

# Perspectivism and Intersubjective Criteria for Personal Identity: A Defense of Bernard Williams' Criterion of Bodily Continuity

*Tristan Guillermo Torriani*\*

**Abstract:** In this article I revisit earlier stages of the discussion of personal identity, before Neo-Lockean psychological continuity views became prevalent. In particular, I am interested in Bernard Williams' initial proposal of bodily identity as a necessary, although not sufficient, criterion of personal identity. It was at this point that psychological continuity views came to the fore arguing that bodily identity was not necessary because brain transplants were logically possible, even if physically impossible. Further proposals by Shoemaker of causal relations between mental states in our memory and Parfit's discussion of branching causal chains created additional complications. My contention in this paper is that psychological continuity views deflected our attention from what should have remained in the spotlight all the time: the intersubjective character (or not) of criteria proposed to decide personal identity in our language game, and ultimately our form of life concerning ourselves as persons. B. Williams' emphasis on the body was not just common sense. It was also recognition of the importance of giving priority to criteria that could be kept under intersubjective control.

**Keywords:** Criterion, Intersubjectivity, Personhood, Personal Identity, Perspectivism, Self-Concept, Will.

**Resumo:** Neste artigo passo em revista os estágios da discussão sobre identidade pessoal anteriores à ascensão de posições Neo-Lockeanas que passaram a privilegiar a continuidade psicológica. Em particular, interesse-me pela proposta inicial de Bernard Williams sobre a identidade corporal como sendo critério necessário, embora não suficiente, da identidade pessoal. Foi a partir deste ponto em que os posicionamentos defendendo a continuidade psicológica vieram mais à tona alegando que a identidade corporal era necessária porque transplantes cerebrais eram logicamente possíveis, mesmo que fisicamente impossíveis. Propostas subseqüentes de Shoemaker sobre relações causais entre estados mentais na memória e a discussão por Parfit de cadeias causais ramificantes criaram dificuldades adicionais. Minha preocupação neste artigo é que as visões sobre continuidade psicológica desviaram nossa atenção do que deveria ter permanecido em foco o tempo todo: o caráter intersubjetivo (ou não) dos critérios propostos para

---

\* Professor colaborador do Instituto de Artes da UNICAMP. *E-mail:* torr@hotmail.com. Artigo recebido em 30.09.2007, aprovado em 06.05.2008.

decidir identidade pessoal em nosso jogo de linguagem e, em última análise, de nossa forma de vida referente a nós mesmos enquanto pessoas. A ênfase de B. Williams sobre o corpo não era apenas bom senso comum. Era também um reconhecimento da necessidade de se dar prioridade a critérios que poderiam ser mantidos sobre controle intersubjetivo.

**Palavras-chave:** Auto-conceito, Critério, Identidade pessoal, Intersubjetividade, Perspectivismo, Pessoa, Vontade.

## 1 Introduction<sup>1</sup>

In our everyday lives, doubts about personal identity are not very common and tend, therefore, to be conspicuous and worthy of becoming a subject of conversation among friends (“Did you see so-and-so? I could hardly recognize him when I saw him the other day!”). A person may be hard to re-identify because he or she went through dramatic changes in appearance and behavior. Some of these changes may be simply developmental: a child who grows very tall in a few years, or a young adult who ages out of grief after many misfortunes. Facial recognition is clearly the most important way we re-identify persons. A mutilated corpse that had its head severed is harder to recognize if the body lacks any characteristics or markings (such as scars and tattoos) that differentiate it to an external observer.

It is also a fairly common experience to see identical twins play games by switching identities. It may be difficult even for parents and relatives to distinguish between them. Such cases of maximal similarity pose the greatest challenge to the external observer who is trying to decide what the person’s identity is. Much may depend on the observer’s capacity to distinguish facial and general physical traits. It is known that there is an ethnic factor in facial recognition: a white Caucasian person placed in countries with a racially different but still rather homogenous environment such as

---

<sup>1</sup> Research for this paper was funded by scholarships from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Deutscher Akademischer Austauschdienst (DAAD). I would also like to thank Dr. José Oscar de Almeida Marques and Dr. Michael Beaumont Wrigley of the State University of Campinas (UNICAMP), Brazil, and Prof. Dr. Arno Ros of Otto-von-Guericke University at Magdeburg, Germany.

Japan, Korea, China, Africa or India will generally have difficulty distinguishing between individuals because, in his perception (not in reality) they “look all alike”. Time is needed to develop a certain familiarity with different facial features so that individuals of other ethnic groups can be recognized sooner. Failure to develop such skill can lead to very embarrassing situations, such as talking to a person while believing that he is someone else, and even mistaking his or her name.

There are, therefore, objective as well as subjective conditions that may generate confusion about personal identity. There is a legal concern about this because persons are subject to the law and must answer for their actions. All countries around the world issue some kind of identity card (in the USA a driver’s license is used) for this reason. A recent development in this so-called information age has been identity theft by electronic means, over the Internet. Once your personal data has been “phished” (or collected), another person may take your place and do all kinds of illegal operations in your name and law enforcement will come after you for it.

Serious as all of this is, we still remain to a reasonable degree able to re-identify persons. Most cases of personal identity are decidable by finding adequate criteria that can be controlled intersubjectively (i.e., among other persons themselves). Persons are beings who have the ability to decide who is a person and who is not, and who is the same person and who isn’t. The ability to re-identify persons is a social skill that varies from individual to individual, but it has to be present in the vast majority for rather evident evolutionary reasons. Humans incapable of distinguishing between friends and foes, between kin and non-kin, simply could not survive natural selection.

As far as living persons are concerned, it seems that most problems can be sorted out in everyday situations by plain common sense. Philosophers started proposing puzzling cases, however, to basically test the logic underlying these cognitive abilities and to clarify what would be the necessary and sufficient criteria for

personal identity. In addition, there was a more spiritual concern. We know that much of Locke's interest in personal identity was motivated by doubts concerning immortality and life after death. While distinguishing between identity of consciousness, substances and men, he pondered about the perplexing possibilities created by reincarnation. Indeed, insofar as we are persons, the issue of personal identity and the conditions of personhood is, despite its technicality, the closest one can get to the ancient Delphic exhortation to know oneself.

Much depends on how we define the problem of personal identity. It is not just a psychological or neurological investigation into how we actually re-identify persons. As Dennett ((Dennett 1984)) pointed out after discussing his six conditions of personhood<sup>2</sup>, an important aspect of this problem is its moral and normative dimension.

Now finally, why are we not in a position to claim that these necessary conditions of moral personhood are also sufficient? Simply because the concept of a person is, I have tried to show, inescapably normative. Human beings or other entities can only aspire to being approximations of the ideal, and there can be no way to set a "passing grade" that is not arbitrary. Were the six conditions (strictly interpreted) considered sufficient they would not ensure that any actual entity was a person, for nothing would ever fulfill them. The moral notion of a person and the metaphysical notion of a person are not separate and distinct concepts but just two different and unstable resting points on the same continuum. This relativity infects the satisfaction of conditions of personhood at every level. (D. Dennett - *Brainstorms*, p. 285).

Philosophical discussion of personal identity tends to focus on the cognitive practicality of necessary and sufficient criteria for continuity, identity, endurance and perdurance of persons through

---

<sup>2</sup> Dennett's conditions are:(1) All persons are or must be rational, (2) All persons are or must be beings to which mental states can be attributed, (3) All persons are treated in a special way, which involves recognition and a special attitude towards them, (4) All persons are or must be capable of reciprocating this attitude, (5) All persons are or must be capable of communicating through language, and (6) All persons are or must be self-conscious in a special way.

time but neglects the conative (voluntary) aspect that is essential in the adoption of self-concepts. In the view of personhood I would like to advance here, both first person (singular) perspective (subjective and phenomenological) and third person perspective (intersubjective and heterophenomenological) should be given their due. However, insofar as the phenomenological perspective provides us with a merely private or subjective criterion of continuity, it is clearly insufficient in face of the possibility of deep self-deception. For this reason I tend to favor heterophenomenological criteria such as bodily continuity and numeric identity (i.e., my position is close to endurantism). I admit, as Dennett does too, that there is a degree of arbitrary preference in this grammatical regulation of our criteria for personal identity. We have to be clear about what we are trying to do here. There is nothing in the “nature of things” that will justify the choice of one set of criteria over the other. It is up to us, in our language game about persons, to decide what criteria we are going to give preference to when we talk about ourselves.

In addition, even though it cannot really validate criteria of personal identity that do not pass checks for intersubjective transparency, logical necessity or sufficiency, the will has a central and decisive place. Our self-concepts are related to our self-esteem. Certain views of personhood may be intolerable to some, either because they are too “thin” or because they are too “thick”.<sup>3</sup>

In this article I revisit earlier stages of the discussion of personal identity, before Neo-Lockean psychological continuity views became prevalent. In particular, I am interested in Bernard Williams’ initial proposal of bodily identity as a necessary, although not sufficient, criterion of personal identity. It was at this point that

---

<sup>3</sup> In his book *Materie und Geist*, Arno Ros ((Ros 2005)) presents a painstaking methodic reconstruction of intersubjective criteria for distinguishing between organisms (Lebewesen), agents (Handlungssubjekte) and persons (Personen). My approach here is, however, somewhat different in that my major concern is in cultural contextualization of the personal identity debate so that non-rationalistic views can be also considered and not unilaterally dismissed. This difference in approach is roughly analogous to Von Wright’s (Von Wright, p. 160, footnote # 85) distinction between analytical and dialectical ways of doing philosophy.

psychological continuity views came to the fore arguing that bodily identity was not necessary because brain transplants and mind/body swaps were logically possible (and intuitively plausible), even if physically impossible. Further proposals by Shoemaker of causal relations between mental states in our memory and Parfit's discussion of branching causal chains created additional complications. My contention in this paper is that psychological continuity views deflected our attention from what should have remained in the spotlight all the time: the intersubjective character (or not) of criteria proposed to decide personal identity in our language game, and ultimately our form of life concerning ourselves as persons. B. Williams' emphasis on the body was not just good old common sense. It was also recognition of the importance of giving priority to criteria that could be kept under intersubjective control.

## **2 Points of Agreement with T. Sider and M. Eklund**

It is noteworthy that in the recent literature, both Sider (Sider 2001) and Eklund (Eklund 2002) acknowledge the importance of Bernard Williams' work, in particular, his 1970 paper "The Self and the Future" (Williams 1973).

Sider's position is closer to perdurantism, except that instead of worms that share common parts through time, he is interested in previous and later stages. If worm theory is true, there would be multiple candidates for continuity and the issue would remain undecidable because of semantic indeterminacy. We basically would be unable to know (or tell) whether we are talking about a  $\psi$ -person (i.e., a person who is psychologically continuous but not necessarily bodily continuous) or a  $\sigma$ -person<sup>4</sup> (i.e., a person who is bodily continuous but not necessarily psychologically continuous). After mentioning B. Williams's torture and memory transplant case, Sider (Sider 2001, p. 197) observes that:

---

<sup>4</sup> This is my terminology, not Sider's. I suggest  $\sigma$  (sigma) because it can stand for  $\sigma\mu\alpha$  (soma, body in Greek).

It appears that we are capable of having either of two intuitions about the case, one predicted by the psychological theory, the other by the bodily continuity theory. A natural explanation is that ordinary thought contains two concepts of persisting persons, each responsible for a separate set of intuitions, neither of which is *our* canonical conception to the exclusion of the other.

I wholly agree with Sider on this. There are two sets of intuitions that we rely on to ascribe personal identity and we need both of them. It makes no sense to want to do away with either one of them. However, insofar as psychological criteria are subjective (phenomenological or in first person perspective), I would plead here for the priority of bodily continuity as the main criterion. I say this is a plea because a suggestion concerning the grammar of our concepts is not (and cannot be) a deductive argument that presupposes logical necessity because inferential rules themselves depend upon a previously accepted grammar. The adoption of a certain language concerning persons ought to be consensual. It is a fundamental misunderstanding of the functioning of concepts and language to believe that the acceptance or rejection of concepts can be compelled by deductive argument. This is why Dennett's observation about the normative character of the concept of a person is something to keep in mind.

Sider (Sider 2001, p. 203) would like, however, to find out what metaphysical option is correct so that we could escape the (to him distressing) predicament of there being no fact of the matter who is right. He compares the views of the perdurantist, the chaste endurantist, and the nihilist (i.e., an atomism that denies the existence of persons) and would like to see one proven right.

I say there exist temporal parts; the chaste endurantist disagrees. And each of us disagrees with the nihilist in thinking there exist composites. These disagreements are not merely over how the world should be described; we disagree about what there is. These disagreements cannot, I think, be dissolved. There must be a fact of the matter who is right.

There seems to be a misunderstanding in Sider's view about what analysis can achieve in philosophy as a problem-solving

method. A couple of metaphors may be of use here. The philosopher may be compared to a mechanic who has to disassemble an engine to discover where the faulty piece is. Or he may be compared to a family member who is chosen to carve the turkey for Thanksgiving. But we know that others in their place may do their job differently. Is there a “correct” way to break something down? Is there only one correct analysis of the concept of a person? Must there be only one? On the contrary, I would say that there should be a plurality of views. Perspectivism, then, appears to unavoidably play a part in contributing to make the personal identity issue unsolvable but at the same time indispensable for our evolving self-understanding as human beings.

Eklund ((Eklund 2002)), p. 481) also realizes the co-existence of psychological and physical criteria. In addition, he makes the very fortunate distinction between a concern with truth conditions and a concern with the meaning of talk about personal identity. If our concern is of the latter kind, it makes sense to accept that both physical and/or psychological continuity are necessary for personal identity. Since I strongly agree with Eklund in this and his view lends support to my own, I shall quote him in full.

It is standardly assumed that since psychological and physical continuity do not always go together, we must make a choice between the psychological and the physical criterion, since a criterion like

(11) person  $x$  at  $t$  = person  $y$  at  $t'$  if and only if  $y$  at  $t'$  is either psychologically or physically continuous with  $x$  at  $t$

or

(12) person  $x$  at  $t$  = person  $y$  at  $t'$  if and only if  $y$  at  $t'$  is both psychologically and physically continuous with  $x$  at  $t$

hardly strikes anyone as plausible, in view of for example the fact that it is hardly the case that (intuitively) both persons die in the mind/body switching cases or that both original persons survive as both the resulting persons.

This standard assumption is justified if our concern is with truth conditions and truth values of statements about personal identity (and this is indeed most people's concern), but not if we are concerned with the meaning of talk about personal identity. For then we can still allow that psychological and physical continuity both count as sufficient and/or necessary for personal identity. The mind/body switching cases simply



reveal that it is a meaning (sic) presupposition underlying talk of personal identity that psychological and physical continuity always go together.

Another point of agreement with Eklund is that our concept of personal identity contains inconsistent intuitions whose very inconsistency is revealed when we consider fission cases.

The fundamental point is that it is precisely by virtue of our competence with the concept of personal identity that we have the intuitions whose joint inconsistency is demonstrated by reflection on fission cases. This point is independent of the possibility of a reductive analysis of the concept.

It seems to be a kind of presupposition of our discourse and thought about personal identity that fission cases do not occur. Speaking loosely, when this presupposition does not hold, our discourse and thought about personal identity *breaks down*. (Eklund 2002, p. 474)

While Eklund seems to conceive the conceptual inconsistency of our talk about personal identity as a matter of pre-linguistic intuitions, I would rather attribute these contradictions to differences in the conceptual grammar of different speakers' perspectives. This creates a clash that may be fundamentally unsolvable. If instead of considering criteria singly we proceeded to a direct confrontation of competing concepts of a person, what should we expect? Do we disagree about the criteria because we are assuming different concepts of a person? How can we ever expect to find a consensual definition of humanity?

I think that bodily continuity is a good starting point for those concerned with intersubjective transparency and a diversity of perspectives. Therefore, in the following I will review what I consider to be the most important arguments about bodily continuity, the bottom line being my belief that "Bernard Williams was basically right all along".

### **3 Bernard Williams and bodily continuity**

B. Williams tried to show that bodily continuity is a necessary condition for personal identity through time. He is credited for the argument of reduplication as an objection against psychological

criteria of personal identity. Williams basically argued that the psychological criterion of memory to determine personal identity failed because several persons could have similar memories and believe to be identical to a past person (in Williams' example, to Guy Fawkes). An identity relation, however, must be one-one, not one-many, as in cases of mere similarity. Memory reports gave us only a notion of similarity between persons, but not identity. We could imagine several persons reporting memories that suggested psychological continuity with Guy Fawkes, but that was not sufficient to establish the identity of one of those persons with Guy Fawkes. In this case, we would need another criterion, such as the body, to decide the identity question. Bodily continuity was a necessary condition because, as a rule, we could not distinguish so easily between the body and the mind of a person. For this reason, physical and psychological criteria could not be easily distinguished and we ended up needing the body to decide the identity or not of a person through time.

S. Shoemaker, however, undertook the defense of Neo-Lockean psychological continuity as a necessary and sufficient condition of personal identity. To do this he needed to theorize memory as a causal concept. This meant that the memory of a past fact would be always caused by the very observed fact. If we granted that an experience caused its later recollection, this meant that my past experiences would cause memories that could only be mine, thus guaranteeing my personal identity. If a person had identical memories to those of somebody who disappeared, they must be the same person, since causality was a necessary relation, and we could only have identical memories if they were caused by the same past experiences. For Shoemaker, by means of this causal concept of memory, we could dismiss bodily continuity as a criterion of personal identity.

D. Parfit went a step further, using this causal view of memory chains to analyze the problem of personal continuity without identity. He was not concerned with having to choose between bodily and psychological continuity. He thought both criteria were acceptable and applicable. On the one hand, he adopted

psychological continuity by means of the causal concept of memory proposed by Shoemaker ("quasi(causal)-memory"). On the other hand, he also conceded that the mental and bodily criteria could not be easily separated, and that, in some moments, bodily continuity could be necessary to decide whether a past person was identical to a present or future person, as Williams required.

Parfit's intention was to question the use of the logic of identity (that is, the *tertium non datur* principle) as a method of analysis of personal continuity through time. In this Parfit distanced himself from Williams and came closer to Shoemaker's position. In spite of not being neither logically nor empirically one-one, he accepted that psychological continuity could be claimed to function as a criterion of personal identity. A and B could be the same person if they were psychologically continuous and there were no other person that was contemporaneous and psychologically continuous with them.

The important step came when Parfit discussed cases of personal survival without identity. At this point, cases were conjured in which there was psychological and bodily continuity, but we did not know how to decide the question of personal identity. For Parfit, this showed that the logic of identity was inadequate to deal with cases of brain fusion and fission. We had better, therefore, think in terms of degrees of continuity between successive selves and abandon the concept of the self as an underlying entity of mental life. Parfit called the traditional Neo-Lockean (and to a degree Neo-Kantian) view of the self as an underlying unity defended by C. McGinn the "simple view", in contrast to his, which he in turn called the "complex view" about the self.

Parfit proposed that the logic of identity, which was of yes or no, ought be exchanged for the logic of survival, which was of more or less (i.e., a matter of degrees). Leaving aside the "simple view" of the self as an underlying unity of mental life, we could be more precise and consider just the degrees of personal survival after brain fission and fusion cases.

I think that Parfit's proposal had a distinctly grammatical character to it, but of a kind one would expect in what Strawson called revisionary metaphysics. As B. Williams pointed out, however, if Parfit's talk of ancestral and descendant selves were actually adopted, our concepts of parenthood and suicide would also need to be revised. To me, the point in all this is that the concepts we use to describe ourselves and the world depend upon shared (and sharable) experience. If brain fission and fusion cases were a matter of everyday experience, we would eventually have developed the language of personal continuity and ancestral and descendant selves to cope with this reality. This, however, was not the case. Parfit's revisionary metaphysics was meant for extraordinary, not ordinary cases, and therefore did not make sense from a practical point of view. As far as perspectivism is concerned, a different view such as Parfit's should be welcome. As far as intersubjective transparency is concerned, however, it would cause the conceptual confusions B. Williams pointed out.

In what follows I will review this argumentative chain in greater detail as space permits.

**3.1 The Body Theory**The body theory of personal identity defended by B. Williams consisted of the following theses:

- (1) the person is the body,
- (2) the necessary metaphysical criterion of personal identity is bodily identity.

Williams suggested (Williams 1973, p. 64) that the mind-body problem could be split into two levels, one micro and the other macro. On the macro level we would have the problems concerning the relations between the mental states of a person and her possession of ("having") a body, including, in particular, their relations to observable behavior. On the micro level we would have the difficulties concerning the relations between the mental states of a person and the internal states of her organism (in particular, brain

states) that could be correlated to mental states in a future psychophysical science.

Materialist positions could be advanced on both levels, micro as well as macro. On the micro level, H. Feigl, U.T. Place, J.J.C. Smart and others proposed what would be known as the "identity theory", in which mental states would be strictly and contingently identical to brain processes. The theory had a stronger version, called "type-type", in which each type or category of mental event (for example, sensations, perception, etc.) would be identical to a type of brain process, making psychophysical laws possible. The weaker version, called "token-token", dropped the claim for psychophysical laws and asserted identity only between any given mental event and its correlated brain process.

On the macro level, materialism could be defended by asserting the identity of a person and her respective body. A person is, after all, a concrete particular object. Identity of a person through time would then consist in the preservation of the identity of that particular object. The firmest criterion to link personal identity to the identity of a particular object would be bodily identity. This would avoid two undesired alternatives: we would not need to say that a person as a particular object was something immaterial (thus avoiding dualism), and we would not have to see a particular person as a type or concept that contained several tokens or instances (thus avoiding conceptual confusions) (Williams 1994, p. 4).

### **3.2 Memory Reduplication Argument**

Against psychological criteria of personal identity, among which memory was the most prominent, Williams presented the *argument of reduplication* in the form of an imaginary case: all of a sudden, in the twentieth century, a man named Charles suffered a radical personality change and started to say he was Guy Fawkes. To prove his personal identity with Guy Fawkes, Charles was capable of remembering all of the events witnessed by his alleged ancestral self. The events he claimed to have witnessed and the acts he claimed to have done suggested the life history of Guy Fawkes. The

verifiable memories Charles produced coincided with the description of Fawkes' life as known by historians and biographers, and those which were unverifiable were at least plausible, making certain up to now unexplainable facts explainable (Williams 1973, p. 7-8).

Since Charles' unverifiable memories provided explanations for hitherto unexplained facts, this proved that he was not deluded, and that he was indeed Guy Fawkes. These unconfirmed memories were not to be found in any existing book about Guy Fawkes and it would be difficult to invent false memories that could systematically explain what had up to then remained unexplained. Thus, according to the memory criterion, Charles would be Guy Fawkes.

But at this moment, all of a sudden his brother Robert showed up and succeeded in satisfying the memory criterion just as well as Charles (this is the reduplication). If it was logically possible that Charles could be Guy Fawkes, then it was logically possible that some other person could also be recognized as being Guy Fawkes. In this situation, it was impossible to decide which of them was Guy Fawkes. Worse, none of them could be identical to Guy Fawkes, because if they were, they would have to be identical to each other, which was impossible, since they lived two contemporaneous and separate lives. They could not both be Guy Fawkes, because if they were, Guy Fawkes would be in two places at the same time. Hence, they could not be simultaneously identical to Guy Fawkes. An identity relation occurs between an individual at a given moment  $t1$  and the same individual at a moment  $t2$ . All of his properties must be identical. It is therefore a one-one relation, and not a relation between several individuals, as in similarity cases. Charles, Robert and Guy Fawkes could be similar, but not identical.

One way out would be to say that only one of them was identical to Guy Fawkes and that the other was only *similar* to him. But we would lack in this case a criterion that would be used to decide to whom, Charles or Robert, this description would apply. Another way out would be to say that both had mysteriously become similar to Guy Fawkes, knew his life well, had great imagination, and so on. But if this was the best description for Charles and

Robert, retorted Williams, then it would be the best description for Charles if he alone had changed. In other words, if after consideration of the reduplication case we came to the result that it was more convincing to say that Robert and Charles were only similar to Guy Fawkes, then that is how we should have responded to the initial stage, when only Charles was similar to Guy Fawkes. Thus, Williams showed that memory could not provide a sufficient criterion of personal identity through time. It suggested at most *similarity*, not identity, and in the reduplication case, we remained without criteria to determine which of the two (Robert or Charles) was identical to Guy Fawkes (Williams 1973), p. 8-9).

Although Williams admitted that bodily identity was not a *sufficient* criterion, he claimed that it was always at least a *necessary* condition of personal identity through time. Without bodily continuity we could not identify persons. If we wanted to identify a particular person P who did particular act A we had to necessarily appeal to her physical presence at the moment in which A happened. By eliminating the body, we remained without objective criteria to distinguish between identity and mere similarity between two persons. There was an enormous difference between 'same body' and 'similar body'. Even if two persons had identical memories, for example, this could only tell us that they were similar. However, if we wanted to distinguish between these two particular persons, we needed the bodily criterion.

### 3.3 Body Swap (Prince and Cobbler Case)

Williams also considered the case (similar to Locke's about the prince and the cobbler)<sup>5</sup> in which an exchange of bodies took place. An emperor and a peasant would have their bodies swapped: the emperor would then be in the body of the peasant, and the peasant in the body of the emperor. Williams believed that in such situations we would probably be left in a state of perplexity. However, we could ask what this experiment implied. Once the bodies were

---

<sup>5</sup> Perhaps one should also remember Mark Twain's *The Prince and the Pauper*.

exchanged, Williams asked himself how we could identify the self of the emperor in the body of the peasant and vice versa. To him, it seemed impossible to imagine the emperor in the body of the peasant, since the peasant's body could hardly come to express the personality of the emperor, and vice versa.

The exchange of bodies showed that the concept of the self as something purely and abstractly mental was not so convincing as it seemed. We are accustomed to associate the selves of persons to their bodies as they manifest themselves in their behavior. When asked to distinguish between the body of a person and his (or her) personality, we would not know what should be attributed to one or the other. Without the embodiment of behavior it would be impossible to determine the identity of distinct personalities in time. Williams argued that if in the Charles-Guy Fawkes case the Fawkes personality changed bodies frequently, identification would become not only difficult but even impossible. The only criterion beyond the body would be that of memory, but then all we could guarantee would be at best some psychological similarity. Therefore, it was a necessary criterion for any claimed identification of persons by means of non-physical criteria that at some moment they should be made on basis of bodily criteria. Williams then concluded that the bodily criterion was indispensable for personal identity through time ((Williams 1973), p. 11).

For somebody who accepted the possibility of identifying personalities without reference to the body, the idea of a body swap would be something conceivable and not problematic. But consideration of body exchange cases showed that at least some of them were problematic. We could hardly identify the peasant in the body of the emperor, or the emperor in the body of the peasant. The strict identity of persons without the possibility of a body swap was always conceivable, but there were body swap cases in which it did not appear at all conceivable. Besides, considerations about body exchange suggested that the distinction between psychological and bodily criteria was not so simple and automatic. For Williams, this was the deeper question, that upon having to distinguish between the personality of a person and his or her body, we did not know where



to draw the distinction. This seemed to suggest that bodily and psychological criteria were inseparable in the end.

### 3.4 Memory Swap

For Williams, memory reports could not even serve as an internal or private criterion to assure the individual of his or her own identity. One way out for the memory criterion would be to argue that it could serve as an internal criterion for the person to solve identity doubts by himself. However, Williams' reduplication argument showed the difficulty of deciding the personal identity of another person (i.e., from a third person perspective) on grounds of memory claims.

Locke, the father of the psychological view, suggested that "consciousness" (understood as memory) would be what made a person be herself to herself.

For, it being the same consciousness that makes a Man be himself to himself, personal Identity depends on that only, whether it be annexed only to one individual Substance, or can be continued in a succession of several Substances. ((Locke 1975), *An Essay Concerning Human Understanding*, Book II, Cap. xxvii, #10, 15, ed. Nidditch, p. 336)

Hence we could imagine memory as a criterion of identity that each of us could use for himself to decide our personal identity through time. Perhaps memory did not work as a criterion for other persons, but at least it could serve as a subjective "internal" certainty about our identity.

However, Williams categorically rejected this variant of the psychological argument. To him, either the question of personal identity would not appear to the person himself, or if it did, then it would be undecidable without the assistance of others as external observers. To show this, he invited us to imagine a man who had previously had a set of memories *S0*, and who now had another set, *S1*. After the substitution of memory sets, the man would supposedly have to have doubts about his personal identity. In this case he would have to deploy the internal criterion of memory to

clear the doubt about his personal identity. But this was not possible, since in having *S0* memories, or in having *S1* memories, he lacked any means or reason to doubt his identity and so the question did not even arise for him. If the doubt were to come up at all, he would have to have *S0* and *S1* *simultaneously*.

To examine this possibility, Williams then asked us to imagine the case in which *S1* included a general memory *W*, by which a person could remember things that he had by now (i.e., after the loss of *S0*) forgotten, such as the very exchange of memory sets. However, only because we forgot something does not mean that we would have to doubt our own identities. Loss of memory and loss of identity are two different things. Williams suggested that we could then imagine that *S1* included a general memory *W* by which the person could remember facts that were empirically incompatible with the memories in *S1*. In this situation the subject would have to try to discover under which memories were truly his. In attempting this, his most economic hypothesis would probably be to suppose that the extended memory *W* itself was an illusion. If he were not yet satisfied with this, or if some parts of *S0* were left over in *S1*, so that he seemed to have completely contradictory "memories", he could do nothing to solve this with the help of his own memory alone. He would have to ask others about his past. In doing this he would be depending on the memories of others about his past. But then memory would *not* serve as a private or "internal" criterion for the person himself. So Williams concluded that there was no way in which memory could be used by a person as an internal criterion for his own identity. In the first case, where *S0* was changed for *S1*, the person would simply believe he was another person and would not have how to doubt his own new identity. In the second case, where *S1* conflicted with *W*, the person would not be able to decide his identity without appealing to others. Thus, Williams showed that the third person perspective and intersubjective control were inescapable.

Bodily criteria cannot be completely excluded from personal identity criteria. The reference to the body is fundamental and inescapable. A person is a particular material object and personal

identity through time necessarily requires bodily identity. These were B. Williams' basic conclusions and in my view they still hold.

### **3.5 Objections (C. McGinn): Brain Transplant and Brain Identity**

There are, of course, objections to the body theory of personal identity, some of which require the thought experiment called the "survival test" as a useful technique to evaluate personal identity criteria. As we consider a change of our self such as a neurosurgery or a psychotic attack, we can ask ourselves "If this happened to me, would I survive?" If we answered yes, this would mean that this change affects something inessential to our self. If we answered no, then probably we would have identified something essential to personal identity. We can use the technique for necessary conditions as well as sufficient conditions. It is important to always check whether the reason for the answer (yes or no) is free from circularity (or not begging the question), and not appealing to the expression 'same person'.

The survival test generates the most basic objection to the body theory: the logical possibility of a successful brain transplant (Shoemaker 1984, p. 43-44; McGinn 1982, pp. 109 and 114). Suppose that my brain were surgically removed from my body and placed in another one similar to my original body, so that now this person has my brain. Applying the survival test, I ask: "If this happened to me, would I survive?" The answer would seem affirmative. By means of brain transplant it is plausible to believe that good part of what I consider my self would survive, for now, thanks to modern science, the brain is seen as the centre of mental life. The fact that the rest of my body subsists as the body of someone else would not therefore constitute neither a necessary criterion nor a sufficient criterion for personal identity through time. It would not be necessary because I could survive in the body of another person through brain transplant. It would not be sufficient because my body could be used as a receptacle for someone else's brain. The argument depends on (a) the non-inclusion of the brain in

what would be considered the person's body and (b) on the idea of the brain as the home of the self. Bodily continuity would not constitute a metaphysical criterion of personal identity, for my body could then be occupied by a new self by means of a brain transplant. Of course, we can call this distinction between the brain and the person's body into question, since the former is part of the latter.

McGinn admits that a way out for the body theorist would be to claim that personal identity is brain identity. A person would be the same if her brain remained unchanged. But the problem of this brain criterion, objects McGinn, is that the analysis of the concept of the brain does not seem to have direct relevance to the analysis of the concept of the self. To have the concept of the self does not depend upon having the concept of the brain, for we can know what is a person without knowing anything about neurology and identity conditions between brains. In this sense it would be incorrect to claim that our judgments about personal identity are grounded on judgments about identity between brains. The criterion of brain identity is not necessary because there is no conceptual necessity that persons have the brains we in fact have. Our concept of the self would not be different if our brains were made differently. Besides, the physical basis of the self could be not an entity, but a process. If we discovered that our brains renewed their tissues continually we would not have to abandon our conviction that today we are the same person as yesterday. Our concept of individuality and persistence of the self is simply independent of the ontological status of our brains. For McGinn, the brain criterion is not sufficient because the self ceases to exist as such upon death, but the brain continues. If somebody received my brain by transplant, this would not be sufficient to establish that that person is me. The self is the reference of the term 'self' and its proper characteristics are mental, seen mainly from the first person singular perspective. The brain is an organ of our bodies, to be investigated empirically, that is, in third person. McGinn's conclusion is that in the same way that it is impossible for the physical states of the brain to present the phenomenal properties of the self's mental states, it is also impossible for the brain to be the self.

In my view, McGinn's objections are fine and well taken, but they concern mostly the logical insufficiency of physical criteria, not their necessity. McGinn's view of the self, after all, follows Strawson and Kant in postulating it as an underlying psychic unity, neglecting our physical dimension.

#### **4 Causal psychological continuity (Shoemaker)**

In response to the body theory of personal identity, S. Shoemaker defended psychological criteria against the objections raised by Williams against memory. The concept of the self advocated by Shoemaker can be expressed in the following theses:

- (1) the self is a construction of mental states,
- (2) the necessary and sufficient metaphysical criterion is continuity in the causal relations between the mental states of a person.

The psychological theory of personal identity states that A is the same person as B if and only if A's mental states satisfy identity relations with B's mental states. Identity relations must obtain between mental states of A and B that occurred through time. The psychological criterion must also explain the identity conditions for a type of object such as the self in terms of relations between entities of another kind that do not already assume an identity relation. The task is to see which relations shall serve this purpose in the case of personal identity through time.

As we saw earlier, similarity could not be a necessary criterion because a person could suffer psychological changes during his or her life history (traumas, brain-washing, mystical experiences, etc.). Similarity was not a sufficient criterion because several persons could be psychologically similar and identity must be a one-to-one relation. So mere psychological similarity among persons could not be a metaphysical criterion for personal identity through time. The difference between similarity and identity was

precisely that identity is a one-to-one relation, whereas similarity may occur among many (Williams 1973, p. 25).

Shoemaker proposed the criterion of causality between successive mental events as a way to buttress psychological criteria and overcome the limitations of mere psychological similarity. Personal identity could be conceived as a series of causal relations between mental events. Thus, mental states would be part of the same object, the self, if and only if they could be causally related in a certain way.

The main case of psychological causality that could serve as a basis to explain the self was memory. Past experiences caused our later memories of them. As we maintained or changed aspects of our personalities we would depend at least in part on causal relations with personal properties we already had earlier. Self-consciousness through time also depended partially on causal relations between successive mental events. Our self-concept was, according to Shoemaker, in great part determined by the memory of our past actions. This self-concept involved the concept the person had of his own character, values and potentialities. In addition, Shoemaker saw what he called our 'future history' as the main focus of our desires, expectations and fears. Shoemaker admitted that this conception of the self might not provide truths about the concept of personal identity but he hoped that it would show the importance of the concept of personal identity in our conceptual scheme and our form of life (Shoemaker 1984, p. 48).

According to Shoemaker's psychological view, persons had a special access to facts about their own lives and identities that could not be had about the life-histories and identities of other persons. This special access had two properties:

- (1) Necessary witnessing - if somebody remembered a past state, then that person must have witnessed that event directly, and
- (2) Infallible self-reference - it was impossible that a person had a complete and precise remembrance of a past event and mistakenly attributed his own action to another

person, or another person's action to himself, on condition that the person had not previously lost his capacity to differentiate his personal identity and was not confused while witnessing the event.

For Shoemaker, it was by means of our memory that we would have access to our personal identity, thanks to the two features of our special access. If we remembered a past event, we would know for sure that we witnessed an event: our identity would not be called into question. If a remembrance was really ours, we could not attribute it to others: again, memory would guarantee our identity.

#### **4.1 Objections (McGinn and Ayers): Mind Swap and Body Consciousness**

McGinn's objections to psychological criteria try show that they are neither necessary nor sufficient (McGinn 1982, p. 110-113). To show that psychological causality is not a necessary condition of personal identity through time, it suffices to identify a case in which there is survival in the absence of mental causal relations. In the mind swap case, we could imagine a person who, suffering from amnesia, lost her memory and personality, receiving then in their place a new set of mental states. In this possibility, later mental states would not be causally related to earlier ones. However, when we ask whether we would survive amnesia and mind swap we could not deny this possibility. It is perfectly possible for McGinn that we would survive after this change, retaining our self. It all depends on the similarity (or not) of the mental states introduced after the swap. If they are different, perhaps we could say that now we are another person, since our new mental states have no causal relation with their immediate predecessors, neither are they similar. But if a strictly identical set of new mental states were introduced in the place of mental states we had before the swap, without any causal relation whatsoever, we would probably continue to be the same

person. Therefore, our survival shows that psychological causality is not a necessary metaphysical criterion for personal identity.

To be able to demonstrate that mental causality is a sufficient criterion for personal identity we need to eliminate any possible circularity. The criterion of causality must be conceived in a way that does not presuppose the identity of the persons between which mental causality occurs. The persons must be identical only because of this causality. Besides, we must restrict the concept of causality so as to exclude cases of causal relations between different persons, for example, when somebody comes to have a belief because of another person's utterances. Thus, constraints are necessary to decide which causal relations between mental states are linked to mental states of the same person. To restrict psychological causality to mental states of the same person in a non-circular way we can try to use the body or the brain as a criterion. We could then say that the causal relations that preserve personal identity are those that occur within the same body or brain. But in doing this the criterion ceases to be strictly psychological. It seems, therefore, plausible to conclude that causality, being too open and unrestricted, allows for the identity of different selves, being thus insufficient as a criterion.

For McGinn, another difficulty for the psychological concept is that the persistence of the self seems to be something deeper and more fundamental than the causal interaction between mental states inside it. The self seems to underlie this interaction, and not be originated from it. Due to this primordality of the self, the causal psychological criterion is not sufficient to determine personal identity through time.

Shoemaker's concept of the self betrayed a clear Lockean inspiration. He not only tried to save psychological criteria as the most important in personal identity, but also endorsed a self that was conceptually disconnected from the body. Michael Ayers (Ayers 1991) criticized Shoemaker's Neo-Lockeanism in a spirit apparently akin to Williams. For Ayers, the myth of a pure subject of thought as something at least conceptually distinct from the bodily self was accepted too uncritically by some Neo-Lockeans.



To illustrate this, Ayers pointed out a passage in which Shoemaker claimed we were introspectively aware of our own thoughts, feelings, beliefs and desires without representing them to ourselves as flesh and blood persons or even as mere objects in the world.<sup>6</sup> For Shoemaker, in the act of remembering our own past actions and experiences (from a first person singular perspective), our bodily self, the physical subject of these very actions and experiences, was *not* given in the content of our memory in the same way as that of other persons (seen from a third person perspective). If, on the contrary, the physical subject were in fact given in the content of experience and memory, then we could in principle mistakenly identify the self as being another person, since the bodily self of other persons did come into the content of our memory. But this was false, Shoemaker argued, because when I am having an experience it is impossible for me to know that somebody is experiencing it and not know that that person is myself. Hence, he concluded, our bodily self is *not* given either in experience or in our memory. Our bodily self is absent *because* we could not mistake ourselves in identifying it as our own. If it were given, we would be able to mistake ourselves, but since this is not the case, then it is absent from our experience.

For Ayers, Shoemaker's argument was clearly invalid and led to several absurdities. If it were valid, bodily sensations would be impossible. Shoemaker could try to argue that our bodily self did not present itself in a special way to our sensory experience (assuming that there was no other bodily self), then we would not be able to have bodily sensations. According to Ayers, Shoemaker would have to argue that, on the contrary, if my bodily self were present in self-consciousness, then I could be conscious of any bodily self whatsoever and still be deceived concerning which bodily self was represented in my consciousness, confusing it with someone else's. For Ayers, the impossibility of mistaking ourselves in identifying the bodily self we are conscious of *in no way justified*

---

<sup>6</sup> As is well known, this was an important issue for Merleau-Ponty in the phenomenological tradition.

*the denial that our bodily self is presented to our self consciousness.* Indeed, we have bodily sensations such as pain and it simply does not make sense to doubt whether the bodily self as it is presented to self-consciousness is ours or somebody else's. But as we experience bodily sensations, we are conscious of our bodily self, and this *is* given in every experience. In this way, contemporary Neo-Lockeans' denial of the importance of the bodily self was as bad as that of radical mind-body dualists (Ayers 1991, pp. 287-288).

#### **4.2 Continuity without identity (Parfit)**

The main theses of Derek Parfit's survivalist theory were the following:

- (1) the self is a mutable entity with degrees of mental and bodily continuity,
- (2) the search for criteria of personal identity is futile. What matters is the question of personal survival, not personal identity.

Briefly we may say that Parfit's intention was to take the idea of the self dividing to its logical limits. This had important consequences for the problem of personal identity. Usually we would think personal identity as the continuity of an indivisible self through time. However, if the division of the self were logically possible (i.e. did not violate any basic classical logical principle), then the continuity of the self through time would not be conceived as personal "identity". The resulting "selves" of an original self would have a relation to it of continuity without identity. These selves would be survivors of the original self. Even while being different from their ancestor, they would preserve their psychological continuity through time. This continuity between the original self and its descendants would be a question of degree, not all-or-nothing. A descendant self could have more or less continuity with its ancestral, depending on the case.

Parfit's influence on the debate was important because he succeeded in identifying cases in which the question of identity

could not receive a straight answer. He showed that bodily and psychological continuity were logically possible without personal identity through time. Using brain fission and fusion cases, Parfit showed that this continuity was given in degrees, not according to the *tertium non datur* principle (either there is identity between *P1* at *t1* and *P2* at *t2*, or not), as in personal identity. In this way Parfit could suggest that the problem of identity was not as relevant as it seemed.

In Parfit's view, we did not have to suppose that the self was an entity beyond its bodily and psychological continuity. We could operate logically with the idea of a division of the self into descendant and future selves. We did not necessarily need to say that we remained identical to who we were in our childhood. We could thus do justice to the developmental process of ageing, in which we feel only continuity in greater or lesser degree with our past selves, but not identity. We did not have, therefore, to be perplexed by the issue "Will my current self be identical to my past self?" We did not have to answer this question according to the *tertium non datur* principle: yes or no. We could calmly answer: more or less. We could reflect about our lives, see how many selves they contained and see what was the degree of continuity between them.

The important point now is to see how Parfit used Williams' and Shoemaker's arguments to defend his idea of psychological continuity without identity. Parfit adopted Shoemaker's defense of psychological continuity as a sufficient criterion for personal identity in the case that it was not one-many. But Parfit also wanted to develop an account of the logical possibility of branching in psychological continuity. For that he had to examine brain fission cases.

In brain fission cases we assume that the two hemispheres of the brain are divided and later transplanted in the bodies of two other persons (one hemisphere for each person). Granted that each hemisphere may perform similar psychological functions, the original person would have become two different persons. We

would then have two separate persons in two different bodies. None of the resulting persons would be identical to the original, since they would be different among themselves and each would be as identical to the original as the other. But the two descendant persons would retain a relation of psychological continuity with the original self. In other words, we would have a branching of the causal chains between mental states considered by Shoemaker. There would be no identity, but only survival, guaranteed by the branched psychological continuity.

Thus, in the brain fission case, Parfit wanted to show that survival of the primordial self as two descendant selves was conceptually possible. He accepted Williams' argument that psychological continuity was neither logically nor empirically one-one, and that psychological continuity was not a necessary criterion because it could be complemented (in the absence of psychological continuity) by bodily continuity, which would then be sufficient to decide the identity issue. Therefore, Parfit actually disagreed with Shoemaker as to the necessity of the psychological criterion, although he accepted its sufficiency. Parfit's contribution began exactly at this point, since he was not satisfied with only defending non-branching psychological continuity as a criterion of personal identity through time. Parfit's idea was to logically explore what would happen if branching in psychological continuity were possible. Hence his interest in brain fission cases. Parfit wanted to account for this possibility of branched psychological continuity in terms of a new language which would render the debate on personal identity obsolete.

As far as personal identity was concerned, Parfit stuck to the psychological view. The important relation that guaranteed personal identity through time was that of psychological continuity (which included causal continuity). When we used the concepts of personal identity, Parfit believed we were trying to suggest such a psychological continuity. Thus, Parfit said that his was a view that favored psychological continuity ((Parfit 1978), p. 149).

Parfit then tried to respond to Williams' argument against the psychological view. For Williams, identity was a one-one

relation, but psychological *continuity* was not logically one-one. Hence, it could not be a criterion of personal identity. Parfit noted that some authors such as Shoemaker replied to Williams saying that it was sufficient that the relation to which one appealed to decide personal identity through time be always empirically one-one.

Parfit's suggestion, however, led to a slightly different answer. He still claimed that psychological continuity could guarantee identity when it was one-one. The difference was that if psychological continuity took a one-many or branching form, according to Parfit we would have to abandon our traditional concept of personal identity. He assured us that this possibility would not compromise the psychological approach to identity. On the contrary, this possibility would even be to its advantage.

This made possible a new defense of the psychological view of personal identity through time. Parfit admitted that judgments about personal identity were of great importance. What gave them importance was the fact that they suggested psychological continuity. That was why, according to the conventional view, when there was such continuity, we ought, if possible, to suggest it by means of an identity judgment. But if psychological continuity took a one-many form, as in the case of successful brain fission, no coherent set of identity judgments could possibly correspond to, or be used to suggest the one-many form of this relation. According to Parfit, what we should do, in such a case, would be to deny the importance that would be associated to an identity judgment and attribute this importance directly to each branch of the one-many relations resulting from brain fission. This case would then help to show that personal identity judgments derived their relevance from their ability to suggest mere psychological continuity and not from personal identity. Parfit regarded the relation of psychological continuity as being the foundation upon which our understanding of personal identity through time rested.

Parfit claimed that his argument employed a principle proposed by Williams. The principle was that "an important judgment should be asserted and denied only on importantly

different grounds" (Parfit, p. 150). He reconsidered the Charles-Guy Fawkes and Charles-Robert-Guy Fawkes cases in which Williams applied this principle. In the first case, Charles was psychologically continuous with the late Guy Fawkes, and in the second case, that of reduplication, Charles and Robert were psychologically continuous with Fawkes. Parfit said that according to Williams, if we took psychological continuity as a sufficient condition to speak of identity, we would say that one man was Guy Fawkes. But we could not say that both men were, although we ought to have the same grounds for asserting this, and this would violate the principle. Williams' solution would be to deny that one man was Guy Fawkes and to insist that the sameness of body was necessary to determine identity.

Parfit realized, however, that Williams' principle could provide another answer. We could imagine a situation in which we considered psychological continuity to be more important than sameness of body and have a man psychologically and causally continuous with Guy Fawkes. If he were so, it would violate Williams' principle to deny that he was continuous with Guy Fawkes, since we had the same important reasons as in the normal identity case. In the two men case, we again had an important reason. So far, Parfit agreed with Williams. The difference in Parfit's solution to the problem started from this point on. He suggested that we ought to remove the importance of identity judgments and attribute it solely to psychological continuity. We ought to say, as in the brain fission case, that each branch of the one-many relation simply survived, and this was in agreement with Williams' principle.

Even if psychological continuity was neither logically nor empirically one-one, it could be a criterion for identity, since we could appeal to the relation of non-branching psychological continuity, which was one-one. Parfit expressed the criterion as follows: 'X and Y are the same person if they are psychologically continuous and there is no other person who is their contemporary and psychologically continuous with the other'. He admitted that we would have to explain what we wanted to say with 'psychologically continuous' and say how much continuity would be required by the

criterion. But having done this, we would satisfy a sufficient condition of personal identity.

However, Parfit accepted that, in the absence of psychological continuity, bodily identity would be sufficient to determine personal identity. Therefore, psychological continuity, although sufficient, was not a necessary condition for identity. This would suffice, according to Parfit, to deal with the usual cases of personal identity

But Parfit went further. He admitted that as we recognize the logical possibility of psychological continuity not being one-one, we ought to say what would happen if it were not one-one, or else his account could be considered incomplete and arbitrary. He therefore intended to discuss the case of psychological continuity in the one-many scenario. His reply was that if psychological continuity took a one-many form, we ought to reconceptualize it and forget the conventional concepts of personal identity. He argued that many different relations were included in or were consequences of psychological continuity. We described these relations in ways that presupposed the continuous existence of a same person, but we could describe them in new ways, without making this assumption of identity ((Parfit 1978), p. 149-151).

This suggested a "wilder" thesis to Parfit. He believed that it would even be possible to think of life experiences in a completely "impersonal" way. But his main aim was to describe a way of thinking about our own personal identity through time that was more flexible and less deceiving than the way in which we usually did. This way of thinking would allow for a meaning in which a person could survive as two in the case of brain fission. An even more important aspect was that he treated survival as a question of degree, not all-or-nothing, as in personal identity.

While trying to estimate the degree of survival of the self, Parfit stressed that we ought to note the distinction between psychological continuity and psychological connectivity. Continuity was a transitive relation and was generally weaker between two or more selves, meaning simply that a self, for example, was a

descendant or ancestor of another self. In addition, it was all or nothing, not a question of degree. Let us suppose that a primordial self were subdivided in branches of ever decreasing descendant selves (first generation self, second generation self, etc.). The third generation self would remember the life of the second generation self, which would remember the life of the first generation self, which in turn would remember part of the primordial self's life. All these selves would maintain a transitive relation among themselves of mere psychological continuity. If X remembered a good part of Y's life, and Y remembered good part of Z's life, then X remembered good part of Z's life. With psychological connectivity transitivity did not occur and its existence was a matter of degree. Connectivity only existed when two selves had between themselves a direct relation between mental events, such as for example between memory and recalled experience, or between an intention and an intended action. The degree of connectivity would tend to be greater between nearer generations. If farther apart in the chain of descendants, then there was less connectivity. A 33rd generation self would certainly have continuity, but little connectivity if any at all with the primordial self. These two concepts, continuity and connectivity, must be kept in mind.

In this sense, personal identity did not have the significance it seemed to have. We could talk about our selves through time without having to assume identity, but thinking in terms of continuity and degrees of connectivity between our selves. Instead of talking about 'the same person' concerning our acts, we could talk of 'descendant selves' and 'ancestral selves' when thinking of psychological continuity, and speak of 'past selves' and 'future selves' when talking about psychological connectivity. Since the concept of continuity did not admit of degrees, it would be better to think in terms of connectivity in speaking of successive selves. The term 'self' could then be used to signify the highest degree of psychological connectivity. When direct relations between two successive selves were few, then we could say that one of the selves was not me, but only a "past (or future)" self. This would have to be explained further by the speaker in terms of degrees of difference



between habits, memories, etc. of his or her self and those of the past or future self.

The argument in favor of survival without identity had, however, some problems (McGinn 1982, pp. 115 and 121). In the brain fission case, Parfit argued that we could survive as two persons, each hemisphere constituting a descendant self in a different body. It was clear that it did not make sense to speak of identity. Identity was a one-one relation, while survival could happen among many. The problem was that in the brain fission case there were three possibilities admitted by Parfit: (1) there was no survival, (2) there was survival of only one of the hemispheres, and (3) the self survived as two distinct persons.

As to possibility (1), Parfit insisted that we had previously agreed that we would survive if our brains were transplanted with success. He claimed that persons had in fact survived with half of their brains destroyed. Hence, he concluded, we could survive if half of our brain were transplanted with success and the other half were destroyed. This being the case, then it did not make sense that we did not survive if the other half were also transplanted with success. This would be absurd. As Parfit famously put it: "How could a double success be a failure?"

Regarding possibility (2), Parfit initially admitted that perhaps one success was really the best possible result of brain fission. Perhaps we would be only one of the resulting persons. But if each half of my brain was exactly similar, just so as each resulting person was, it did not make sense to survive as only one of two persons when both were similar. If similar, both should have the same chances of survival. If both were similar and could survive, both should be able to survive, not only one. Therefore, Parfit concluded that the possibilities (1) and (2) - that we did not survive, and that we only survived as one of the persons - were highly implausible. He suggested that who accepted these possibilities must have ignored other possible solutions to the problem.

In rejecting (1), maybe the most medically plausible result, Parfit took it for granted that we had previously agreed that if the

brain were transplanted with success, then there would be survival. But now this was precisely the problem at issue. Parfit added to this argument the empirical fact that some persons survived with only one hemisphere. From this he wanted to infer that we could survive if half of our brain were transplanted with success. But if the transplant worked for one half, then it should work for the other as well. Therefore, Parfit saw option (3) as "highly plausible", for there was no reason why one hemisphere should survive and the other not.

### **4.3 Objections (McGinn and Williams): Mind-Brain and Conceptual Revision**

It was unclear, however, how we could extrapolate the fact that a person survived with only one hemisphere to the claim that a transplant of a hemisphere would be successful. Besides, Parfit's assumption that if the transplant were successful, then there would be survival could be empirically and contingently false. It seemed more plausible that there would be brain death after such a transplant, or at least great neuronal damage. What was worse, Parfit did not explain his criteria of "success" and how we could decide them in brain fission and transplant cases. There was no way to know up to what point the self survived this procedure. Parfit argued as if this were simple and not controversial.

This questionable use Parfit made of neurological cases extended to the operation against epilepsy. Parfit chose only one of the interpretations that had been given to hemisphere separation. But the patients' reactions could be interpreted in several ways, without a definite conclusion (cf. (Nagel 1979), "Brain bisection and the unity of consciousness" in *Mortal Questions*). There were also difficulties concerning the limit as to what each person could accept as intuitively plausible. Parfit saw no problem in the idea of a divided mind corresponding to the idea of divided hemispheres. But McGinn, for example, was critical of this automatic association between the division of the brain and the division of the self. Parfit was assuming too much in this issue, and not only in the concept of the self as a brain, but also in the idea that any brain *section* could be a *centre of self-consciousness*. For him, we could really imagine a

self with two separate spheres of consciousness, with the person having two simultaneous experiences but without simultaneous consciousness of both.

The second problem, pointed out by McGinn, was that Parfit's argument went against our usual concept of the self. No matter how strong the argument for the changeability of the self might be, our conventional concept of the self assumed a certain minimal unity of the self through time. No matter how great the difference between my younger and older selves was, there must necessarily be something that guaranteed my unity through time. If it were not so, then we had better to seriously consider the possibility that all us after all had multiple identities. However, we were then faced with huge complications when trying to identify the authors of particular actions or to attribute moral and legal responsibility for reprehensible acts. As B. Williams noted, the very idea of a future self became problematic without identity.

Williams argued that if we imagined our descendant selves as our children, then to commit suicide would be a contraceptive procedure, since we would be killing not only our own self, but also our descendant selves. But this was absurd, for when I commit suicide, I am clearly killing only my self, not my descendants (Williams 1994, p. 6). A society made up of persons with multiple identities would be extremely problematic. The idea of mutable and multiple selves would bring more problems than solutions. This line was well developed by Williams. He pointed out that persons have interests, desires and life projects that they try to realize throughout their short lives. These interests, desires and projects not only provide the reason for our interest in the future, but also constitute the necessary conditions for us to have such a future.

For Williams, by speaking like Parfit in terms of future selves that are like descendants, we neglected the central issue of our relation towards our interests, projects and future. For example, to commit suicide and leave descendants were two separate decisions that did not imply one another. We could have sons or daughters before committing suicide. A person might even decide to

do this deliberately, for certain reasons. But she might refrain upon realizing she would not be there to help her children survive. Williams reminded us that that is how we would normally regard the relation between suicide and parenthood.

However, when we started considering our supposed future selves, they would have the strange property of, on the one hand, be born only from the brain fission of their ancestor, while, on the other hand, the suicide of their ancestral would abort them completely. For Williams, the analogy became even worse when we had to conclude that the failure of our projects, and our subsequent suicides, would also kill all our "descendants", although they were descendants that would be born only with the fission and death (for Williams) of our ancestral self. This confused the issues of suicide and parenthood. In the first, we had to consider whether, our project having failed, we ought to commit suicide, killing our ancestral self. In the second, we had to decide if by means of brain fission we would have descendants with their own and different projects. Williams noted that the analogy confused the first kind of question with a question of the second kind, misrepresenting the importance of the first question for the theory of the self. Williams argued that suicide only made sense because we would be eliminating our future self, whose projects had failed. If the future self of a person were not another descendant self, but the future of his own self, then we could understand why this future must be eliminated with the failure of the project that kept the person motivated to live. But it did not make sense to *prevent the birth of descendant selves because the life project of the ancestral self failed*. Hence, from Parfit's view, according to Williams, suicide would be a strange kind of contraceptive act. The suicide case showed that our self was more fundamental exactly because it was he, and not his descendants, who would not be any more in the world after our self-destruction. It was the failure of the ancestral self's project that justified his suicide, and the abortion of the descendant selves was not necessarily related to this issue (Williams 1981, p. 11-12).

McGinn also agreed with Williams that to abandon our contemporary concepts of personal identity just to adopt Parfit's

view and terminology was not a feasible proposition. To accept the unraveling of the self in ancestors, descendants, past and future selves, and abandon the conventional view of the self we have today would require a wider consensus of our linguistic community. The adoption of such a radical revision would have to be based on something more convincing than merely to avoid perplexity while considering extraordinary brain fission and fusion cases.

## 5 Concluding Remarks

With hindsight, it seems to me that Bernard Williams' emphasis on bodily continuity was important and remains valid. Later developments, such as Shoemaker's non-branching causal memory, Parfit's branching descendant selves and the narrative identity differently proposed by P. Ricoeur, Alasdair MacIntyre, Charles Taylor and D. Dennett, if taken solely by themselves, would suffer from a neglect for the body which is ultimately the principal criterion for intersubjective control of statements regarding personal identity. As far as perspectivism is concerned, if priority is given to the first person perspective<sup>7</sup>, such criteria also suffer from reductionism, which is more harmful than helpful because it impoverishes the language game in which we try to make sense of ourselves as persons.

## References

- AYERS, M.: *Locke: Epistemology and Ontology. The Arguments of the Philosophers*. London & NY: Routledge, 1991.
- DENNETT, D.C. : *Brainstorms: Philosophical essays on mind and psychology*. Cambridge (MA): MIT Press, 1984.
- EKLUND, M.: "Personal Identity and Conceptual Incoherence". *Noûs*, 36, 3, pp. 465-485, (Sep., 2002).
- LOCKE, J.: *An Essay concerning Human Understanding*. Oxford:

---

<sup>7</sup> This is clearly not D. Dennett's case for he has understandably stressed the importance of heterophenomenology. My intention here is just to point out the insufficiency of psychological and narrative concepts of the self if reference to the body is missing.

Claredon Press, 1975.

MCGINN, C.: *The Character of Mind*. Oxford: Oxford University Press, 1982.

NAGEL, T.: *Mortal Questions*. Cambridge: Cambridge University Press, 1979.

PARFIT, D.: "Personal Identity". In *Philosophy of Mind*, J. GLOVER(ed.) Oxford : Oxford University Press, 1978, pp. 142-162.

ROS, A.: *Materie und Geist: Eine philosophische Untersuchung*. Paderborn: Mentis, 2005.

SHOEMAKER, S. : *Identity, Cause, and Mind, Philosophical Essays*. Cambridge: Cambridge University Press, 1984.

SIDER, T.: "Criteria of Personal Identity and the Limits of Conceptual Analysis". *Noûs*, 35, Supplement: Philosophical Perspectives, 15, Metaphysics., pp. 189-209, (2001).

WILLIAMS, B. : *Problems of the Self, Philosophical Papers 1956 - 1972*. Cambridge: Cambridge University Press, 1973.

WILLIAMS, B. : *Moral Luck, Philosophical Papers 1973 - 1980*. Cambridge: Cambridge University Press, 1981.

WILLIAMS, B. : "An Interview with Bernard Williams". *Cogito*, 8, 1, pp. 3-19, 1994.