judged as ‘it should have been so’, where “the Day of Judgment” only functions as a regulative ideal to secure honesty. I think that this reading could fit in well with Locke’s text.

In other words, I am proposing such a way of reading Locke as accepting the radical theory (Strawson calls) literally as a bold and original claim, but understanding “consciousness” by overlapping its original meaning, “common knowledge” (*conscientia* in Latin), or its synonym, “conscience”, to have affinity to ‘from-the-outside’ point of view. An issue on whether a certain agent has “conscience” or not can be judged not only from his/her first-personnel point of view, but also from the third-personnel, external, point of view. That could be well understood by remembering the court procedures or moral judgements on someone’s actions, where it is not the case at all that only the first-personnel point of view is decisively authoritative. This is exactly the point that Locke highlights as to the case of a man drunk. As long as Locke’s argument aims at offering a philosophical perspective of taking “person” to be a responsible agent to whom punishment could be inflicted, it must agree to Locke’s text to interpret his basic concept, “consciousness”, in a normative context (rather than in a merely mental-phenomenal one) related to something like “conscience”. Please refer to my argument (if possible) in Ichinose (1997) (in Japanese) as to this point.

4. Person as the Basis of Property Rights

This way of reading naturally leads us to understand that Locke's theory of personal identity is linked to the Lockean concept of "person" as the basis of property right in his other main work, *the Second Treatise in Two Treatises of Government*. The reason I am convinced of this linkage is that Locke's concept of property right, although it originates in principle from labours of the person, must satisfy what is called "Lockean proviso", where a judgement on whether every particular property right or ownership actually satisfy "Lockean proviso" cannot help being made by people concerned from the third-personnel point of view. For "Lockean proviso" is intrinsically and eventually concerned with a question on whether a certain agent's owning a particular property would infringe other's property rights seriously or not.

In addition, Locke describes how questions at issue on property rights could be ultimately judged in terms of "an appeal to Heaven" (Locke 1960, the Second Treatise, chapter 3, section 20 et al.), which is precisely corresponding to Locke's argument of "the Day of Judgment" in his *Essay*.

Moreover, this Lockean concept of property right arising from "person" is evidently based upon the law of nature; hence, the property right is conceptually connected with the notion of punishment in the case of violation of the law of nature. Actually, Locke seems to define punishment through the concept of property rights. Locke gives an explanation about reparation, which is one of the basic elements of punishment, like the next in the context of arguing an unjust war, which is a kind of crime;

Let the Conqueror have as much Justice on his side, as could be supposed, he has no right to seize more than the vanquished could forfeit; his Life is at the Victors Mercy, and his Service and Goods he may appropriate to make himself reparation; but he cannot take the Goods of his Wife and Children (Locke 1960, the Second Treatise, chapter 27, section 183).

In this argument we should pay special attention to Locke's introduction of the concept of 'forfeit', because it seems that Locke offers here the logic of punishment in terms of forfeiture. Namely, his argument sounds to rely upon such an idea that; "by performing

acts contrary to natural law one forfeits that portion of one's own rights against others that will make an interference in one's own life, proportionate to one's interference with others, morally permissible" (Simmons 1992, 149). Even if we think in a general way, we should logically suppose that how to attribute responsibility in the case of violation of rights must be clarified in order to make sense of the notion of rights; hence, in this sense as well, "person" as a subject of property rights or ownership is intrinsically connected with "person" as a subject of responsibility. I am utterly convinced that we cannot reach adequate understanding of Locke's theory of person or personal identity at all without considering seriously this connection between two appearances of "person" in Locke's main texts.

In reality, Locke himself, even in his *Essay*, describes how person and its consciousness work with regard to its actions as in the following;

That with which the *consciousness* of this present thinking thing... attributes to it *self*, and owns all the Actions of that thing, as its own (Locke 1975, 2.27.17).

Evidently, Locke takes the notion of ownership into account in discussing the notion of person or personal identity. Additionally, Strawson himself stresses how "concernment" matters a lot to the issue of personal identity, then refers to 'one's family or one's property' (Strawson 2011, 158) as objects of our concern. There is no doubt that problems on property rights or ownership naturally come up with discussion about personal identity in any case. Strawson's emphasis of "concernment" must be much to the point in this respect.

I really hope that the understanding of Locke's theory of person or personal identity would be advanced and deepened from all directions (in particular including two main texts of Locke with considering the conceptual connection between responsibility and property rights) in the international contexts of Locke studies.

[References]

- Ichinose, Masaki. 1997. *The Rise of the Person-Knowledge Theory: The Moment of John Locke*. The University of Tokyo Press. (in Japanese)
- Locke, John. 1960. *Two Treatises of Government*, ed. by P. Laslett, Cambridge University Press. (Quotations are made by section numbers)

— 1975. *An Essay concerning Human Understanding*. ed. by P. H. Nidditch, Oxford University Press. (Quotations are made by numbers of book, chapter, and section)

Simmons, Alan John. 1992. *The Lockesn Theory of Rights*. Princeton University Press.

Strawson, Galen. 2011. *Locke on Personal Identity: Consciousness and Concernment*. Princeton University Press.

*This article is based upon my book review on Galen Strawson's *Locke on Personal Identity: Consciousness and Concernment*, which was initially published in Japanese in *Studies in British Philosophy* issued by "The Japanese Society for British Philosophy" in March 2014. In publishing an independent article here in English, I have drastically revised my original review by adding new perspectives.