CHAPTER – 1

PROBLEMS OF PERSONAL IDENTITY:
DIFFERENT PHILOSOPHICAL APPROACHES
In the present chapter we shall discuss the concept of personal identity from different philosophical standpoints. At the beginning we shall explain what personal identity is, and how the concept of personal identity raises various philosophical problems. Next we shall mention various types of criteria proposed by distinguished philosophers. The rest of the chapter is the elaboration of such criteria from various philosophical standpoints.

1.1 What Is Personal Identity?

Theories of personal identity are concerned about identities of persons over time. It is generally believed that persons are existential beings, and similar to other beings various physical and psychological changes occur in the life-span of a person. Nevertheless we identify a person in spite of those changes. We assert identity judgments about our own identities, and we assert identity judgments about the identities of others. Identity judgments are generally expressed in sentences of the form, ‘A existing at $t_1$ is the same as B existing at $t_2$ where A and B refer to the same person, and $t_1$ and $t_2$ refer to different times. Terence Penelhum comments,

“[The problem of personal identity is]...the problem 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.”

This sort of practice is very common to human nature and it involves a strong conviction within itself. The conviction is about the identity of one’s own self and about the identity of the ‘other’. The conviction is important, for it helps us distinguish others’ identities from personal identity. For example, ‘The Morning Star is the same as the Evening Star’ is an identity statement that does not involve any subjective conviction within it. Here ‘identity

---

means identity in reference. This statement, though expresses a theoretical truth, is true for all time and hence, does not require any conviction of identity.

As we noted, a general expression of any personal identity statement is of the form: ‘The person \( P_2 \) at time \( t_2 \) is the same as the person \( P_1 \) at another (supposedly earlier) time \( t_1 \).’ Here ‘identity’ means persistence or re-identity\(^2\). Any personal identity statement can be viewed from two angles: from first person point of view and from third person\(^3\) point of view. From first-person point of view the person \( P_2 \) recognizes (or reidentifies) himself as \( P_1 \), because, he has a strong conviction about his identity or persistence from \( t_1 \) to \( t_2 \). Simultaneously other persons can identify \( P_2 \) at \( t_2 \) to be the same as \( P_1 \) at an earlier time \( t_1 \), for they also have the strong conviction that they are identifying the same person who is continuing as identical from \( t_1 \) to \( t_2 \). The latter is the third person version of personal identity.

According to some philosophers, the root cause of the problem of personal identity (specifically from first-person point of view) is the fact that we think that we (i.e., persons) have a special or privileged access to our own mental states. Among those states there is one about my own identity over time. I know that I am identical with all my previous existences. Persons are distinct from other beings, if the term ‘person’ refers to only first-persons, because only a first-person can remember his own past

---

\(^2\) P.K. Mohapatra divides the problem of personal identity into two distinct questions: the “Unity question” and the “Re-identity question” The unity question is about the nature of persons, and the question is, “whether persons are, as they are commonly supposed to be, single and unitary beings and if so, what is the principle of this unity.” “What makes a later person (to be) the same person as any earlier person” is the question of re-identity and it concerns the nature of personal identity. Mohapatra prefers to use the term “Re-identity” instead of “identity,” for according to him the term “identity” sometimes refers to the “non-trivial issue” of identity. The identity between the Morning Star and the Evening Star is the “non-trivial” issue of identity, where the object comes under different descriptions. In the later case the question of re-identity does not arise. The nature of personal identity should be distinguished from such cases. The nature of personal identity is also distinct from “identity across kinds”, that is, the supposed identity between mental states and one or the other brain-state. Theorists of personal identity are generally concerned with the identity of an object (i.e., person) with itself. Mohapatra calls this the “trivial issue” of identity. [See P.K. Mohapatra, *Personal Identity* (New Delhi: Decent Books, 2000), 3.]

\(^3\) Instead of the term ‘third-person identity,’ philosophers used to express it as the identity of the others.
experiences and actions and can identify himself with all his past existences. It should be noted that some thinkers distinguished between the term ‘person’ and the term ‘man’, for, according to them the ideas annexed to the term ‘person’ are different from those attached to the term ‘man’. We shall revisit this point at later sections. The problem of personal identity is also known as the problem of self-identity when viewed from the first-person perspective. Some philosophers use the terms ‘personal identity’ and ‘self-identity’ interchangeably. They argue that the problem of personal identity deserves a special status only for the peculiarity of the notion and nature of first-persons. Finally, the underlying presupposition of any theory of personal identity is the belief that persons are enduring objects.

1.2 Philosophical Problems Surrounding Personal Identity

The question of personal identity is related to various philosophical problems. For example, the problem of defining characteristics and nature of persons is a core issue in the literature on personal identity. The problem of persistence is another issue related to the constitutive criteria of personal identity. If $A$ is the necessary and sufficient condition for $B$, then occurrence of $A$ is always congruent with the occurrence of $B$. A criterion always refers to that account, which the object of the criterion necessarily (and sometimes sufficiently) consists in.

Personal identity also deals with the problem of evidence. My fingerprint on a glass is the evidence that I am that person who held that glass. Some philosophers hold the view that memory continuity or bodily continuity provides the criterion of personal identity; whereas according to some other philosophers, continuity of memory and body are two commonly accepted evidences in favour of personal identity.

According to most of the philosophers, however, identities of persons could not be understood without the notion of substance. Persons are persistent objects and continue through time. Only the notion of a spiritual substance can explain the notion of personal identity, that is, identity of a person as a single unitary being. Philosophers tend to believe that substances are the unifying entities that unite various existences of a person. Persons are subject to changes, and without the notion of one unifying substance
identities of persons cannot be explained. The problem is, whether there is one continuing substance that underlies different moments of personal existences or whether different substances underlie different existences of a person is contentious. Doubtless there are also some philosophers who altogether deny that persons are substances. There are several other problems that pose serious conceptual difficulties for the notion of personal identity. For example, the problems of fission and fusion and the problem of survival challenge the very notion of ‘identity’. The problems of fission and fusion are two interrelated problems. The root cause of these problems is a belief that questions on personal identity have definite answers. Consider the following ‘thought-experiment’ – a common tool employed by philosophers for analysing difficult problems of personal identity: suppose $X$ is a person and a surgeon removes $X$’s brain from $X$’s body and destroys $X$’s body. $X$’s two brain-hemispheres are severed and transplanted into two brainless bodies called $Y_1$ and $Y_2$. Since $Y_1$ and $Y_2$ share with $X$ all psychological characteristics (as both hemispheres are responsible for $X$’s psychological features, i.e., consciousness) both are candidates for being identical with $X$. But this is absurd. $X$ could not be identical with both $Y_1$ and $Y_2$. $X$ could, in principle, be (numerically) identical with either $Y_1$ or $Y_2$; since, $X$ and $Y_1$ or $Y_2$ do not exist at the same time. There is, therefore, a possibility that $X$ continues to exist as $Y_1$ or $Y_2$. But $Y_1$ and $Y_2$ themselves cannot be (numerically) identical, as $Y_1$ and $Y_2$ exists simultaneously. Naturally, what happens to $X$ (i.e., whether $X$ remains the same post-operation) is a debatable matter. This problem is known as the problem of ‘fission’. The problem of ‘fusion’ is just the converse of ‘fission’. Here two distinguished persons, each of whom is an independent individual, having different characteristics, desires, intentions etc., fuse into one. The resulting person shares some compatible features common to each of the original persons, though some of the competing characteristics of the original persons cancel each other out and are consequently lost. E.g., if one person loves music and the other loves to travel (compatible features), the resulting person may be both a connoisseur of music and a travel enthusiast. On the other hand, if one is a music-lover and the other despises music (competing features), then the resulting person may have no feeling for music at all. From traditional philosophical standpoint fusion is equivalent to death, for fusion would imply changing some of our essential
characteristics and some of our desires, i.e., loss of identity. But, at the same time fusion does not rule out the possibility that part of the original person will continue to survive in the new incarnation of the resulting person. Parfit suggests that fusion, while not entirely survival, is not entirely failure to survive. Thus, according to Parfit, survival, unlike identity is not a matter of all or nothing but can have degrees.

It is worth mentioning that the problems mentioned above have a common cause, and that is the belief that despite the changes in the characteristics of a person the person retains its identity. All the said problems directly or indirectly are concerned with the identities of persons; all of them try to find an answer to the question: if all the observable features of a person are subject to change then what makes us say that these features are features of a single person rather than those of many different persons? The answer to this question may be found in different criteria of personal identity to which we shall now devote our attention.

1.3 Different Criteria of Personal Identity

In the present section we shall look into different approaches towards the question of the constitutive criteria of personal identity. “A criterion is a set of non-trivial necessary and sufficient conditions that determines, insofar as that is possible, whether distinct temporally indexed person-stages are stages of one and the same continuant person. (A temporally indexed person-stage is a slice of a continuant person that extends in three spatial dimensions but has no temporal extension.)”

The criteria of identity of an object depend on the nature, concept and essence of the object that we want to identify. Ordinary material things like trees, stones, table etc. require only spatio-temporal continuity as a criterion. But the criteria of a regiment, corporation, and baseball team etc. are obviously different. Likewise, the criteria of personal identity should be determined by the nature and concept of persons.

Philosophers have come up with mainly two types of criteria for personal identity – the bodily continuity criterion that ascribes identity to the spatio-temporal continuity of a

---

4 The Internet Encyclopedia of Philosophy - http://www.iep.utm.edu
person and the memory criterion which bases identity on the consciousness of past experiences.

### 1.3.1 Bodily Continuity Criteria

Bodily continuity criterion is the most natural and primitive criterion of personal identity. According to the bodily criterion of personal identity, $P_2$ at $t_2$ is identical with $P_1$ at $t_1$ if and only if $P_2$ has the same body that $P_1$ had. Personal identity is not essentially different from identities of material objects, for, like other material objects persons are spatio-temporally continuous, and persist through time. For example, an oak tree that goes through various characteristic changes from the beginning till the end remains the same oak, because of the retention of a particular type of matter and a specific form. “What required for the identity of person $P_2$ at time $t_2$ and person $P_1$ at time $t_1$ is not that $P_2$ and $P_1$ are materially identical (i.e., both $P_2$ and $P_1$ have the same matter) but merely that the matter constituting $P_2$ has resulted from the matter that constituted $P_1$ by a series of more or less gradual replacements in such a way that it is correct to say that the body of $P_2$ at $t_2$ is identical with the body of $P_1$ at $t_1$.”

Bernard Williams brings bodily criterion to the centre stage. His arguments purport to establish the claim that bodily continuity (that is spatio-temporal continuity) criterion is a necessary condition of personal identity and continuity of memory does not provide any sufficient condition of personal identity. The criterion of bodily identity includes the notion of spatio-temporal continuity. Though Williams criticizes memory continuity as an independent criterion of personal identity, he argues that bodily continuity does not completely rule out memory continuity.

According to Williams, all claims of remembering or all kinds of memories are not veridical. Only in some cases the meaning of the expression ‘$A$ remembers $x$’ means $x$ really happened. If a person Charles claims to remember some actions and events (suppose we call these actions and events as ‘$E$’) of the past, then to verify his claims we

---

have to accumulate some positive “corroborative” evidences. Those evidences will show that either Charles was bodily present at the time of occurrence of ‘E’ or he was not. If Charles was bodily present during the occurrence of E then Charles’ memory is veridical or accurate and therefore he is the person who did those actions or witnessed those events. If the evidences show that Charles was not bodily present at the time of occurrence of E, then there are no evidences to verify his claims. It seems that we could hardly distinguish between bodily and psychological or memory criteria of personal identity and that “the normal operation of ‘mental’ criterion involves the bodily one.”

Without the reference of body memory claims cannot be verified. If the events remembered by Charles are supposedly or certainly un witnessed, Charles may be inferred to be clairvoyant. As Williams explains,

“To describe Charles as clairvoyant is certainly not to advance very far towards an explanation of his condition; it amounts to little more than saying that he has come to know, by no means, what other people know by evidence. But so long as Charles claimed to remember events which were supposedly or certainly un witnessed, such a description might be the best we could offer.”

The point was that if a person Charles claims to remember doing certain actions in the past, and if those actions were un witnessed then Charles could not claim his identity and he would be considered as clairvoyant. Any memory claim regarding a past action or event must be witnessed by some other persons, who was present at the time of occurrence of that action, and observed the bodily continuity of the person who is making the claim. It follows that a person could not assure his bodily continuity and therefore could not claim his identity by making sincere memory claims if there is no other who verifies his claims.

But there is a problem. Suppose a witness has observed a person, P’s, bodily presence at times $t_1$ and $t_2$. But he may not have watched P’s bodily continuity incessantly from $t_1$ to $t_2$. Is this sufficient to establish P’s identity? Williams describes a thought experiment,

---

7 Ibid., 5.
8 Ibid., 5.
better known as the ‘reduplication argument’, to elucidate this point. The argument is as follows.

Consider a person Charles, who, after a sudden change of all of his mental-states, acquires exactly the same character of an ancient person Guy Fawkes. His sincere memory claims fit exactly the events and activities of Guy Fawkes’ life. From these given premises the most natural conclusion that can be drawn is that Charles is identical to Guy Fawkes or Charles is Guy Fawkes. After that Williams imagines another person Robert, who at the same time but in a different place undergoes a sudden character change and make memory claims exactly similar to Charles. Robert’s claim also perfectly fits the life of Fawkes. Again, from all of the above premises one can spontaneously draw the conclusion that both Charles and Robert are Guy Fawkes.

But two distinct persons at the same time cannot be identical with a third one, since either Robert or Charles is not spatio-temporally continuous with Guy Fawkes. Moreover, if both were identical with Fawkes, “Guy Fawkes would be in two places at once, which is absurd.”\(^9\) On the other hand, Charles and Robert are not identical themselves, since they exist simultaneously. Williams goes on to say that,

“We might instead say that one of them was identical with Guy Fawkes, and the other was just like him; but this would be an utterly vacuous manoeuvre, since there would be ex hypothesi no principle determining which description was to apply to which. So it would be best, if anything, to say that both had mysteriously become like Guy Fawkes, clairvoyantly knew about him, or something like this. If this would be the best description of each of the two, why would not it be the best description of Charles if Charles alone were changed?”\(^10\)

Clearly, Charles and Robert cannot both be Guy Fawkes; but, more importantly, not even Charles alone can claim to be identical to Guy Fawkes, for sincere memory claims alone cannot provide sufficient ground for identity. To claim identity only on the basis of memory, according to Williams, is therefore vacuous. Williams explained the last

\(^9\) Ibid., 8.
\(^10\) Ibid., 8. Emphasis mine.
point by drawing a distinction between identity and exact similarity. The fact that two persons live in the *same* house is different from the fact that two persons live in exactly similar houses. When we say that two persons are spatio-temporally continuous we actually claim that they are *same*, i.e. identical. But saying that two persons have the same character means that they have exactly similar, but not same, characters. Likewise, two or three distinct persons – who have different bodies – can have exactly similar, but not same memories, and hence, no two persons cannot be identical with each other only on the basis of their memory claims. In case of Charles, “all we actually know is that Charles’ claims fit Fawkes’ life”, but we do not know how to identify Charles with Fawkes only on the basis of memory.

Only in case of body (like physical matter) the distinction between identity and exact similarity can be clearly understood. If a person wants to relate his/her past with the history of another in terms of exact similarity of everything except the body, “we are going to have a serious difficulty in finding a suitable description in these terms of his memory claims.”\(^\text{11}\) To verify Charles’ memory claims we have to identify a person in the past, namely, Guy Fawkes, and we have to know what Fawkes did in the past, and to do that we have to witness Fawkes’s action and therefore the person Fawkes with full bodily continuity. At the same time we have to know whether Charles was bodily present at the time of occurrence of those events and actions which he claims to remember. Otherwise Charles will be considered as clairvoyant. If a later person \(P_2\) claims to remember something of an earlier person \(P_1\) then the justification of this sort of claim depends upon the references of those witnesses who observed the relevant bodily continuity of \(P_1\) with \(P_2\). That is to say, those witnesses observed that \(P_1\) was physically present at the time of those actions which \(P_2\) claims to remember. The witness must also presume that the body of \(P_1\) is spatio-temporally continuous with that of \(P_2\)\(^\text{12}\). If these presumptions prove to be false then \(P_2\) may not be the same as \(P_1\) and \(P_2\)’s memory claim may not be genuine but merely apparent memory claim or clairvoyance. But if these assumptions prove to be true then \(P_2\) is certainly identical with \(P_1\).

\(^{11}\) Ibid., 9.
\(^{12}\) The witness is not likely to have watched \(P_1\) incessantly from time \(t_1\) till time \(t_2\), hence the presumption is required.
No doubt Williams is right when he says that both Charles and Robert cannot be identical with each other and cannot both be identical with Fawkes. What is not clear is that why Charles alone could not claim to be identical with Guy Fawkes. If there are “corroborative evidences” of Charles’ memory, i.e. if someone witnessed Charles at the time of occurrence of his remembered action, and if the person also verified Fawkes’ memory claims and admits that both memory claims are the memory claims of one and the same person, there is nothing that should restrict us to accept Charles’ identity with Fawkes. Williams explained the notion of personal identity in terms of third person eye witness. But it is difficult for an eye witness to observe the spatio-temporal continuity of a person. If the term ‘spatio-temporal continuity’ intends to imply a continuous third-person observation of different spatial and temporal existences of a person then only on the basis of his own memory Charles cannot claim to be identical with Fawkes. The other possibility is Charles is not a spatio-temporal continuer of Fawkes (i.e. they have different bodies) and therefore cannot be claim to be identical with Fawkes. Most probably Williams wants to emphasize the last point.

According to Williams, simple bodily continuity alone is often also not sufficient to explain personal identity. Memory must be invoked. Sometimes we may require to identify a person in terms of memory if the person frequently goes through random physical changes (i.e., changes of some bodily features). If a lady singer often changes her hair colour and make up, the only way to identify her is to take resort of memory, i.e., how she sang, what she did in past stage performances and how she herself recalled those events. If a person like Fawkes had undergone frequent physical changes, identification would become impossible, and “the only other resource would be the memory criterion.”13 In this sense, bodily criterion is necessary for establishing identity-claims but may not be sufficient by itself.

No doubt Williams is right in that bodily continuity is necessary to establish personal identity, but what he could not realize is that not the entire body, but only a part of it is required to continue; that essential part being brain. Continuity of the whole body may

13 Ibid., p.11.
not be compatible, for instance, with Shoemaker’s 14 Brown-Brownson thought experiment. If a surgeon is able to put Brown’s brain in Robinson’s brainless body and if post-operation the new person Brownson survives he will definitely show character, memories and personality quite indistinguishable from Brown. If Brownson claims identity with Brown, Williams’ theory would clearly fail to explain it. As it appears, we have to adopt a new criterion where identity should be established not by means of continuity of the entire body, but in terms of that of the brain.

An ambiguous aspect of Williams’ bodily continuity theory is that he did not clarify if spatio-temporal continuity would necessitate continuity of some parts of the body or that of the whole. If whole body is required and if a limb of a certain person is amputated following an accident then there will be no continuity between the pre-accident and post-accident persons. But probably Williams would avoid this alternative, i.e. continuity of the whole body, because it leads to absurd consequences. A possible solution to this problem could be that Williams probably implied brain continuity by spatio-temporal continuity. We have already noted that even though Williams considered spatio-temporal continuity as necessary condition for personal identity, he nonetheless admits continuity of memory and character as supporting criteria. But to admit memory and character we must take cognizance of the underlying person, whose memory and character we are referring to. Again, without reference to a person’s body it is not possible to understand the causal relationship between the person and his memory etc. Here ‘body’ will essentially imply brain, as only brain is responsible for all the psychological faculties of a person. Since brain is itself a physical object hence continuity of brain will certainly mean spatio-temporal continuity of the person. Even though Williams implied spatio-temporal continuity by bodily continuity, the concept of the former has not been properly elucidated in Williams’ theory.

Probably it is because of this reason that by spatio-temporal continuity Wiggins 15 has specifically implied brain continuity. But in order to understand why he has made this implication we must first understand what Wiggins has meant by the terms ‘person’ and

‘personal identity’. According to Wiggins, there is a very close relationship between the notions of person and the criterion of personal identity. The criterion of personal identity can be derived by analysing the notion of person. Perhaps that is why an indispensible part of Wiggins’ thesis is that a criterion for the identity of persons must be “potentially analytical”. According to Wiggins’ theory in order for being a person one must not only possess a mind or mental capacity but he must also be able to remember a part of his past. Additionally, one should also be capable of being aware of one’s own identity through remembrance of past experiences. From this one should also infer that other persons, who are very similar to oneself, are also aware of their own identities. Wiggins has referred to this capacity of any person his “characteristic functioning.”16 From this it is clear that personal identity is logically connected to memory.

Let us closely examine this logical connection. If having characteristic functioning implies that person is a purely mental or non-physical being and if person does not possess any physical aspect then it must be impossible to relate the notion of spatio-temporal continuity with that of a person. As a result, we shall have to search for such a physical property as will be responsible for the memory and character of a person and whose continuity will help us ascertain the continuity, and therefore identity, of a person. The idea which we believe to be contingently linked with the notion of a person (but which at the same time is an empirical truth) is that a person possesses a body. But according to no theory the limbs of a person, for instance, are responsible for a person’s psychological faculties. This is why Wiggins has held the brain responsible for characteristic functioning of a person. Brain is that part of a person’s body which is spatially locatable, temporally continuous and the causal seat of memory, personality or character of a person. It is to be noted in this context that by brain continuity Wiggins has implied continuity of the entire brain, not that of just a part of it: “It would be better, after a conceptual analysis of the essential and characteristic vital functions, to analyse person in such a way that coincidence under the concept person logically required the continuance in one organized parcel of all that was causally sufficient and causally necessary to the continuance of essential and characteristic functioning, no

---

16 Ibid., 55.
autonomously sufficient part achieving autonomous and functionally separate existence.”\textsuperscript{17}

The advanced version of brain criterion is known as the physical criterion of personal identity. It is a known fact that human brain is divided into two hemispheres, each of which plays an important role in controlling human behaviour, linguistic and physical abilities, pattern recognition skills etc. The left hemisphere plays a major role in the control of the limbs on the right side of the body and the right hemisphere plays a major role in the control of the limbs on the left side of the body. If cerebral commissurotomy (fission of brain) occurs then according to the physical criterion what is necessary for personal identity is not continuation of both the hemispheres, but only any one of the two. For, only one half of the two hemispheres is enough to control all personal abilities as well as all psychological traits. It is worth emphasizing that each hemisphere can relearn to control all physical and psychological traits of the other if the other becomes inactive. As Noonan observes, “It might be possible one day to remove a whole hemisphere without killing the patient, the other hemisphere taking over its function, as sometimes happens when one hemisphere is incapacitated by a stroke”\textsuperscript{18}. Although here bodily criterion stands\textsuperscript{19}, but, to preserve memory, character and personality of the continuer, his body is neither relevant, nor necessary.

An objection of re-duplication may be raised against philosophers who take the stance that person will continue if only one brain-hemisphere is retained; since a single hemisphere is sufficient for keeping all mental faculties of a person intact. The objection is one of producing two continuers from a single person. If the brain of a person is bisected and each hemisphere is transplanted in a different body then the memory and character of the donor will be transferred to both recipients. As a result, whole brain transfer will be consequentially similar to the transfer of a hemisphere. But split-brain transfers give rise to a lot of peculiar situations which are mutually incoherent. E.g. either every recipient will have to be identical with the donor or each recipient must become two completely new persons. But none of these results are easily

\textsuperscript{17} Ibid., 55.
\textsuperscript{18} Noonan, \textit{Personal Identity}, 6.
\textsuperscript{19} That is to say body with half brain is enough to constitute personal identity.
comprehensible. According to Wiggins, we may also trisect or multisect a brain and there is no logical limit to this process. Rather if we take the position that the continuity of an organized physical parcel such as brain is causally necessary and sufficient for the continuity of person then understanding personal identity becomes easier and identity is portrayed as a one-one relationship.

The above version of ‘physical criterion of personal identity’ was first proposed by Wiggins in 1967 and elaborated by Parfit\(^\text{20}\) in 1984. The simplest formulation of this version is, “what is necessary for personal identity is not identity of the whole of the brain, but identity of enough of the brain to be the brain of a living person: person \(P_2\) at \(t_2\) is the same person as person \(P_1\) at \(t_1\) if and only if enough of the brain of \(P_1\) at \(t_1\) survives in \(P_2\) at \(t_2\) as the brain of a living person.”\(^\text{21}\)

The physical criterion of personal identity is also not unobjectionable. Though Shoemaker’s Brown-Brownson case and brain-hemisphere transplant case do support the physical criterion there are some thought experiments that make the physical criterion unsatisfactory. Bernard Williams imagined a thought-experiment which shows that the claim of physical criterion is wrong. Suppose, the person \(A\) is the subject of an experiment. An investigator while observing \(A\)’s internal brain functions remove some information (i.e., some memory, some personality traits) from \(A\)’s brain and kept it in a scientifically proven storage device which is theoretically equivalent to the total state of the brain, i.e. where it was originally placed. Williams imagines, “thus we can imagine the removal of the information from a brain into some storage device (the device, that is, is put into a state information-theoretically equivalent to the total state of the brain), and is then put back into the same or another brain.”\(^\text{22}\) After sometime the investigator puts back the stored information into \(A\)’s brain. According to Bernard Williams after the experiment \(A\) will treat himself as the same \(A\) before the operation. For, there is nothing to say that \(A\)’s memory is not genuine memory. If somebody loses his memory and recovers it through some storage device (suppose he relearns his past by reading his own diary) then we can say that this is not a case of genuine memory or remembering. But


\(^{22}\) Williams, B., “Are Persons Bodies” in *Problems of the Self*, 79.
what happens to $A$ is not learning again. Rather he acquires a new brain but identifies it as his own. The purpose of this sort of experiment is to show that the physical criterion does not provide any necessary condition for personal identity. For, in this case we are in doubt whether we should call $A$’s replaced brain as identical to his original brain. In the above case $A$’s identity is thus secured without continuity of the same brain or enough of the brain.

The failure of this sort of physical criterion compels us to look for an alternative view, according to which personal identity is constituted of memory or psychological elements. John Locke provides the simplest version of memory criterion.

1.3.2 Memory Continuity Criteria

John Locke introduced the notion of memory continuity as criterion of personal identity. ‘Identity’ according to Locke is the sameness of an object throughout its existence. Locke differentiates among identities of different sorts of objects (for example, identity of atoms, identity of masses, identity of vegetables, identity of trees, identity of animals etc.). He distinguishes between identity of man and identity of person. He says, “...the Identity of same Man consists; viz. in nothing but a participation of the same continued Life, by constantly fleeting Particles of Matter, in succession vitally united to the same organized Body.”\(^{23}\)

Again he says, “For I presume ’tis not the Idea of a thinking or rational Being alone, that makes the Idea of a Man in most Peoples’ Sense; but of a Body so and so shaped, joined to it; and if that be the Idea of a Man, the same successive Body not shifted all at once, must, as well as the same immaterial Spirit, go to the making of the same Man.”\(^{24}\)

On the other hand, “Person stands for;... 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


\(^{24}\)Ibid. 335.
Problems of Personal Identity: Different Philosophical Approaches

thinking, and as it seems to me essential to it: It being impossible for anyone to perceive, without perceiving, that he does perceive."²⁵

No doubt Locke’s theory of ‘man’ and ‘person’ is a radical departure from the ordinary view of man and person, where these terms are generally interchangeable. ‘Identity of man’ for him is nothing but continuation of a specific bodily form. Any psychological changes may take place in a man but that does not restrict us to call him ‘the same man’ when there is bodily continuity. Continuity of the soul or immaterial substance may be joined with the concept of man, but cannot be taken as the criterion for the identity of man; otherwise, strange result follows. For example, if we assume for the sake of argument the hypothesis of reincarnation, we should have to say that X, living in ancient Greece, was the same X as Y (i.e. X is the same man as Y), living in medieval Europe, simply because the soul was the same. But this way of speaking would be very strange; for, in case of a man’s identity, body matters. According to Locke, men of distant ages and different tempers could not be the same man, as he claims that identity of the same man consists in, “nothing but a participation of the same continued life, by constantly fleeting particles of matter, in succession vitally united to the same organized body.”²⁶

On the other hand, the defining characteristics of a person are rationality, consciousness and self-awareness (i.e., reflexive consciousness). Consciousness constitutes personal identity. From Locke’s theory it follows that ‘consciousness’ means self-awareness and memory, that is, consciousness of one’s own past selves. Consciousness constitutes the identity of a person as far back as his memory reaches²⁷. Identities of persons and identities of substances are two separate issues. Though a person may be joined with a number of spiritual substances, it does not matter whether the consciousness is annexed to one immaterial spiritual substance or a number of different spiritual substances. Consciousness is the only constitutive criterion for personal identity; consciousness unites different substances into one person. To the extent the same consciousness is

²⁵ Ibid., 335.
²⁶ Ibid., 335.
²⁷ According to Locke, if a person could not remember some of his past deeds, that is, if his consciousness does not reach up to that part of his previous life which he could not remember, which is otherwise very natural, he would not remain any more responsible for those actions.
preserved in spite of sameness or differences in substances, personal identity is preserved.

Noonan clarifies what Locke intends to imply by the term ‘memory’, which is ‘memory of events witnessed or participated in’, which can be expressed through the statement ‘I remember X’s F-ing’ (as opposed to the typical report of factual memory, ‘I remember that X F-ed’). For example, ‘I remember the event of John's attacking me’ – is a genuine memory statement. Memory of one’s own experiences and actions is a special sort of event-memory that is reported by first-person memory claim. For Locke, this sort of event or experience-memory is the criterion of personal identity as it involves self-awareness or the consciousness of one’s own self.

According to J. L. Mackie, Locke’s theory of personal identity plainly rejects two related views in this area. Locke’s distinction of ‘man’ from ‘person’ and ‘same man’ from ‘same person’ makes it sufficiently clear that he wants to banish bodily continuity from the realm of personal identity. More specifically, Locke rejects the view that persistence of human organism constitutes personal identity. At the same time, by this distinction he denies the relevance of the continuity of one immaterial, spiritual soul-substance. He does not deny the existence of spiritual substance, but, their existence or identity does not matter for personal identity. Quinborough's present mayor may possess the soul of Socrates. But until he does not possess that consciousness (i.e., self-awareness) which Socrates had of his own actions and experiences the mayor is not the same person as Socrates. A person can have a number of different spiritual substances at different times, and a single soul-substance can be joined to alternate conscious thoughts at different times. In this way Locke makes it clear that soul-substances are irrelevant for the concept of personal identity, because, the latter can be understood without any mention of soul-substance.

‘Person’ is a forensic term representing the name of the self of the morally responsible agent, who is responsible for whatever he has done in the past and which he can distinctly remember as his own action. If someone does not remember some actions of

---

28 Noonan, Personal Identity, 9.
his past, then he is a different person who has no link with the earlier one who did those actions; therefore, he cannot remain the author and is not any more responsible for those past actions (though this doctrine is contrary to the common human belief).

Contrasting this view of Locke, Sydney Shoemaker points out that forgetting some past events merely implies that memories are interrupted and therefore, full of gaps. Forgetting is an indispensible part of memory and, according to psychologists, is a necessary means for acquiring new information, but in Locke’s theory the act of forgetting creates an obstacle in establishing one’s identity. I have no way to know that whether the subject or substance of the remembered thoughts and experiences (who existed in the past) is identical with the subject or substance of the present thoughts and experiences (who exists in the present); yet I cannot doubt that all those experiences are mine. Therefore, the criterion need for establishing personal identity is different than the criterion need for establishing identity of a substance. From this it does not follow that persons are not substances or personal identity does not involve the identity of substances, but only that the requirement to establish the identity of substances (whether material or mental) and the identity of persons or selves are not same.

From Locke’s view it seems that he adopted the memory continuity criterion. Here ‘memory’ means direct first-personal experiences or consciousness of a past event. A person becomes conscious of an action if and only if he can distinctly remember that action as his own past deed and refers to it in a way as if he did that action. Though for Locke the concept of person involves certain psychological considerations, yet we strongly believe that memory cannot be the sole criterion of personal identity. Fredrick Copleston points out that though Locke does not refer to the continuity of an object’s spatio-temporal history as one of the criteria for persisting self-identity, it is undeniable that in all kinds of identity conditions there must be a continuous existence which has some relation to spatio-temporal co-ordinates. Secondly, when one claims to remember what he witnessed or experienced in the past, the only source to know whether his

---

30 To establish the identity of an object continuity of that object is required. Locke interprets identity in terms of continuity of consciousness, i.e. of memory in case of persons. Obviously no person could be entrusted to remember all his past actions and events. Hence, forgetting some past events creates obstacle to establishing self-identity.

remembering is genuine or apparent is the inner conviction of that person. The question of the degree of reliability of that person suggests that memory criterion is not self-sufficient as Locke says it is, for in order to know if it is satisfied on a given occasion we have to apply bodily criterion first. No doubt what Bernard Williams claimed for bodily criterion, till now has not been rejected on sufficient grounds.

There is a vast range of critiques against Locke’s view of personal identity. But before that we would like to discuss the views of Thomas Reid and Joseph Butler. Although both support versions of memory criterion, each of them raised some important objections against Locke’s theory.

According to Reid32, every man has a conviction of his own identity and this conviction goes at least as far back as one’s memory reaches. This conviction is inherent in human nature. No memory is possible without the conviction that if I can remember something as my own experience and action, then it implies that I existed at that time. It is a contradiction to suppose that my memory reaches a moment further back than my belief and conviction of my existence.

It is evident that the notion of identity supposes an uninterrupted continuance of existence. An object, if it is going to cease to exist at time \( t \), cannot be identical with another object that comes to exist after \( t \). Otherwise, it would be a contradiction in thought. As Reid explains, “[...] a being to exist after it ceased to exist, and to have had existence before it was produced, which are manifest contradictions.”33 According to Reid, memory is the ‘proper’ evidence of personal identity. Through memory I can identify myself with all my earlier existences. My memory testifies not only what I now remember was actually done by me, but that it was also done by the same person who now remembers it. Every person believes that everything which he can distinctly remember implies that he existed at the time remembered. Memory provides the ‘most irresistible evidence’ for the identity of every person. Beside memory there are other good evidences for personal identity. Those evidences may have the firmest assurance

---

33 Ibid., 202.
that I did such and such things, though I cannot remember those things. Reid Comments, “I know who bore me, and suckled me, but I do not remember these events.”

‘A part of a person’ or ‘part of a personality’, according to Reid, is a manifest absurdity. Part of a person does not mean bodily parts or correlated external parts like wealth, health, strength etc. Without these parts a person remains same. More clearly, if a person loses his health, wealth, strength, his personality loses nothing; if his bodily parts like arms or legs are amputated, he remains the same person he was before. Person is something indivisible such as what Leibniz called ‘monad’. Hence, Reid says, “My personal identity, therefore, implies the continued existence of that indivisible thing which I call myself. Whatever this self may be, it is something which thinks, and deliberates, and resolves, and acts, and suffers. I am not thought, I am not action, I am not feeling; I am something that thinks, and acts, and suffers.”

The grounds (or evidences) of the judgments about the identity of other persons, on the other hand, are different. Sometimes it produces the highest certainty and sometimes lowest probability. Reid does not explicitly mention what those grounds are. He only says that these grounds are sometimes similarity (i.e. similarity between appearances), and sometimes determined from a variety of circumstances. But identity from first-person-point of view is always certain.

We ascribe identity on natural and artificial bodies only for convenience of speech. Identities of external bodies are merely matters of words. “But identity, when applied to persons, has no ambiguity and admits not of degrees, or of more and less. It is the foundation of all rights and obligations, and of all accountableness; and the notion of it is fixed and precise.”

34 Ibid., 203-204.
35 Reid seems to confuse between ‘personality’ and ‘person’. In general there is a manifest distinction between the term ‘personality’ and the term ‘person’. ‘Personality’ is a psychological term that includes memory, character traits, attention, intelligence etc. in its meaning. Reid uses these two terms almost interchangeably.
38 Ibid., 206.
It is to be noted that Reid not only put forward his own views on identity and personal identity, but also commented on and criticized Locke's theory of personal identity. Locke stated, “...in this (consciousness) alone consists personal identity, i.e. the sameness of a rational Being: And as far as this consciousness can be extended backwards to any past Action or Thought, so far reaches the Identity of that Person; 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.”

Reid claims that some strange consequences follow from Locke’s thesis. First, if it is at all conceivable that consciousness can be transferred from one intelligent being to another (which is not inconsistent for Locke, because he explored the thought experiment of Prince and Cobbler, where Prince's consciousness had been transferred to cobbler) then two or twenty intelligent beings may be the same person. Here, the same consciousness is shared by different persons. Again, if an intelligent being loses his consciousness of the actions done by him and if any recovery of that consciousness is not possible then following Locke he is not the person who did those actions. Hence one intelligent being may be two or twenty different persons if he so often loses the consciousness of those former actions and is not able to restore these. Locke claims, “[...] if the same Socrates waking and sleeping do not partake of the same consciousness, Socrates waking and sleeping is not the same person. And to punish Socrates waking for what sleeping Socrates thought, and waking Socrates was never conscious of, would be no more right [...]”

Let us, for a moment, ponder over the meaning of ‘transfer of consciousness’. Read does not make it clear what he implies by that phrase. If he means to say that more than one person can be the same person at any point of time, then his objection is vacuous. What Williams said in his reduplication argument is that more than one person cannot be identical at the same time, for it breaks the principle of numerical identity; that identity must be one-one. In fact, Locke in his Essay never mentions this sort of possibility. He has only explored different examples to demonstrate the necessity of consciousness for personal identity. The example of Prince and Cobbler is not inconceivable, since

39 Locke., 335.
40 Locke., 342.
according to psychological criterion by means of brain transfer all sorts of psychological faculties can be transferred from one person to another (for example, Shoemaker’s Brown-Brownson case). But this is possible between only two persons, not more than that.

Reid’s next objection against Locke's theory is that a man may be and at the same time not be the person who did a particular action. Reid imagines a thought-experiment to establish his point. “Suppose a brave officer to have been flogged when a boy at school, for robbing an orchard, to have taken a standard from the enemy in his first campaign, and to have been made a general in advanced life: suppose also, which must be admitted to be possible, that, when he took the standard, he was conscious of his having been flogged at school, and that when made a general he was conscious of his taking the standard, but had absolutely lost the consciousness of his flogging.”

Now from Locke's theory it follows that the boy who was flogged at school, and the young officer who had taken a standard, are one and same, because the officer’s consciousness can reach up to the boy’s experience. For the same reason the young officer is the same person with the general. By the relation of transitivity it follows that the general and the boy are identical, but, since the general’s consciousness could not reach up to the boy’s experience, then according to Locke’s thesis the general and the boy are different persons. Logically and empirically this is absurd (for it breaks the rule of transitivity). The obvious though peculiar consequence which follows from Locke's theory is that the general is and at the same time is not the same person with the boy who was flogged at school.

This objection rests totally on third-person viewpoint, while Locke wanted to emphasize on first person perspective of identity. That a statement is logically valid does not imply that it is psychologically acceptable (or valid) too. Locke actually offers a psychological criterion and not a logical one. This objection actually tends to imply that there can be two types of views regarding identity: one is extensional, and the other is intentional.

---

42 In reply to this it can be said that though Locke emphasizes first-person view of personal identity, sometimes in his thesis he comments on third-person identity. A person is identical as far back as his consciousness reaches and not accountable for those actions which he could not
In reply to this criticism Sydney Shoemaker and Derek Parfit offer different solutions in favour of Locke's theory. Let us consider Shoemaker's reply first.

Shoemaker\(^{43}\) points out that if Reid, by the phrase, “absolutely lost consciousness of the flogging” means that the concerned memory was lost without any possibility of recall, then, the only way to defend the Lockean theory is to modify it. Shoemaker conveniently constructs this revision in person-stage terminology. Two successive person-stages belong to the same person, if the later person-stage contains the memories which comprise the preceding stage. This revision allows for the fact that the current person-stage contains all those memories as long as one has the potentiality\(^{44}\) of remembering it. Two person-stages that are so related are ‘memory-connected’. “This comes to saying that two stages belong to the same person if and only if they are the end-points of a series of stages such that each member of the series is memory-connected with the preceding member.”\(^{45}\) One such series consists of the stages of the boy at the time of his flogging, young officer’s stage at the time of his valiant deed, and the old general’s stage at the time at which he remembers the valiant deed but not the flogging. The significance of this revision is one’s memory continuity with one’s past selves – “memory continuity consisting in the occurrence of a chain of memory-connected person-stages of the sort just described.”\(^{46}\)

Derek Parfit\(^{47}\) offers almost a similar solution against the objection in terms of the chain of overlapping memories. For example, if I can remember the experiences that a certain person had yesterday, and that person was able to remember the experiences which another person had the day before yesterday, then by the chain of overlapping memories distinctly remember. The term ‘accountability’ brings the concept of third-person identity. To myself, doubtless I am accountable for those actions which I remember, but to others I am accountable for all those which I did (whether I may or may not remember these). When the concept of accountability is conjoined with the notion of identity, it naturally brings the third-person viewpoint with it.


\(^{44}\) Precisely, if I forget some of my earlier memories but if there is any possibility of remembering it, that is, if I have the potentiality to re-discover it, then all those memories are contained in my current person-stage.

\(^{45}\) Ibid., 81.

\(^{46}\) Ibid., 81.

I am identical with the person who had those experiences the day before yesterday, even though I cannot now remember any of the latter experiences. Philosophers are in doubt whether Locke himself would have welcomed any one of these proposed revisions. But, undoubtedly these revisions of Locke’s account of personal identity does not suffer from the apparent difficulty of self-contradiction.

Joseph Butler, another supporter of the memory criterion argues that the identity or sameness of a person is implied in the notion of my living now and hereafter or in any two successive moments. He agrees with Locke that ‘consciousness of what is past’ ascertains our personal identity. But, if consciousness makes personal identity, then, Butler comments, “[...] a person has not existed a single moment, nor done one action, but what he can remember; indeed none but what he reflects upon.” Precisely, consciousness is a kind of remembrance through which I can reflect on my whole past as an identical self. Therefore, it is self-evident that consciousness of personal identity presupposes personal identity and hence cannot constitute criterion for personal identity or cannot be used to explain it.

In other words, according to Butler, there is a fallacy of circularity in Locke’s theory. Locke defines personal identity in terms of consciousness, i.e., memory, but, memory itself presupposes the identity of a person. ‘I can remember some of my past experiences and actions’ is a memory statement which presupposes that the person to whom we attribute memory and the person who had those experiences are one and same. Clearly, whenever I say that I can memorize some of my past actions, I also want to say that I am that person who did those actions, otherwise I could not remember. Particularly, Butler’s point was that the precondition for attributing genuine memory of some past experiences is that the person to whom one is attributing the memory was indeed the person who had the experience.

Lowe points out that Parfit, who is a neo-Lockean revisionist, seeks to evade this apparent difficulty of circularity. To do this Parfit brings in the concept of quasi-memory.

---

49 Ibid., 167.
(q-memory\textsuperscript{51}) which is supposed to be just like ordinary memory, but without the implication that “it cannot in principle be attributed to one person, in respect of experiences that were undergone by another.”\textsuperscript{52} Hence, by using the concept of ‘quasi-memory’ we can overcome the objection of circularity at ease.

Lowe raises some doubt as to whether Locke would accept those new revisions of his theory, since Locke was not a ‘technical’ philosopher and might not have been sympathetic to other philosophers who tried to overcome Locke’s difficulties. Lowe anticipates that Locke’s own response to these objections would have been to stick to his original account and to deny the force of these objections. As Lowe comments, “Both Butler and Reid held human persons to be genuine substances in the strictest sense of the term and, given that assumption, their memory based objections may well have some force. But, as I have tried to make clear, it seems that Locke himself did not share that assumption, and in its absence, it is not so clear that such objections are compelling.”\textsuperscript{53}

Let us return to Butler. For him the notion of identity of vegetables is utterly different from that of personal identity. In case of vegetables or pure physical matters (suppose the identity of a tree fifty years old is concerned here) the word ‘same’ means ‘same as to all the purposes of property and uses of common life’. In case of persons the identity is strict and philosophical. Therefore, if we say that a fifty years old tree is identical with all its past existences, and if here ‘identity’ means “strict and philosophical sense of identity”, then it will be a contradiction in terms. Only in the loose and popular sense identity of life forms and that of organizations may be understood. In case of persons identity is ‘philosophical’, because personal identity cannot consist in diversity of substances. On the other hand, the identities of physical objects or life forms and organizations can persist through the perpetual changes of parts and diversity of substances.

\textsuperscript{51} According to Parfit, “I am q-remembering an experience if (1) I have a belief about a past experience which seems in itself like a memory belief, (2) someone did have such an experience, and (3) my belief is dependent upon this experience in the same way (whatever that is) in which a memory of an experience is dependent upon it.” [Derek Parfit, “Personal Identity.” The Philosophical Review 80/1 (1971): 15.]

\textsuperscript{52} Lowe, Locke, 96.

\textsuperscript{53} Ibid., 96.
Let us examine more closely what Butler meant by the term ‘strict and philosophical’. In Butler’s own words, “[…] in a strict and philosophical manner of speech, no man, no being, no mode of being, no anything, can be the same with that, with which it hath indeed nothing the same. Now sameness is used in this latter sense, when applied to persons.” It seems that according to Butler, “strict and philosophical sense of identity” means identity of substance and therefore identity of properties of substance. In other objects it could not be guaranteed that they are linked with one and only one substance; it may be the case that the substance $S$ that is linked with the object $O_1$ at time $t_1$ altered at $t_2$. If an object is linked with diverse substances, it is also a substratum of diverse properties, for, “the same property cannot be transferred from one substance to another”. According to Butler, Locke and his followers hold that consciousness is not permanent, but successive and momentary. It is self-evident that one’s personality is also successive and momentary as consciousness constitutes personality or personality consists in consciousness. From this it necessarily follows that we cannot charge our present selves for anything we did in the past, since our present self is not the ‘same’ with the self of yesterday. Though these theorists use the word ‘same’ or ‘identity’, it may appear that what we have just said is not an accurate interpretation of their thesis. Nevertheless according to Butler, “they cannot, consistently with themselves, mean, that the person is really the same […] but only that he is so in a fictitious sense.” Therefore the Lockean concept of personal identity employs the term ‘identity’ only in a fictitious sense. Butler argues, “The bare unfolding this notion, and laying it thus naked and open, seems the best confutation of it.”

In response to Butler’s criticism it can be replied that Locke never used the term ‘identity’ and ‘sameness’ in a fictitious sense. Rather he uses it in a strict and philosophical manner. Similar to Reid and Butler, Locke also distinguishes persons from all other animate and inanimate beings. Butler’s criticism also goes against the views of David Hume, which we shall consider later on.

55 Ibid., 170.
56 Ibid., 170.
Further, Butler argues that every person is conscious of the fact that he is identical with all his earlier existences that he can distinctly remember. I am certain of the fact that I did some actions in the past and at present when I reflect upon or become conscious of those actions, I can conceive myself as identical to the person who did those actions. That is to say, I know that the person who did those actions and the person who is now reflecting upon it are same. ‘Those actions actually happened’ – the assurance of these facts arises wholly from consciousness that ‘I myself did it’.

In opposition to Locke, Butler prominently opines that all beings “confessedly” continue to be the same during the whole span of their existences. All the actions, enjoyments, and sufferings of a living being are actions, enjoyments and sufferings of the same living being, since remembering or forgetting can make no alterations in the truth of past matters or facts. Despite all these facts a living being can know itself as the same living being. ‘I am conscious of some of my past actions’ means I can distinctly remember those actions as if now I am experiencing them. Reid also interprets ‘consciousness’ in Locke’s theory as memory. He points out that the meaning of the term consciousness is not clear from Locke’s theory. According to Reid, if Locke wants to imply by the term consciousness a kind of memory or remembrance then his theory is intelligible but if not then Locke’s theory of personal identity needs more clarity.

Reid argues that memory is the only faculty by which we have an immediate knowledge of our own past actions. In ordinary language there is no fixed line of difference between consciousness and memory. As he says, “it is unnecessary, in common discourse, to fix accurately the limits between consciousness and memory.”57 Remembrance is sometimes called ‘sense’ sometimes ‘consciousness’ without any inconvenience. But in philosophy there should be a sharp distinction between these two. Reid specifies the distinction. That is, consciousness is an immediate knowledge of the present, while memory is an immediate knowledge of the past. Therefore, in Locke's theory consciousness is confounded with memory and more strangely the criterion of personal identity is confounded with evidence. For Reid, memory cannot constitute criterion of personal identity; it only provides the requisite evidence for personal identity.

Reid’s last objection against Locke is, however, not entirely appropriate; for Reid himself clarifies two meanings of the term ‘consciousness’. He explicitly mentions that in popular discourse the term ‘consciousness’ is often used as a synonym for the term ‘remembrance’. Only in the discourse of philosophy the meaning of the term ‘consciousness’ is just the opposite of the meaning of the term ‘remembrance’. Clearly Locke did not confound consciousness with memory, rather he only followed common language.

Butler and Reid, contrary to Locke, opine that consciousness or memory is not the criterion of personal identity; rather, it provides the evidence. As Reid said, “memory gives the most irresistible evidence of my being the identical person that did such a thing, at such a time [...] This remembrance makes me to know assuredly that I did it.”

Memory, according to them, is an important factor which reassures their convictions of their own identities. Nowhere in their discussion they use the word ‘criteria’, but use the word ‘evidence’. Consciousness could not be the criterion of personal identity, as consciousness could not be identical in any two moments. Successive consciousnesses are not same, but only of same kind.

The views of David Hume are considered as one of the greatest contributions in the literature on personal identity. Hume’s notion of person and identity is just the contrary of the views we have discussed so far, because he analyses the concept of personal identity in rigid empirical terms. Hume denies the notion of identity and the notion of person altogether. According to Hume, every idea invariably follows from an impression that precedes the idea. The term ‘person’ does not evoke any such impression in one’s mind and consequently, there are no such ideas. Mind or self is “nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement.”

Identity or sameness and diversity are two different ideas, though they have some resemblance between them for which they are generally confounded with each other in our common way of thinking.

58 Ibid., 203-204.
Mistakenly, we fall into it before we are aware. Identity or sameness according to Hume is “a distinct idea of an object” that remains invariable and uninterrupted through a supposed variation of time. Diversity is “a distinct idea of several different objects existing in succession, and connected together by a close relation.” These two ideas are distinct and contrary, but they are generally confounded with each other by our common way of thinking. It is our imagination that creates an uninterrupted and invariable object by reflecting on succession of related distinct objects. Because of the resemblance among those distinct objects, the relation (that holds between related objects) “facilitates the transition of the mind from one object to another, and renders its passage as smooth as if it contemplated one continu’d object.” There is no invariable object but succession of parts connected together by resemblance, contiguity and causation. Hume comments, “For as such a succession answers evidently to our notion of diversity, it can only be by mistake we ascribe to it an identity.”

Hume explicitly differentiates between numerical and specific identity. An object always remains numerically identical with itself \((a = a)\). But when the parts of an object are going through a continuous process of change we call it specific identity, because the parts generally serve a common end, and at any one point of time we do not have the idea of difference or multiplicity; every object \((P_1)\) is in a manner annihilated before the second \((P_2)\) comes into existence. Identity is a kind of quality and it is our ‘uninterrupted progress of thought which constitutes the imperfect identity’ instead of diversity by overlooking the gradual and insensible changes (however rapid these may be) of that object.

Let us try to make the last point clearer. Hume differentiates between numerical and specific identity. An object always remains numerically identical with itself. By the term ‘specific identity’ Hume actually means qualitative similarity. When the parts of an object are going through a continuous process of change he calls it specific identity. For example, hearing a sound does not mean that the sound which originated from the source and the sound that reaches the listener are numerically identical. Sounds are frequently

---

60 Ibid., 253.
61 Ibid., 254.
62 Ibid., 255.
interrupted and renewed and it is evident that sounds have only specific identity. But whenever we hear a sound we do not have the idea of difference or multiplicity, we have only the idea of identity.

According to Hume, a considerable change in any part of a body can destroy its identity. To know the identity of an object one should know what changes have occurred within that object. If any part of that object has been altered one has to know the proportion of the part in respect to the whole. Hume does not explicitly mention the criteria of personal identity, but from his discussion it appears that it must be psychological. As he said, “since this interruption makes an object cease to appear the same, it must be the uninterrupted progress of the thought, which constitutes the imperfect identity.”

Hume holds that personal identity does not consist in the identity of a substance, but consist in some relationship (i.e., relation of resemblance) between present experience of a person and the earlier ones. Resemblance connects different perceptions (earlier and later) into one thread, so that we are able to build the idea of one continuing object in our imagination. But according to Sydney Shoemaker this concept is mistaken. If a present experience does stand in a particular relation with a past experience, then it need not always be the case that the past experience equally stands in relation (resemblance) with the present one. But if is the ‘observable empirical relationship of co-personality’, and if my past experience does not stand in relation with my present one, then these experiences are not my experiences. But this is absurd. Because, in spite of the fact that I do not perceive any empirical relationship between my remembered experience and the present one, nevertheless, my past experiences still remains my experiences. Hume tries to make the self empirically knowable, but he did not clarify how it was possible.

Resemblance, contiguity and causation are three ‘natural relations’ that can produce a union in the imagination. According to Hume, the relation of contiguity is irrelevant for our present purpose. Our attributions of identity to different related perceptions results from the easy transition of the imagination from one perception to another. Since, resemblance and causation are only two relations that can facilitate such a transition,

---

63 Ibid., 256.
these two relations are enough to produce in us the ‘fiction’ or mistake of a continuously existing self or mind.

Memory is an important factor for discerning personal identity, because it alone acquaints us with the continuance and extent of the succession of related perceptions. Hume claims, “Had we no memory, we never shou’d have any notion of causation, nor consequently of that chain of causes and effects, which constitutes our self or person. But having once acquird this notion of causation from the memory, we can extend the same chain of causes, and consequently the identity of our persons beyond our memory, and can comprehend times, and circumstances, and actions, which we have entirely forgot, but suppose in general to have existed.”

The problem of memory is that I may not remember some of my actions and thoughts of 2nd January 2009 or of 24th December 2005 etc.; but I cannot deny that my present self is identical with the self of that time. So, memory does not so much produce as discover personal identity by only showing us the relation of cause and effect among different perceptions.

However, memory also provides us the way to access to our past selves and thus gives us a sense of our endurance through time. We think of our selves as one enduring thing partly because, we can remember. Remembering is a kind of perception that occurs in the mind. To remember is to have a perception, which represents, and therefore resembles, the past perception it is a memory of. Resemblance is one of the two relations that leads the imagination to slide more easily from one member of the series to another in diverse but similar, transient objects, and hence to think of it as a continued view of the same object.

Barry Stroud, a renowned commentator of Hume points out that it is doubtful whether the relation of resemblance plays such a good role in constituting a person’s mind. For example, suppose we consider a bundle of perceptions which is composed only of perceptions of the Eiffel Tower. Suppose views of it are taken from a particular place on the bank of the Siene. Following Hume, here every perception resembles every other, in

---

64 Ibid., 261-262.

65 Hume explains, “The passage of the thought from the object before the change to the object after it, is so smooth and easy, that we scarce perceive the transition, and are apt to imagine, that ’tis nothing but a continu’d survey of the same object”. (Ibid., 256).

being about the same thing. Obviously each of these perceptions has resemblance to every other for all are perceptions of the same thing. Since the perceptions of the bundle belong to different people we are not for a moment inclined to think of that bundle as constituting one mind. Therefore, resemblance alone is not sufficient. Stroud comments, “[...] so even if resemblance can and does have an effect on us, it might be that the perceptions must resemble each other, because they all occur in one person’s memory before we come to think of them as constituting one mind. And then the idea of one person’s memory would be doing all the work. But that, as we saw, could not be what we are going on when we get the idea of a single mind in the first place, since it uses the very notion of one person’s memory.”67

It may appear that the problem of artificially constructed bundles of perceptions which we do not regard as constituting persons can be overcome by the relation of causality. If each of the members of a bundle is effect of its previous member and cause of the succeeding one, then imagination slides easily along the causal chain and comes to think of them all as one mind. But in case of Hume’s account of causality, it is not as obvious as it may appear.

To explain this Stroud gives another opposite example. Suppose there were long standing regularities among perceptions of various kinds, so that whenever one of the A-sort appeared, then one of the B-sort appeared and when one of the B-sort appeared then one of the C-sort appeared and so on. Following Hume that would imply that a causal chain held between particular perceptions like A, B, C and so on, belonging to those kinds. Moreover, this causal chain also holds when A occurs in my mind, B occurs in your mind, and C occurs in some other person’s mind. In other words, there is a constant conjunction between various types of perceptions regardless of which mind those perceptions belong to. So, in that bundle the effect of my perception is your perception, and the effect of your perception is someone else’s perception. In considering such a bundle, we would not find ourselves inclined to regard them all as constituting one mind. But, Hume does not commit this answer because he would regard such bundles as individual mind, since all their members are related causally, and as a natural relation

67Ibid., 124.
that facilitates the mind’s passage along the series, thus leading us to think of the bundle as one mind.

Here the only satisfactory thing is that the above far-fetched situations do not happen in our actual world! The bundles of perceptions which constitute our minds are not such that one of its members is my perception, one is yours, and the other one is somebody else’s which are causally linked and which mutually produce, destroy, influence and modify each other.

Stroud does not want to say that there are no causal connections among all our perceptions. Certainly there are. The relation between impression and idea is causal. But this sort of causality does not help Hume to solve the problem of the identity ascriptions to ourselves. Because, according to Stroud, “Those causal connections run ‘vertically’, so to speak, from the impression up to the idea, and then perhaps to other ideas and impressions. What Hume needs a causal chain that runs ‘horizontally’, as it were, along the whole series of incoming perceptions that we get from moment to moment. That is what I am arguing does not exist.”\textsuperscript{68} In two successive continuing moments I may have two impressions successively – one is the impression of a tree, and the other one is the impression of a building; but the first is not the cause of the second. The first does not belong to a class of perceptions each of which has been followed by a member of a class of perceptions to which the second impression belongs. There are no such regularities in our impressions of sensations.

From Stroud’s criticism it follows that the relation of resemblance and causality are not sufficient to explain the idea of an individual mind or self that endures through time. It is not easy to say whether Hume was conscious about this difficulty. Though he found his theory to be defective, but did not clearly explain what the defect was. He was particularly dissatisfied with his principles of resemblance and causality; because these principles are not enough to explain the origin of our ideas of the self or mind. Hence, reviewing his former strict opinions of personal identity, Hume confessed that he found himself in such a labyrinth that he neither knew how to correct his former opinions nor how to render them consistent. In Hume’s own words, “But upon a more strict review of

\textsuperscript{68} Ibid., 126.
the section concerning personal identity, I find myself involv’d in such a labyrinth, that, I must confess, I neither know how to correct my former opinions, nor how to render them consistent. If this be not a good general reason for skepticism, ’tis at least a sufficient one (if I were not already abundantly supplied) for me to entertain a diffidence and modesty in all my decisions.”

It can be said that similar to Locke, Hume is also a memory theorist. The only difference is while Locke thinks that ascribing identity to persons or other objects is something ‘real’, for Hume it is utterly fictitious. Hume does not mean to say that identity is just a matter of words, confined only in the expressions of language. He knows that in the ordinary way of thinking the ascription of identity is very natural. His point is that by thorough investigation we find that there is only the notion of diversity in both external and internal experiential world, not any notion of identity.

So far we have discussed different classical philosophical views regarding personal identity that somehow supports memory as a criterion or as evidence of identity. Locke is a prominent memory theorist and declared memory as a criterion of personal identity. Butler, Reid and Hume, though did not support memory as criterion, but invariably accepted that memory is one of the strongest evidences of personal identity. No doubt memory theory also has some shortcomings. As Williams claimed, memory could not establish personal identity as one-one relation. And it faces the famous circularity objection. Now we shall explore whether advanced versions of memory theory could overcome those difficulties.

1.3.3 Advanced Psychological Continuity Criteria

First, we would like to discuss Richard Swinburne’s theory of personal identity. He uses psychological continuity criterion, which is an advanced version of simple memory theory.

70 Locke never mentioned the term ‘real’. But he never said it was ‘unreal’.
One of the greatest objections against any memory theory is Williams’ reduplication argument, which has been discussed earlier. According to some philosophers one way to overcome the reduplication argument is to adopt some sort of bodily continuity. For them bodily continuity means brain continuity which is the core of the body and which is causally responsible for all types of memory and character continuity. Only the continuity of brain constitutes personal identity. But according to Swinburne, the brain continuity criterion cannot avoid the duplication objection. It is possible that there are more than one later persons who satisfy memory criterion or brain criterion (e.g. Williams’ Guy Fawkes case or Wiggins’ Fission case) or a combination thereof for being the same as an earlier person.

Swinburne’s version of memory criterion is a slightly advanced version of psychological continuity criterion. This criterion claims that for defining personal identity, in addition to experience-memory, other psychological facts should also be taken into account. For example, there exist direct psychological connections between intentions and the acts later carried out by those intentions. Such direct psychological connections can link childhood experiences with adult character traits, fears, prejudices etc. These connections are important constituents of personal identity. The definition of this criterion is that $P_2$ at $t_2$ is the same person as $P_1$ at $t_1$ if and only if $P_2$ at $t_2$ is psychologically continuous with $P_1$ at $t_1$.

According to Swinburne, to define personal identity we first need to find out the essence of a person. He shows that mere bodily matter (i.e., a particular chunk of flesh) is not essential for a person, not even certain apparent memories. Something substantial is necessary to define personal identity. He cites Aristotle, “[…] a substance at one time is the same substance as a substance at an earlier time if and only if the later substance has the same form as and continuity of matter with, the earlier substance.”72 What is necessary for personal identity, according to Aristotle, is therefore continuity of the same form and continuity of the matter.

72 According to Swinburne, form and matter are two essential properties of a substance. ‘Matter’ is the thing out of which objects are made, and ‘the form’ means those properties the possession of which is essential if a substance is to be the substance in question. Ibid., 5, 26.
Admission of the possibility of the existence of ‘disembodied persons’, however, creates the possibility of existence of a person without any particular body or particular chunk of matter. The possibility of ‘disembodied persons’, if not contradictory, serves to demonstrate that without any bodily matter two persons can be identical. What natural law can determine is that every person has a certain body having a certain constitution. But it could not determine which body is mine or which is yours, i.e., it is not able to show any necessary connection between my body and myself. As Swinburne says, “Just the same arrangement of matter and just the same laws could have given to me the body (and so the apparent memories) which are now yours, and to you the body (and so the apparent memories) which are now mine.”73

The possibility of disembodied existence indicates the insufficiency of Aristotelian account of identity of substance in case of persons. The presence of some mental capacities (for example, thoughts, sensations, conscious experiences, ability to perform intentional actions) is required for being a person, which is not required for being a table or a tree. It seems either we have to admit that personal identity is distinct from other identities or we have to adopt a more general account than Aristotelian account of identity of substances, because to identify a person with his earlier existences only form and matter is not sufficient. For this reason, Swinburne formulates a “wider Aristotelian account of the identity of substances”. That is, “[...] two substances are the same if and only if they have the same form and there is continuity of the stuff of which they are made, and allow that there may be kinds of stuff other than matter.”74

“Stuff other than matter” – means a kind of immaterial soul-stuff. A person living on earth consists of two parts – material bodily part and immaterial soul-part. Soul75 is the essential part of a person and its continuity is essential for continuity of the same person. This is the main thesis of Swinburne’s ‘simple theory’.

74 Ibid., 27.
75 According to Swinburne, soul is the essential part of a man, which enjoys the mental life which is the subject of sensation and thought and the originator of actions. It is the necessary core of myself which must continue if I am to continue, i.e. which is necessary for my continuing existence. A person is the soul together with whatever, if any, body linked temporally to it.
Simple theory admits the existence of disembodied persons. Though in general persons are constituted of both material and immaterial stuff, the immaterial stuff is the essential core for being a person. This sort of view is also proposed by the classical dualists, viz., Plato and Descartes. According to them each and every person is indivisible, because the essential core (i.e., soul) of each person is indivisible. It is inconceivable and therefore impossible that each of the two later persons is partly identical with an earlier one, because persons are not identical with pure material objects and any chunk of matter, no matter how small it is, is divisible (at least logically divisible), whereas, the soul-stuff has the property of indivisibility.

Swinburne admits that the dualistic version of simple theory actually originated in Descartes. An actual fact follows from the mere logical possibility of my continued existence, that is: “[...] there is now more to me than my body; and that more is the essential part of myself. A person being conscious is thus to be analyzed as an immaterial core of himself, his soul being conscious.”

Swinburne claims that his arguments in favour of the simple theory involve Descartes’ argument in support of the person’s disembodied existence. Descartes described certain possible cases of thought experiments where the person continued without having the formal body. Swinburne explores some similar thought-experiments which imply that the presence of body is not logically necessary for the existence of a person. From the wider Aristotelian account it follows that the continuation of a substance involves the continuation of some of the stuff of which it is made and because my continuing existence does not necessarily involve my bodily matter, there must be something that continues as part of me. This is the (Cartesian) soul and it forms the essential part of myself.

Now the fact that a person has an immaterial soul-stuff does not mean that by acute observation through a powerful microscope we can realize its existence. Rather, “it is just a way of expressing the point within a traditional framework of thought that persons can – it is logically possible – continue, when their bodies do not.” Nevertheless, any

---

76 Ibid., 30.
description of a person or a living being is incomplete without the reference of his material body. The main point is that my body, though part of me, is only contingently and temporally related to me.

Personal identity is unanalysable in terms of brain continuity and apparent memory, though they provide only the flexible evidence for personal identity. Normally, when we comment, ‘this is the man whom I saw last week’, we do this only on the basis of similarity of visual appearance. Similarity of appearance covers not only instantaneous appearance, but appearances over time in the way in which the person manifests his presence – the way he walks, and the gestures he makes. Similarity of appearances (though it allows for gradual changes of facial and bodily structure) is the indirect evidence of personal identity. For it is the direct evidence of bodily identity which in turn is the indirect evidence of personal identity. Bodily identity, again, is the indirect evidence of brain identity (because brain transplants do not yet occur in reality) which in turn is the evidence of personal identity.

Brain continuity is important because its continuity guarantees continuity of apparent memories and character.

Swinburne argues that in absence of counter examples apparent memories should be taken at their face value as evidence of that of which they are apparent memories. This follows from the principle of credulity, which is the most basic principle for making inferences from worldly experiences. “That states that probably things are as (in the epistemic sense) they seem to be – e.g., if it seems to me that there is in front of me a brown table, or a Greek vase then probably there is; and I ought so to believe, unless counter-evidence turns up. My apparent memory that I did or experienced so-and-so is just it seeming to me that I did or experienced so-and-so; that I am justified in believing my apparent memories in the absence of counter evidence follows from the principle of credulity.” Principle of credulity is the only way to justify our beliefs about the external world. Otherwise, if we think that we require something more in addition to the principle of credulity, we do not have a normal belief system.

All access to the past depends ultimately on apparent personal memory. My personal memory is the key to knowing what I did and experienced personally. The principle of

credulity\textsuperscript{79} is essential to justify all my past deeds and experiences. In absence of any counter-example we should rely upon apparent personal memory because as a criterion of personal identity apparent personal memory is justified a priori.

In some cases brain continuity provides a strong evidence for personal identity. Nonetheless, brain continuity is only an indirect evidence of personal identity, for it is actually the evidence that there must be apparent personal memory, which in turn is the direct evidence of personal identity. In brain-transplant cases continuity of the brain matter provides the assurance of the continuity of apparent personal memory and vice versa.

Apparent personal memory is accompanied by sameness of character. If \(P_2\) has the same attitudes, habits, reactions towards the world as \(P_1\) then it is evident that \(P_2\) is \(P_1\). Sometimes, what a person apparently remembers is actually experienced by some other person (such as ‘clairvoyance’ à la Williams), and if so, then that apparent memory is subject to correction, and hence, not reliable. Otherwise, I have every justification for relying on my memory. Whether my apparent memory is infallible or not, can be checked in two ways. Either I have to make public memory claims which have been observed by persons other than me, or in case of my privately accessible experiences only the presence of my body is necessary. In the latter case the continuation of my body from time \(t_1\) to \(t_2\) provides sufficient evidence that I am that person who makes such and such private memory claims. But according to Swinburne, apparent personal memory can be cross checked by apparent memory, and without any reliance on any bodily continuation. If, for instance, a person \(P_3\) at \(t_3\) remembers the acts of \(P_2\) at an earlier time \(t_2\) and of \(P_1\) at a still earlier time \(t_1\), then the evidence that \(P_3\) is identical with both \(P_2\) and \(P_1\) may be secured by verifying whether \(P_3\) is capable of remembering the apparent memories of \(P_2\). This can happen if \(P_3\) and \(P_2\) both made public memory claims at times \(t_3\) and \(t_2\) respectively about \(P_1\) at the time \(t_1\). According to Swinburne, “They may do this by being embodied and speaking through mouths – and this can happen and be seen to

\textsuperscript{79} Even according to Swinburne, my experience about my identity in a very short period is justified by the principle of credulity. This experience is known as 'apparent experiences'. He gives the example, "It seeming to me now epistemically that I am the common subject of such-and-such successive or simultaneous experience." Here 'seeming' is (apparent) experience, because it is not inference based, and though reliable, but fallible.
Problems of Personal Identity: Different Philosophical Approaches

happen without our needing to take for granted that they are embodied in the same bodies.”\footnote{Ibid., 62.} The primacy of apparent memory as a criterion of personal identity is thus vindicated.

From Swinburne’s version of apparent memory it follows that a person has to make public memory claims to verify his apparent personal memory by means of his other apparent personal memory. But public memory claims cannot be made unless one is bodily present in front of the public. Clearly without the help of a body an apparent personal memory cannot be verified by another apparent personal memory. Let us see what Shoemaker’s position is in this respect.

Sydney Shoemaker\footnote{Sydney Shoemaker, and R. G. Swinburne, \textit{Personal identity} (Oxford: Basil Blackwell, 1984).}, another memory theorist, distinguishes between the words ‘person’ and ‘man’. Similar to Locke, he holds that every person occupies a special role in the world of beings, for persons can think, argue and judge; whereas the term ‘man’ only denotes a certain biological species. Shoemaker anticipates various objections against memory theory and tries to overcome them as far as possible. Among these objections one of the most direct objections is that one does survive in a state of total amnesia. If it happens, then there is no chain of memory-connected person stages, and it counts against any kind of simple or revised memory theory. Here the term ‘amnesia’ does not mean loss of memory that can be recovered. Here ‘amnesia’ means the possibility of ‘philosophical amnesia’ i.e., the irreversible loss of all memory of the past.\footnote{‘Irreversible loss of all memories’ means loss of all memories of past experiences and action. After a certain time the person cannot remember any experiences before that time. But ‘irreversible’ loss does not mean irretrievable loss of memories of all kinds, because, if this kind of amnesia happens then the most probable outcome is that the \textit{person} will no longer exist, as persons are self-conscious beings and aware of their own identities through memory.}

In this problematic situation Shoemaker invokes the notion of psychological continuity. To do this he cites his famous thought-experiment of Brown-Robinson-Brownson case, where the brain-surgeon’s assistant mistakenly put Robinson’s brain into Brown’s head and Brown’s brain into Robinson’s head. One of the two resulting persons (with Brown’s body and Robinson’s head) does not survive. The other one with Robinson’s
Problems of Personal Identity: Different Philosophical Approaches

body and Brown’s head survived and is called Brownson post operation. According to Shoemaker, the following argument is called the *third-person version of the change of body argument*. Brownson’s memories are ‘memories from the inside’ of Brown’s past life, for he possesses Brown’s brain. If so, then it is very much possible that in all psychological respects Brownson resembles Brown, i.e. Brownson’s interests, tastes, talents, personality, and character traits all are similar to Brown’s. If something happen otherwise, i.e., if Brownson remembers from the inside Brown’s past, but if Brownson’s other psychological traits are similar to Robinson’s, then we can hardly attempt to identify Brownson with Brown.

From the fact of having Brown’s brain in Brownson’s head, we assume that there is a causal relationship or a counter-factual dependence between Brown’s psychological traits (before the operation) and Brownson’s psychological traits (after the operation). What Shoemaker wants to claim is that the hierarchical status of memory and the status of all other psychological traits are almost same. If the circumstances are such that evidence of similarity of all psychological traits between two person-stages is evidence of such a causal or counter-factual dependence then that evidence of similarity is also the evidence of identity; just as Brownson’s memories provide evidence that he is Brown. Shoemaker comments, “Thus the status of similarity and continuity of personality traits as evidence of personal identity seems no different than that of memory continuity; both are evidence only in so far as they include, or are evidences for, causal relations between earlier and later states.”

Henceforth, by ‘psychological continuity’ Shoemaker understands both types of causally grounded continuities – memory continuity and the continuity of other psychological traits. Two person-stages (which are temporally distinct) can be psychologically directly connected (if there are appropriate causal relation) and can be stages of one and the same person if and only if, (1) they are connected through a series of stages, where each and every member of the series is psychologically directly connected with the immediately preceding member of that series (i.e., there is appropriate relation of causal dependence between the earlier stage and the later stage of the same person), and (2) there is no such

---

83 Ibid., 90.
series that contains a member which is psychologically directly connected to two different stages occurring at the same time. Identity of persons consists in this kind of psychological continuity, which involves the relation of causal dependence between two successive stages of the same person. Shoemaker calls this the “causal continuity account” of personal identity.

According to Shoemaker, the phrase ‘unity of consciousness’ is ambiguous. It refers either to the unity of conscious states or consciousness of a variety of mental states as unified into a single consciousness, namely one’s own. Shoemaker, in his well-known treatise *Self-Knowledge and Self-Identity*, points out that “third-person version of the change of body argument” apparently demonstrates that personal identity does not imply bodily identity and that criteria of personal identity include psychological criteria. This view is compatible with philosophical behaviourism, for behaviourism holds that all psychological states can be analysed in terms of behavioural states. Thus, the behaviour of Brownson inclined us to say that he is Brown. Behaviourism admits that identities of persons are distinct from identities of human bodies or other material objects (because bodily-identity involves spatio-temporal continuity). But as a materialist theorist, a behaviourist could maintain at the same time that “all the properties of persons can be regarded as physical properties (behavioural dispositions and so forth) and that there is therefore an important sense in which persons are material objects.”

The main point is that Shoemaker wants to differentiate persons from the Cartesian idea where persons are essentially non-physical entities.

The third-person version of the change of body argument, according to Shoemaker, is a borderline case of personal identity, for nobody knows whether Brownson is Brown. Some may want to call him Brown, others Brownson. No factual consideration can decide the answer. What happens here is just a dispute about words. To solve it we have to adopt either a universally accepted definition or a new or modified criterion of personal identity.

Shoemaker imagines a first-person version of the change of body argument. Suppose I am Brownson. After recovering consciousness following my brain operation I see that

---

my body has been replaced with another one and therefore my new body is not the body I had before the operation. I do not know about the past experiences and actions of my new body, but surprisingly, all my interests, tastes, mannerisms, personality traits are as it is as it was before. Now I shall respond to those who calls me as Brown, but obviously their reasons (behind my identity with Brown) are distinct than mine. Shoemaker comments, “I say that, I am Sydney Shoemaker, not because I claim to remember doing the things Shoemaker did, but because I do remember doing those things.” For other persons I am Brown, because I have Brown’s brain states and other psychological traits. All these considerations are relevant for them to decide whether I am Brown or not, though none of these is relevant for me. Exhibitions of my mannerisms, my interests etc. are required to prove my identity to a third-person, whereas all these are useless for me. I know that I am Brown and that is all.

When I say, “I have a toothache”, I can know it to be true without knowing anything about my present bodily-states or any of my bodily behaviour. Likewise, first-person memory statements are not grounded in bodily identity or any kind of spatio-temporal continuity. They are not grounded in the knowledge of any physical relationship between one’s present body and a past one. If I remember taking a walk last night then I know it to be true and there is no need to prove my bodily identity for making a claim to that effect.

If a person makes a genuine memory statement about his own past experiences then that statement is actually a judgment of identity. Through first-person memory statements every person claims his own identity. These statements imply persistence of a person through time but they are not based on any criterion of identity.

According to Shoemaker, there are two kinds of memory statements; one is a simple memory statement, i.e. a report or description of what I remember and the other one is the conclusion drawn from memory statements, which is based on resemblance, spatio-temporal continuity and other such considerations. According to Reid, first-person past tense statements are based on memories in the sense that they are report or descriptions of what I remember. Conclusive statements, on the other hand, need criteria for providing good grounds for their truth claims. What Shoemaker tries to explain is that, “I remember that P” is a simple memory statement, because of the meaning of the word
‘remember’. “I remember I broke the front window yesterday”, is a memory statement because it is not a conclusion drawn from anything. If I remember that I broke the front window yesterday then the question, ‘Am I the person who broke the front window yesterday?’ does not arise. To be skeptic about the truth of the statement is to be skeptic about the occurrence of that event (breaking the front window) and therefore, to be skeptic about whether one is identical with oneself.

In the broad sense of the word ‘remember’, if a person can remember some past actions then he must be a witness of that action. Putting it in another way, ‘remembering a past action’ implies being directly aware of that action. First-person past tense statements are generally memory statements in the sense that I do not need to know whether that event actually occurred or not, and if it actually occurred then whether I am identical with the observer or not. I can simply remember that I broke the front window because I witnessed that event and have direct knowledge of that event. Clearly according to Shoemaker, identity of one’s own is dependent on his direct memory claims and the person does not need any criteria to verify his memory claims. No criterion is needed from the point of view of a first-person. This is different from Swinburne’s notion of personal identity, as Shoemaker clearly explains how a memory claim can be verified by direct acquaintance to one’s own past.

Shoemaker points out that despite the differences between the psychological and non-physical criteria of personal identity and the no-criterion criteria of personal identity there are important similarities. Both distinguish personal identity as logically independent from mere bodily identity in general. Both stress on the fact that persons have minds and have priviledged access to first-person psychological states. But while one analyses the identity statements about oneself in terms of criteria of memory and other psychological phenomena, the proponents of other hold that mind is ‘sui generis’, i.e., unanalysable, and the truth of identity judgments about oneself can be known by direct acquaintance or intuition. Shoemaker defends the latter view simply because it is a ‘special and limiting case of the former’.
In response to Swinburne’s view on personal identity Shoemaker comments that the use of ‘I’ does not involve any recognition or identification of a person.\(^{85}\) In *Personal Identity* he claims that identification always includes the possibility of misidentification. But to misidentify oneself as a subject of first-person experience-ascriptions like, “I have an itch” is impossible. Likewise, on the basis of memory when I know that, ‘I did so-and-so in the past’, it is not the case that I remember someone doing that thing in the past and then identify that person as myself by what I remember about him. Neither I do this identification in terms of bodily criterion nor in terms of non-bodily criterion. My past self is the subject of my remembered actions and experiences, but when I remember those actions and experiences ‘from the inside’ my past self does not become a part of my memory, the way other persons do. If someone knows that he remembers an event ‘from the inside’ then it is a special case of directly produced self-knowledge. For he directly knows that his self enters into the content of his memory as a subject of the present remembering; not as the subject of the past actions and experiences. If anyone remembers his past experiences and actions from the inside, then “it is a short step from the knowledge that one remembers a past action or experience from the inside to the knowledge that it was oneself that did the action or had the experience.”\(^{86}\)

Shoemaker holds that bodily, non-bodily, and dualistic theories of personal identity are complementary to each other, and each one arises out of the incompleteness of the other. If someone argues for no-criterion criteria then it is a totally independent theory, and according to this theory there are no constitutive criteria of personal identity at all, i.e., personal identity is indefinable, and does not consist in anything. Nevertheless, Shoemaker admits that the facts of self-knowledge are entirely compatible with functionalist psychological continuity view and the latter is entirely compatible with the materialist view of the world.

If we admit the possibility of changing bodies like the Brown-Brownson case then it will be incorrect to conclude that a person is simply identical to his body. But according to

---


Shoemaker, this conclusion is, however, entirely compatible with a materialist view of personal identity. According to a version of materialism (identity theory), sensations and other mental states are identical to brain-processes. For example, pain is identical to C-fibre firing in the central-nervous system. And following the functionalist structure of the world it is highly plausible that the same pain (which is a functional property) is variably realized by different species in different worlds. What materialism wants to prove is that each and every mental state is realized by one or other physical state. Variable realizability is compatible with materialism. Analogously, some philosophers hold that though personal identity is normally realized by identity of body, it is plausible that it can be realized by some other kind of physical medium. For example, many philosophers think that to realize personal identity brain is sufficient. These opinions are clearly materialistic, and if personal identity is realized in any of these ways then it is fully compatible with materialism.

We want to point out that contrary to Swinburne, who himself describes his theory as a kind of dualism, Shoemaker claims that his theory of personal identity is compatible with materialism. According to him, ‘copersonality’ or ‘psychological unity’ might itself be defined functionally. In his own words,

“Given the intimate internal connections there are between the notions of the particular mental states and the notion of co-personality, it is impossible for the mental states to be realized physically without the relation of copersonality also being realized physically. For the realization of a mental state involves the existence of a mechanism whereby it stands in the causal connections that are definitive of it – i.e., a mechanism whereby it produces copersonal successor states in conjunction with co-personal states simultaneous with it. If this mechanism is entirely physical, so will be the realization of the relationship of copersonality.”

The dualistic view of personal identity does not provide any sound basis for the nature of self-knowledge. The implication of the functionalist account is that there is no a priori reason that shows why mental states might not be realised in terms of physical states. Again there are no empirical evidences which compel us to admit that mental states

---

87 Ibid.,107.
might be realised non-physically. Hence, we ought to believe that mental states are realized physically.

So far we have discussed different versions of memory criterion and psychological continuity criterion. We now turn to Derek Parfit’s views on personal identity. Parfit is the most contemporary among the significant contributors in the field and at the same time his views are in some respects quite remarkable. Because, contrary to the previous philosophers, Parfit thinks that partial continuity of a person (i.e., partial survival) is possible. According to most of the philosophers person is something indivisible and cannot continue partly like a car. Questions related to personal identity have strict yes or no answers. Parfit shows that this idea of personal identity is based on some false beliefs.

The first belief is that the questions related to personal identity have definite yes/no answers. A person believes that whatever happens to him in future, either he will exist, or he will not (as survival is generally held to be synonymous with identity). The second belief is about the importance of personal identity. Parfit claims that, “Certain important questions do presuppose a question about personal identity. But they can be freed of this presupposition. And when they are, the question about personal identity has no importance.”

According to Parfit, those who believe that certain important questions do presuppose personal identity do not think that all these questions have definite answers, rather we must decide upon an answer.

To analyse the first belief Parfit explores the thought-experiment originated by David Wiggins. In the experiment two hemispheres of a person are separated and transplanted into two brainless bodies. Both people survive as a continuant of that person and get his character and apparent memories. The question is, after the operation where does the original person go, and if it happens to me then where shall I be? If someone replies that I shall not survive then it is an implausible answer for Parfit, for I should survive if each half of my brain was successfully transplanted, and double success could not fail. On the other hand, if anyone suggests that I shall survive as one of the two resulting persons then also it is absurd. If each half of my brain were successfully transplanted there

---

89 Wiggins, Identity and Spatio-temporal Continuity, 50.
would be nothing to make me one of them instead of the other. Even one cannot claim that he would survive as both, if by ‘survive’ he means being identical with his pre-operation self. Because that would go against the principle of identity as the relation of identity is one-one. Therefore, there is no way to solve the problem and it should be freed of the presupposition that questions of personal identity have definite answers.

In Wiggins’ experiment the brain of a person is divided and housed into two separate bodies. As a result each of the two resulting ‘products’ has all the attributes of a person. In Parfit’s own words, “Suppose we admit that they are different people. Could we still claim that I survived as both, using ‘survive’ to imply identity?” Parfit replies, “We could, for we might suggest that two people could compose a third. We might say, ‘I do survive Wiggins’ operation as two people.’ They can be different people, and yet be me, in just the way in which Pope’s three crowns are one crown.” According to Parfit, this kind of reply demands that to keep the language of identity one should change the concept of a person. Otherwise, one has to give up the language of identity. Parfit argues in favour of the second alternative. He suggests that one can survive as two different people without implying that he will be these people.

Philosophers who believe that any question on personal identity deserves a definite answer generally try to avoid Wiggins’ case, as it is doubly perplexing. If we give up the belief that the question has a definite answer, as Parfit opines that we should, these problems disappear. In that case there is nothing perplexing in Wiggins’ case, it is just like many others, “in which, for quite unpuzzling reasons, there is no answer to a question about identity.”

In this way and through Wiggins’ example Parfit shows that the first belief about identity is implausible. Regarding the second belief Parfit’s viewpoint is that we can solve important questions only by separating them from the question about identity. Once we can do this, according to Parfit, the question about identity has no further importance.

---

91 Ibid., 8.
92 Ibid., 8.
To establish the last point Parfit explores the question of survival. “Shall I survive” is a question of survival, that seems as if equivalent to “Will there be some person alive who is the same person as me?” No doubt the last question is a question about identity. If these two questions are equivalent, then our accepted results of Wiggins’ experiment (i.e. I survive with two bodies and a divided mind) seems highly unsatisfactory. Supporters of the second belief would say that “I ought to regard Wiggins’ operation as death”. The person who accepts that he would survive with two bodies and a divided mind, might believe that his relation to each of the resulting people fails to contain some elements which is contained in survival. According to Partfit, this could not be true. His conclusion in his own words is this, “The relation of the original person to each of the resulting people contains all that interests us – all that matters – in any ordinary case of survival. This is why we need a sense in which one person can be survived as two.”

Parfit therefore holds that personal identity is ultimately indeterminate and we have no idea how to answer a question about personal identity. Certain important questions, for example, about memory, survival etc, does presuppose a question about personal identity. But they can be freed of this presupposition. And when they are, the question about identity no longer has any importance.

In Reasons and Persons Parfit explores two thought experiments to show that personal identity is indeterminate. In both experiments someone is travelling from earth to Mars. In the first example after entering the teletransporter the traveler becomes unconscious for one hour. A scanner on earth destroys his brain and body while recording the exact details of all his states, and then creates an exact replica in Mars out of new matter with the recorded states of the original person. The new person, when becomes conscious finds no change at all in his physical and psychological features. Everything remains exactly alike, ‘same’ (or better to say exactly similar) body, ‘same’ brain, exact memories with ‘same’ personality.

In the second example of teletransportation the person remains alive with full consciousness. The scanner records only the blueprints without destroying the traveler’s

---

93 Ibid., 9.
94 Ibid., 10. Emphasis mine.
brain and body and the person can even speak to his replica. That is to say in this example the traveler and his replica co-exists. Parfit calls the second example Branch-line-case. Both experiments are constructed to show the indeterminacy in the nature of personal identity, that is, all questions of personal identity could not claim a strict yes/no answer.

According to Parfit, supporters of physical continuity criterion would reject teletransportation, for there is no spatio-temporal continuity between the earlier and the later person (i.e., the original person and his replica). Two persons have distinct bodies and, therefore, are distinct. The followers of psychological criterion also hold the same view, for according to that criterion, remembering an experience means actually having that experience in the past and the apparent memory of the experience is causally dependent on having the experience in the normal way. Obviously this view would reject teletransportation, since the causal dependency between my experiences and the experiences of my replica is not a normal one. Parfit proposed a wider version of psychological continuity criterion, which would not reject teletransportation.

In *Reasons and Persons* Parfit also distinguishes between two kinds of identity: numerical and qualitative. A billiard ball is numerically identical with itself; two billiard balls are qualitatively identical with each other. Different problems of personal identity are usually concerned about numerical identity. Qualitative identity is generally a matter of degree. Nevertheless in some cases of personal identity qualitative identity matters and that happens if suddenly I cease to exist and a new person comes into existence as in Parfit’s example of the first thought experiment.

From Parfit’s point of view ‘memory’ is to be redefined as a sort of weak memory that does not presuppose personal identity. He introduced the notion of *q*-memories that are almost like ordinary memories, except that they are memories of experiences which could be had by someone else. *Q*-memories do not presuppose identity. In case of *q*-memory one ceases to assume that his apparent memories are, simply, *q*-memories of one’s own experiences. He wants to drop the concept of memory and use in its place the
wider concept of $q$-memory which does not presuppose the relation of personal identity. $Q$-memories are seem like memories, so one can $q$-remember having experience.\textsuperscript{95} Parfit anticipates an objection against $q$-memories, which is that $q$-memory, like ordinary memory, seems also to presuppose identity. It may be said that, “my apparent memory of having an experience is an apparent memory of my having an experience. So how could I $q$-remember my having other people’s experiences?” Parfit’s reply is that when one seems to remembers having an experience, he seems to remember having it (from inside that is, in Shoemaker’s terminology). But the underlying feeling of the identity that the person who is now remembering the experience is same as the person who had the experience is not a part of his recollection. Therefore, $q$-memory does not presuppose identity. Hence, if someone knows what a $q$-memory is he should no longer assume that his apparent memories were about his own experiences. To differentiate apparent memories from $q$-memories one should ask the following questions: “(1) Does it tell me about a past experience? (2) if so, whose?”\textsuperscript{96}

In Wiggins’ case the resulting people, both have apparent memories of living the life of the original person. If they agree that they are not this person, they will have to regard these memories as only $q$-memories. According to Parfit, every memory is $q$-memory. From now we should describe the relation between an experience and what we call a ‘memory’ of this experience in a way which does not presuppose that they were had by the same person. This way of describing this relation has certain merits. It vindicates the ‘memory criterion’ of personal identity against the charge of circularity.

Parfit’s standpoint resembles to some extent the Humean theory of the self with regard to personal identity, according to which sometimes it is simply indeterminate whether we would like to maintain it was the same or a different person. In the Treatise Hume expressed that he could not explain the problem of personal identity and held that personal identity is a fiction. Such a view supports anti-realism with regard to the notion of self. Normally we assume without questioning that there is a hard fact about personal identity, that self is something which either determinedly exists or fails to exist. But Parfit tries to explain that this is not so. On the contrary, we have to conceive personal

\textsuperscript{95} We have mentioned Parfit’s definition of $q$-memory in the footnote of page 26.

\textsuperscript{96} Ibid., 15.
identity as consisting of various relations which may hold in varying degrees, sometimes generating the indeterminate cases.

Parfit distinguishes between psychological connectedness and psychological continuity. He prefers to use the expression ‘psychological connectedness’ when the relations in question are directly causal and the expression ‘psychological continuity’ when we speak of overlapping layers of psychological connections. Psychological continuity is transitive. If, for example, $A$ is supposed to be survived (in the Parfitian sense) by $B+1$, $B+2$; again $B+1$ by $B+3$, $B+4$ and $B+2$ by $B+5$, $B+6$ and so on up to $B+30$, then $A$ will be psychologically continuous to $B+30$, but not psychologically connected. In Reid’s brave officer example, there is a psychological continuity between the brave officer and the boy who was flogged though there is no direct psychological connectedness, because, the officer could not remember the flogging. Parfitian survivors, therefore, are not psychologically connected to the original person, though they are psychologically continuous; whereas survivors of Wiggins’ experiment are psychologically connected. The phrase ‘a past self of’ implies psychological connectedness. ‘Being a past self of’ is treated as a relation of degree, so that this phrase can be used to imply the varying degrees of psychological connectedness. By the notion of ‘psychological connectedness’ Parfit explains personal identity.

From the views of Derek Parfit, it follows that there are some questions of personal identity, specifically those that are branch-line cases, which do not deserve yes or no answer. It is a true fact that in every field of science and philosophy there are some grey areas which we do not know how to treat. Actually Parfit repeatedly stresses on the aspect of qualitative similarity, but not on numerical identity. Hence, in a sense Parfit’s entire discussion is one sided.

Till now we have explored various philosophical theories of personal identity. It is clear more or less that every theory wants to explain personal identity either in terms of bodily continuity criterion or in terms of memory criterion (no matter whether it is genuine memory, apparent memory or $q$-memory), whereas at the same time each one claims that the problem of personal identity is something unique in character. This is not so much because of the term ‘identity’; rather, the term ‘person’ seems to be peculiar in character. To understand the problem of personal identity we should therefore understand the
nature of persons. So our next attempt should be to analyse the concept of a person. That is the subject matter of the next chapter.