

A NEO-HUMEAN BUNDLE THEORY: A REDUCTIONIST ACCOUNT OF
PERSONAL IDENTITY, CONSCIOUSNESS AND SELF-CONCERN

by

SINEM ELKATIP

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of
the requirements for the degree of Doctor of Philosophy,
The City University of New York

2009

© 2009

SINEM ELKATIP

All Rights Reserved

This manuscript has been read and accepted for the
Graduate Faculty in Philosophy in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Stephen Grover

09/09/09

Date

Chair of Examining Committee

Iakovos Vasiliou

09/09/09

Date

Executive Officer

Michael Levin

John Greenwood

Steven Cahn

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

A NEO-HUMEAN BUNDLE THEORY: A REDUCTIONIST ACCOUNT OF PERSONAL IDENTITY, CONSCIOUSNESS AND SELF-CONCERN

by

SINEM ELKATIP

Adviser: Professor Michael Levin

Although our understanding of mental states ordinarily assumes something that has those mental states, viz. a subject, or a self, it is far from clear what the subject or the self is supposed to be. While philosophers have been critical enough of thinking subjects in a certain way for instance as Cartesian egos, or as brains, they have not been critical enough of the need for subjects. Since there is no well-articulated theory of what subjects are really supposed to be, I contend that it is time to challenge the commitment to subjects in mental lives, and see if we can do without them.

Both Hume's bundle theory and Parfit's reductionism challenge the idea of subjects as separately existing entities, but Hume's atomistic view of perceptions makes it very difficult to further develop an understanding of the unity of mental states in a single bundle. And Parfit at times seems to give in to the idea of persons as subjects of experiences.

I contend that the idea of persons as subjects needs to be made explanatorily redundant, for reducing subjects to experiences and endorsing an eliminative approach towards subjects. There are various phenomena that seem to call for subjects to explain

them, which might justify the existence of persons as subjects. Hence, methodologically speaking, only if there were a way to account for phenomena that seem subject-requiring without referring to subjects, would it be reasonable to suggest that there are no such things as selves and that persons reduce to bodies and a series of interrelated mental states.

I call such accounts reductionist and start by arguing for a reductionist account of personal identity and unity of experiences. Then I argue for the possibility of conscious mental states without a conception of oneself as the subject of one's experiences. Finally in order to show that reductionism is compatible with basic human experiences, I argue for an explanation of why there would still be self-related concerns in the absence of selves.

Acknowledgements

I would like to begin by thanking my advisor Michael Levin. His willingness to listen granted me the courage and the mental freedom to entertain ideas thoroughly and to grow philosophically. It has been a privilege to rely on his philosophical insight when I needed it.

I would also like to thank John Greenwood, Steven Cahn, Catherine Wilson, Richard Sorabji for their helpful comments on earlier drafts of this work and the remaining members of my committee Stephen Grover and Rohit Parikh for their insightful comments. I am forever indebted to Ozan Hatipoğlu, for his untiring efforts to keep me both philosophically and personally on track. This dissertation is dedicated to him.

I know that my parents Selmin and Cengiz Elkatip had to endure through difficult times in my absence, and I would like to thank them for their loving support and patience. I would also like to thank my dear friend Ayça Boylu for the long telephone talks we had which provided me with a bit of fresh air to breathe during difficult times.

To Ozan Hatipoğlu,

Your presence made this possible and special.

TABLE OF CONTENTS

Introduction.....	1
Chapter 1. Undermining the Subjects.....	10
1.1 Mental unities.....	13
1.2 Ownerless experiences, the impersonal description claim and reducibility.....	22
1.2.1 Impersonal description of experiences and atomism.....	26
1.2.2 Subjectivity.....	31
1.2.3 Criticism of the non-substantial subject thesis.....	36
1.3 The unity problem.....	39
Chapter 2. The Bundling Problem and Spatio-Temporal Continuity.....	44
2.1 The Bundling Problem.....	46
2.2 Preliminary Assumptions.....	47
2.3 Spatio-temporal properties and the possibility of the transfer of psychological states.....	52
2.4 The Spatio-temporal criterion of personal identity.....	61
Chapter 3. The possibility of conscious mental states without self-consciousness....	72
3.1 Pre-reflective self-consciousness.....	74
3.2 Individual principle vs. state principle.....	80
3.3 Individual principle and the subject thesis.....	84
3.4 Four forms of consciousness.....	92
3.5 Four forms of self-consciousness.....	96
3.6 The neo-Kantian argument and reductionism.....	107

3.7 Parfit's selfless community.....	118
Chapter 4. A reductionist account of one's concern for one's future.....	125
4.1 Is explaining the special concern a responsibility of identity theories?.....	127
4.2 What is the main reason for the extreme claim?.....	129
4.3 Do identity theories that are not based on continuity succeed in explaining the special concern?.....	135
4.3.1 Schechtman's narrativity thesis.....	136
4.3.2 Criticism of the narrativity thesis.....	139
4.3.3 Personhood and Personal Identity.....	141
4.4 The reductionist account of one's concern for one's future.....	147
4.5 Is the special concern an illusion.....	153
Concluding Remarks.....	156
References.....	160

INTRODUCTION

This work was initially inspired by Frankfurt's (1971) theory regarding the distinguishing properties of persons from other sentient beings. Although persons seem to be the only ones that are capable of wanting to be different from what they are, an interesting thing to note about persons is how frequently they fail to do so. One wishes to quit smoking but can't, another wants to lose weight but can't. Some make New Year resolutions that never get to be fulfilled. Others want to be more successful at what they do, but somehow can't bring themselves to put in the necessary amount of effort. It was worth noting that a property that allegedly distinguished persons from other sentient beings has also been a property that they most often miserably failed to realize, and it seemed that those that managed to also *be* different from what they are, rather than only *wish* to be different were already disposed to be so. That is when I first started thinking that persons might be less than what they seemed to be.

My naïve interest was then directed at the question of free will and agency, like when a bank teller is threatened by a robber, in what sense is he really forced to hand over the money? Or for that matter, in what sense would he refuse to do so? My intuitions and thoughts about free will always seemed to me immature without an understanding of what persons essentially are and what they are capable of and perhaps more importantly what makes them capable of whatever it is that they are capable of. And that's what I finally turned my attention to.

Undoubtedly, mental states; i.e, beliefs, pains, thoughts, emotions etc. provide persons with the properties that distinguish them from other sentient beings. However, mental states themselves somehow do not seem to be enough. Only a few of us, if any, is

happy to think of himself merely as a thought, or a desire, or as a Humean bundle. Reid (1785/2008) captures this rebellious intuition against allegedly belittling persons by reducing their existence to a bunch of mental states when he says “I am not thought, I am not action, I am not feeling; I am something that thinks, and acts, and suffers” (p.109).

Reid's (1785/2008) view is further supported by the way we think about mental states. We do not think of mental states in the way we think about, say, weather states. We assume that where there are mental states, there are subjects of those mental states. I call this the subject thesis. Where there is a pain, there is something that suffers, where there is a belief, there is something that believes and this seems to be the most natural thing to say. Besides, a state is a state of something and it doesn't have an existence outside of the thing it is a state of. While this is true for all states including mental states, it does not immediately translate to the subject thesis. For instance a roll of thunder, insofar as it is a state of the weather, does not have an existence outside of the relevant medium of atmosphere. But we do not thereby posit a subject of the thunder, or we do not necessarily think of something that thunders.

However, one might argue that mental states are supposed to be different from states of the weather. For instance, as Strawson (2003) argues, if there is pain, but nothing that suffers from it, the incentive to seek relief from the pain will be meaningless (p. 281).¹ In short, it seems only natural to endorse the immediate inference from mental states to subjects of those mental states, and the inference is rarely second-guessed. The unique properties of persons are invested in their being subjects of experiences, or selves, even though no one seems to know what exactly these mysterious things called subjects

¹ On a related note, see Carruthers (2004) for an understanding of suffering without subjectivity.

are.²

While the influence of Descartes' *Cogito Ergo Sum* on the inference from mental states to subjects of those mental states has been undoubtedly huge, it is again Descartes' arguments that seemed to have muddied the waters, and gave rise to a certain sense of skepticism about subjects especially regarding what they are. Cartesian egos, which is one way of thinking about subjects, have long been the unwanted guests in our ontologies. The resultant skepticism about Cartesian Egos brings with it the question of what persons as subjects or selves are.³ An important and an inspiring reply can be found in Hume's famous (1739-40/1978) lines;

For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch *myself* at any time without a perception, and never can observe any thing but the perception. (p.252)

Perhaps there are no such things as selves, or subjects of experiences and persons are nothing but a bundle of mental states. Or perhaps there are no selves, only bundles of mental states mistaken for selves.

Although Hume's views are inspiring, the experience of oneself as a self or lack thereof alone is not a good reason to either accept or reject the idea of persons as selves. In the beginning of chapter 1, I provide more reasons to be skeptical of the separate existence of subjects of experiences. Given the long standing problems with the idea of subjects as Cartesian egos and the idea of subjects as brains, I consider a more tempting

² See Shoemaker (1985) for a criticism of this.

³ I should note here that I am using the terms subjects of experiences, subjects of mental states and selves interchangeably. My understanding of selves is restricted to selves as subjects of experiences.

way of thinking about persons as subjects and that consists in Crane's (2003) idea of persons as mental unities. I argue that it is not clear what mental unities are supposed to be, hence not clear what purpose they serve. As such, I suggest that it is time to question the immediate inference from experiences to subjects of experiences, and consequently of subjects. This lays the ground for seeking a metaphysical view of persons that does not take the existence of persons to consist in the existence of selves or subjects. I turn to Parfit for such an account.

It is worth noting here that my initial criticism of subjects is directed at subjects as separately existing, ontologically distinct entities. However, there are other ways of thinking about subjects. Firstly, subjects can be seen as grammatical constructs arising from the frequent usage of subject-predicate propositions in our language. I am neither concerned with nor bothered by the idea of subjects as grammatical constructs insofar as they have no metaphysical significance. Secondly, subjects can be seen merely as a logical requirement or a formal condition of certain phenomena, rather than a separately existing entity, which of course is the Kantian view. Even though Kant is not committed to the existence of a self, he takes the numerical identity of the self to constitute a formal condition of certain types of experiences such as the experience of persisting objects. I address these issues in chapter 3. Thirdly, one may argue that subjects just are sums of experiences, so even though they are distinct entities, they can be reduced to experiences, and this I address in chapter 1. To summarize, my criticism of subjects are intended not only for subjects as ontologically distinct entities, but also for subjects in the second and third sense; viz., as a formal condition and as a distinct mereological sum. It should also

be noted here that there is a rich literature regarding the absence of subjects in eastern philosophy which I am not competent to discuss.

According to Parfitian reductionism, the existence of persons consists in the existence of bodies and a series of interrelated psychological events or states and not in a further fact such as a self (Parfit, 1984, p. 211). In other words, persons are not selves, bodies and psychological states. Rather they are just bodies and psychological states. In the remaining sections of chapter 1, I articulate how reductionism with respect to persons should be understood. I argue that the implications of the reducibility of y to x depends on the kind of thing that y is. So, I contend that while the reducibility of persons does not necessarily give us good reasons to think that persons do not exist, the reducibility of subjects does give us good reasons to think that subjects do not exist. So my understanding of reductionism is slightly different from Parfit's reductionism in the following way. I start out thinking of persons as bodies, a series of interrelated experiences *and* as subjects of those experiences. Then I endorse an eliminative approach towards subjects because I contend that subjects reduce to experiences, and subjects are the kinds of things such that if, they reduce to something else, that is a good reason to think that there are no such things as subjects. I do not however endorse an eliminative approach towards persons because they reduce to bodies and a series of interrelated experiences as a result of the reducibility of subjects to experiences.

Parfit (1984) on the other hand works directly on persons and their reducibility. Since Parfit does not want to endorse an eliminative approach towards persons as a result of their reducibility, he ends up saying things like: persons are what have experiences,

therefore they are subjects and they are distinct from their bodies and experiences, but he adds that they are not thereby separately existing entities, that is, separate from their bodies and experiences (p.223). This way of handling reductionism subjects Parfit to certain inconsistencies which I bring out in chapter 3.

In order to argue for the reducibility of subjects to experiences and endorse an eliminative approach towards them in this dissertation, the idea of persons as subjects needs to be made explanatorily redundant, since there are various phenomena that seem to call for subjects to explain them, which might justify the existence of persons as subjects. So, methodologically speaking, only if there were a way to account for phenomena that seem subject-requiring without referring to subjects, would it be reasonable to suggest that there are no such things as selves and that persons reduce to bodies and a series of interrelated mental states.

A substantial part of this dissertation focuses on providing such subjectless accounts for what I believe to be among the most important phenomena that seem to call for subjects. I call these accounts reductionist accounts, or sometimes, impersonal accounts. To repeat, my purpose in this dissertation is to eliminate the explanatory power subjects might have, power that normally prevents us from being concerned about unnecessarily overcrowding our ontologies and our theories with these mysterious things that are subjects which we hardly know anything about.

In order to cover a sufficient variety of options, I choose three different ways to undermine subjects; firstly from a logical point of view, secondly from a phenomenological point of view, and thirdly from a psychological point of view. That is

to say, I focus mainly on three phenomena that allegedly call for subjects and suggest a reductionist account for each; firstly the unity of experiences of a single person and personal identity, secondly the possibility of conscious mental states without a conception of oneself as the subject of one's experiences, and thirdly the possibility of concern for one's future.

Chapter 1 lays the groundwork for the reductionist accounts to follow in the remaining chapters. After motivating skepticism about subjects and introducing the need for various impersonal accounts of phenomena that seem to call for subjects (to prepare a defense of reductionism), I note an important distinction between my account and Hume's bundle theory. Hume's bundle theory partly depends on his atomistic view of perceptions according to which perceptions can exist independently of each other, and are separable from each other. I, on the other hand, endorse a non-atomistic view of experiences according to which each experience is in need of a relational matrix⁴ of experiences, or more primitive forms of mental states. Much of the reductionist accounts to follow make use of this non-atomistic view of mental states.

In chapter 2, I provide a reductionist account of personal identity, by using a spatio-temporal continuity criterion for the unity of the experiences of a single person. This is another issue where I differ from Parfit (1984), who appeals to a psychological continuity criterion to explain how the experiences of a single person are united. In this chapter, I argue against the traditional distinction between psychological and physical continuity theories and I argue that thought experiments that are similar to the prince and the cobbler case (Locke, 1694/2008, p.44) and involve a transfer of psychological states,

⁴ I borrow this term "relational matrix" from John Greenwood.

are misguided.⁵ I argue that psychological states cannot follow a different spatio-temporal path than the one they originated in and I suggest a criterion of continuity based on the spatio-temporal properties of psychological states. This continuity criterion is based not on the content of psychological states and the various overlaps involved, but on the particular time interval during which they endure and the region of space they occupy.

In chapter 3, I argue in favor of the possibility of conscious mental states without self-consciousness; i.e., without an implicit conception of oneself as the subject of one's experiences. First, I examine the arguments suggested by Gallagher and Zahavi (2008) for the requirement of a certain form of self-consciousness for consciousness and argue that the argument for the requirement is not tenable. Then I discuss two approaches to consciousness, one inspired by Kriegel (2004) and another by Rosenthal (2005) and his transitivity principle. I suggest that it may be possible to account for conscious mental states without relying on self-consciousness if the latter approach is endorsed. Next I critically examine Kriegel's arguments for the dependence of consciousness on self-consciousness and point out the inconsistencies they lead to. Also in chapter 3, I criticize Cassam's neo-Kantian argument that in order to have the experience of persisting objects, one must have a conception of oneself as the numerically identical, persisting subject of one's experiences.

I end chapter 3 by discussing whether Parfit's (1998) suggestion that there could be intelligent beings that have experiences and thoughts like us, at least no worse than us, but with no conception of themselves as the subjects of their experiences, is tenable.

Among the psychological states that seem most unlikely for the members of a community

⁵ For a discussion involving such thought experiments see Williams (1970).

without a conception of themselves as the subjects of their experiences, but also important for them to be considered to be similar enough to us, humans, are states like regretting past deeds or hoping for a certain sort of future. I take it that such psychological states represent time sensitive forms of Frankfurtian type of second-order desires. And since Frankfurtian second-order desires involve a sense of being concerned about oneself – one's future and one's past for that matter – I zoom in on the phenomenon of one's concern for one's future in chapter 4. I take it that if it is possible to be concerned about one's future in the reductionist framework, other psychological states too would be possible. That is why in chapter 4, I suggest a reductionist account of one's concern for one's future.

Indeed, one of the criticisms of reductionism consists in its supposed failure to justify one's concern for one's future.⁶ This criticism is articulated in great detail by Schechtman (1996) among others. Schechtman suggests her narrativity thesis in place of Parfit's views and other personal identity theories that depend on continuity so that important psychological states of human lives can be justified, such as one's concern for one's future. In chapter 4, I critically examine this attack on reductionism and argue that it does not succeed. I also critically examine Schechtman's narrativity thesis and argue that while it may succeed in explaining why one is concerned about one's future, it does not do so as an identity theory. Since Schechtman suggests her narrativity thesis as an identity theory that justifies the concern, its failure to be an identity theory is shown to result in certain inconsistencies. Finally in order to support reductionism on independent grounds as well, I suggest a reductionist account of the concern.

⁶ Kind (2004) notes that such objections date back to Butler's criticism of Locke's account (p. 536-7).

1 UNDERMINING THE SUBJECTS

Introduction

Our understanding of beliefs, desires, memories, emotions, perceptions etc., ordinarily assumes something that believes, desires, remembers, feels, perceives etc. However there seems to be no consensus regarding what the nature of this thing that desires and remembers is. Because of this, philosophers have been careful about making inferences from mental states⁷ to entities that have the mental states.⁸ However, as Crane (2003) notes, it seems only natural to think there is a subject of mental activity when there is some mental activity, so it is not the idea of an entity that is the subject of mental activity that is cautioned against. Instead it is “an entity *of a certain kind*”, viz., a mental substance in the Cartesian sense that is cautioned against for long-standing reasons (Crane, 2003, p.230).⁹

I contend that the implications of rejecting the owner or the subject of mental states as a mental substance in the Cartesian sense have not been fully endorsed, one of which I believe to be reductionism. Reductionism is the view that the existence of a person consists in the existence of a body and interrelated psychological and physical

⁷ Mental states, psychological states, psychophysical states, experiences are used more or less interchangeably throughout the text. Only in chapter 3, does the difference between experiences and mental states become significant and I note there that I take experiences to be *conscious* mental states.

⁸ See for instance Putnam (as cited in Crane, 2003, p.229-30) who appeals to Dewey who in turn says, “Unfortunately, an influential manner of thinking has changed modes of action into an underlying substance that performs the activities in question. It has treated mind as an independent entity which attends, purposes, cares and remembers.”

⁹ As Kim notes “The idea of minds as souls or spirits, as entities or objects of a special kind, has never gained a foothold in a serious scientific study of the mind and has also gradually disappeared from philosophical discussions of mentality. . . There has been a near consensus among philosophers that the concept of mind as a mental substance gives rise to too many difficulties and puzzles without compensating explanatory gains” (as cited in Crane, 2003, p.231, fn 6).

events, or experiences in short, and not in a further fact (Parfit, 1984, p.211).¹⁰ In other words the existence of a person consists in the existence of a body and experiences, and not in addition in the separate existence of a subject of experiences or the self.¹¹ One reason why reductionism is not thought to be entailed by the rejection of Cartesian souls is that there remains another alternative that can serve as the further fact, viz. the brain.¹²

It seems that the mere possibility of thinking about the further fact in different ways, viz. as Cartesian souls or as brains suggests that the further fact does not have to be either one of them. So despite the problems with Cartesian dualism and materialism, there is still room for a belief in the further fact, since it is conceivable that the further fact is a fact about neither a Cartesian soul nor a brain but an entity of another kind. For instance Crane (2003) argues that it is in virtue of being a mental unity that persons as subjects of experiences, are different from their experiences, which would mean that they cannot be reduced to experiences.

I contend that insofar as the subject is taken to be something over and above experiences, and by “over and above” I mean ontologically distinct from, the nature of the subject of experiences cannot be articulated other than by saying that the subject is a Cartesian ego or a brain. Given the long standing problems with both ways of thinking about the subject, I suggest that it is time to reconsider the inference from experiences to

¹⁰ Throughout the text, the difference between psychological states and events is not significant so I use the words ‘states’ and ‘events’ somewhat interchangeably unless noted otherwise.

¹¹ See Strawson (2003) for a minimal understanding of the self as a thing that is a subject of experience.

¹² Although the thesis of reductionism and materialism are usually run together, it should be noted that neither entails the other. It is possible for a dualist to be a reductionist, defending the view that persons reduce to a series of interrelated physical *and* mental events, as it is possible for a physicalist to be a non-reductionist, defending the view that persons are not reducible to physical events, that there is a further fact but whatever the further fact is, it is also physical. See Parfit (1984, pp. 209, 241) and Korsgaard (1989, p.106).

subjects of experiences which has been taken for granted for so long, and perhaps too long, partly because of the influence of Descartes' ingenuous *Cogito, Ergo Sum*.

I call the thesis that whenever there is an experience there is a subject of experience the subject thesis. I borrow the phrase “subject thesis” from Strawson (2003, p. 280), but it should be noted that Strawson says that no inference about the nature of the subject can be made from the subject thesis, not even an inference that it is ontologically distinct from the experience (p.293). I also do not think that subjects are distinct from experiences, but I don't see the point of the subject thesis if one agrees that the subject may not be different from the experience. So when I argue against the subject thesis, it is a subject of a substantial nature, something that is considered to be different from the experience that I have in mind. Otherwise I do not see the significance of the subject thesis except under two conditions; one where the subject is taken to be a grammatical construct and that I have no problems with. Two, and this is more important, if the subject is composed of experiences in such a way that it just is a bundle of experiences, but it is also true that a bundle of experiences compose yet another object. I refer to this as the non-substantial subject thesis and address it in 1.2.3.

In what follows, I first argue against Crane's idea of persons as mental unities that are neither Cartesian egos nor brains. Next, I discuss what we should understand from reductionism and its implications. I argue that the implications of the reducibility of y to x depends on the kind of thing y is. Then, I introduce the idea of an impersonal description; viz., a description or an account of some phenomenon that does not refer to subjects or selves, meaning that it is a reductionist account. I argue that an impersonal description of

experiences does not imply and is not implied by the kind of atomism with respect to perceptions endorsed by Hume. I endorse a non-atomistic view of experiences and introduce the idea of a relational matrix of experiences, that is, a series of interrelated experiences, or more primitive forms of mental states that are associated with and are required by the occurrence of an experience. Next, I suggest a reductionist account of subjectivity. Then, I provide a criticism of the non-substantial subject thesis that I described above. Finally, I end by articulating the problem of unity; i.e., the difficulty in explaining what unites the experiences of a single person without referring to the person or presupposing the identity of the person and suggest a solution in chapter 2.

1.1 *Mental unities*

One reason why undermining the inference from experiences to subjects of experiences is not tempting is that persons seem to be excellent candidates to be subjects of experiences. There is a very deeply embedded intuition that we are not just “our” mental / bodily states, but the ones who have them and that we are not just grammatical constructs but things that have independent existence. Persons are also “good candidates for being things, at least in the sense of being particular objects, the bearer of properties” (Crane, 2003, p.231). Undoubtedly, bodies of persons make persons good candidates for being particular objects in the same way rocks, chairs and trees are good candidates for being particular objects. But what would make persons good candidates for being the kinds of things that have mental states? In other words, what would make persons subjects of experiences, granted that there are such things? It seems unlikely that the

thing that has experiences, that attends and cares can be just the body.¹³ While that pretty much leaves a Cartesian ego, there is quite a bit of resistance to thinking of persons as Cartesian mental substances for reasons I don't need to repeat here. It is thus important to take note of Crane's (2003) arguments for a way to think about persons as mental substances that are neither Cartesian souls nor brains, but mental unities.

According to Crane (2003), one way of thinking about persons as mental substances can be derived from one possible way of thinking about substances in general. Crane claims that substances are real unities. In other words, there are some entities that are unified in the way they are, not because of some "stipulation, postulation or theory-building" but because of some "underlying natural principle which is responsible for their organization, activity and boundaries" (Crane, 2003, p.246). This way of thinking about substances is exemplified by van Inwagen's (1990) theory of material beings. Van Inwagen (1990, p.81-2) says that only insofar as the activity of *xs* constitute a life can it be said that *xs* compose another object *y*. And it is for this reason that he thinks only animal bodies exist. Crane says that in the case of persons the principle of unity depends on "the way the person's mental life is unified, around a single point of view; and what would explain the identity of a person over time are fundamental connections between the elements of a person's mental life" (p.248).

Granted that there are indeed such real unities, one must note that persons could belong with other real unities such as the ones exemplified by material objects or living organisms. But to think of persons primarily as *mental* unities is also rightly motivated, since a person's ability to have mental states is not captured by one's thinking about

¹³ Unless panpsychism is true.

persons as having the same type of real unity that material objects or living organisms possess.

While having mental states is undoubtedly a very important attribute of persons, I contend that the idea of persons as mental substances in Crane's (2003) sense leads to a dead end unless the nature of the thing that is a mental unity can be articulated to be something other than a Cartesian ego or a brain. Insofar as a subject of mental states is posited and taken to be something over and above a mereological sum of mental states, it should be possible to say something about what it is. But for reasons that I discuss below it is only in virtue of the subject being a Cartesian ego or a brain that anything might be said about its nature. Given the problems with both, it might prove to be more fruitful if persons are not considered as mental unities but as real unities of another kind e.g. as a bodily unity that has, or rather happens to have, mental states. And due to the difficulty in explaining why the bodily unity of persons happens to have mental states,¹⁴ while other bodily unities such as rocks and chairs do not, it might prove to be even more fruitful to endorse a reductionist approach, say that persons just are bodies *and* interrelated experiences¹⁵ and give up on the idea of a unity of a substantial nature all together.

I should emphasize here that this is not a criticism directed against the idea of substances as real unities but against the idea of persons as mental unities. There are other reasons to argue against the idea of substances as real unities. As Crane (2003, p.245-6) notes, one may argue that there are no non *ad-hoc* principles of unity, or one may defend

¹⁴ Not that it would be any easier to explain how a mental unity is capable of having mental states without being circular.

¹⁵ It is worthwhile mentioning here that reducing a person to a body and a series of interrelated mental events does not require endorsing some sort of dualism – bodies on one side, mental events on the other. It just acknowledges the present difficulties in coming up with theories that relate the two.

an ontology of four-dimensional entities and a metaphysics of unrestricted mereological composition.¹⁶ While these are important issues, I do not have the space here to address them thoroughly and they are not really essential to what I am going to say.

My criticism of persons as mental substances in Crane's (2003) sense is directed at the ambiguity of what persons as mental unities are supposed to be. It is not immediately clear what it is about the idea of a person as a mental unity that is not captured by a mereological sum of some psychophysical states. If it is possible to show that whatever it is about the idea of a person as a mental unity that allegedly cannot be captured by a mereological sum of psychophysical states, can in fact be accounted for by appeal to psychophysical states only, then it may prove to be quite unnecessary to posit something that does the same job but whose nature we know nothing about, hence have no articulate understanding of how it does the job.

As mentioned before, the idea of persons as mental unities is nevertheless tempting, because it satisfies a deeply embedded intuition that we are not just the mental / bodily states but the ones who have them. And that in itself would be a reason why the idea of a mental unity is not captured by a mereological sum. It seems as if a mere sum lacking a substantial nature cannot act as the possessor of the things it is a sum of, but a mental substance in Crane's (2003) sense can.¹⁷ In addition, a mental unity is *prima facie* safely guarded against being understood as a Cartesian soul or a brain since nothing about

¹⁶ For different views regarding mereological composition, see van Cleve (2008) and Markosian (2008).

¹⁷ Well, perhaps a mere sum of *xs* cannot act as the thing that has *xs*, but it is possible that a sum of *xs* that is yet another object *y* is the thing that has *xs*. I will come back to this point. The reason why I don't address it here is that *y* composed in this way lacks substantial nature and right now I'm concerned with subjects that have substantial natures.

the idea of a unity implies that “the only way something could be a mental substance is either by being a Cartesian soul or by being a brain” (Crane, 2003, p. 247). Hence it escapes the problems involved with both Cartesian dualism, and materialism. It is the perfect candidate for the non-reductionist’s further fact.

But what is it? Despite attempts to refute reductionism, there has been almost no attempts at articulating what precisely the further fact might be. And how informative is it to know that it does not have to be a Cartesian soul, and that it does not have to be a brain either when we cannot say anything about what it is? We seem to rely merely on the idea that given beliefs, desires and thoughts, there must be something that believes, desires and thinks. Yet nothing can be said of this believer's or thinker's nature. The only thing that can be said is that it believes, desires and thinks etc. and that its beliefs, desires and thoughts are unified in a special way. And there is no understanding of how this particular unity is capable of mental states whereas other types of unities are not, and saying that it is in virtue of being a *mental* unity that this particular unity is capable of mental states is obviously circular.

The question is, how do we get to a separately existing and persisting entity that believes, desires and thinks from those beliefs, thoughts and desires being unified in a special way? More fundamentally, as also pointed out in Lichtenberg’s (1959) objection to Descartes, how do we get from mental states to the idea of something that has those mental states? Consider the following analogy. When there is a thunder, there are number of things happening that collectively give rise to the thunder. But one does not thereby conclude that there is *something* thundering. My purpose is to adopt a similar way of

thinking about mental states. For instance, for there to be a state of fear, there needs to be a family of things happening that collectively or jointly give rise to the fear but one does not thereby have to conclude that there is *something* fearing. And if one does, he is responsible for saying something about what kind of a thing it is that fears apart from the fact that it is the thing that has the fear. For Descartes it took no less than a Cartesian soul to fill the shoes of the thing that has the mental states. And for the materialists it takes no less than brains to do the same job. So if one wants to reject Cartesian egos and / or brains but keep the further fact, he'd better come up with an explanation as thorough as Descartes' about what the further fact might be.

On a related note, if it is granted that whenever there is an experience, there is a subject of experience, it needs to be pointed out that in so far as the subject is taken to be a separately existing entity, there could be millions of them, one for each experience, or one for groups of experiences, in the course of a single person's life.¹⁸ That is why if the substantial subject or the owner of experiences is thought to be singular and persisting in time, it would be something that needs to be further argued for. And whatever the further fact as a mental unity might be, it needs to be persisting and singular for experiences at different times to be unified in the manner required – that is unified around a single point of view. Now it is not clear how the further fact as a mental unity can be not only an independently existing but also a persisting entity without being so either by being a brain or a Cartesian soul. That is why although Crane (2003) says that mental substances as mental unities do not have to be Cartesian egos or brains, I contend that it is inconceivable how a mental substance is singular, independently existing and most

¹⁸ For related discussions see Strawson (1997, 2003) and also Rosenthal (2002).

importantly persisting without being either a brain or a Cartesian ego.

Now just as the non-reductionist is responsible for articulating the further fact in the way and for the reasons described above, the reductionist is responsible for explaining the unity of mental lives without the further fact. At this point one might recall Hume's (1739-40/1978) despair as expressed in the Appendix where he says;

But all my hopes vanish, when I come to explain the principles, that unite our successive perceptions in our thought or consciousness. I cannot discover any theory, which gives me satisfaction on this head. In short there are two principles, which I cannot render consistent; nor is it in my power to renounce either of them, viz. that all our distinct perceptions are distinct existences, and that the mind never perceives any real connexion among distinct existences. Did our perceptions either inhere in something simple and individual, or did the mind perceive some real connexion among them, there would be no difficulty in the case. (p. 636)

Hume mentions two alternatives; inherence of perceptions in something simple and individual and the mind's ability to perceive a real connection between perceptions.¹⁹ For my purposes only the first alternative is relevant, an alternative which basically amounts to saying that it would be easy to explain the principles that unite successive perceptions if they belonged to the same subject. Of course Hume is skeptical about there being a separately existing entity that is the subject of experiences, since when he looks most intimately into what he calls himself, he sees nothing but perceptions (p.252). But the reason why Hume is unwilling to endorse the inherence alternative as expressed in the

¹⁹ Also, for a variety of interpretations of Hume's despair, see Garrett, 1997, p. 163-185 in particular.

quotation above is different and is mainly due to his atomism with regards to perceptions, that is to say because he believes that “distinct perceptions are distinct existences” and they are “separable from each other, and may be separately consider'd, and may exist separately, and have no need of any thing to support their existence” (Hume, 1739-40/1978, p. 252). However my skepticism about there being a subject in which experiences inhere does not depend on an atomistic view of experiences, as will become clear later. In fact I deny atomism with respect to experiences for reasons that I discuss in section 1.2.1.

My purpose right now is to discuss another reason why the inherence alternative is not tempting regardless of Hume's dispositions. If the unity of experiences was explained by appealing to the fact that they inhered in the same subject, then it would be possible that the subject amounts to no more than a bearer of properties. But the idea of a substance as a bearer of properties is too weak because when substances are defined as bearers of properties, they amount to nothing in the absence of properties. What is supposed to distinguish substances from properties is thus lost, since there is no longer “a sense in which substances have independent existence and attributes do not” (Anscombe, as cited in Crane, 2003, p. 239). That is why Crane does not accept the idea of substances as bearer of properties in the first place and develops an understanding of substances as real unities.

In the light of the discussion above, one may argue that rejecting mental unities as subjects of experiences just because it is not easy to articulate what they are and there is a good chance that they are either Cartesian souls or brains is not a fruitful criticism. After

all we do have a long way to go before agreeing on a well-articulated theory of the mind. Hence there may be no harm in positing mental unities and distinguishing them from others as a start. There is after all a deeply experienced sense of unity when it comes to mental lives, and mental unities are unlike other kinds of unities. But without articulating the nature of the thing that is allegedly distinct from the mental states and has the mental states, the idea of a mental unity as a separately existing entity does not get off the ground. To repeat, for Descartes it took no less than a Cartesian soul to fulfill the role of the thinker given that he assumed it to be something separate from the thought. And he was right, that is to say, right in attempting to articulate the nature of the thing that is supposed to be distinct from mental states, because it is not tenable to attribute an independent existence hence an ontological significance to something without saying at least a few things about what it is.²⁰ Likewise, the non-reductionist materialists are right in thinking that it was the brain that had the mental states, that is to say, right in attempting to articulate the nature of the thing that is supposed to be distinct from mental states.

Given the problems with Cartesian souls and brains – and it is inconceivable to me what else a substantial subject can be even when it is considered as a mental unity – I merely suggest that it may be worthwhile to stop positing and start questioning the allegedly immediate inference drawn from mental states to subjects of those mental states. Maybe it is time to be skeptical of making an inference from there being mental activity to not just an entity-of-a-certain-kind that has the mental activity but to an entity full stop. Philosophers have been critical enough of Cartesian Egos and of brains, but not

²⁰ Other than the reasons why it was posited in the first place.

of the need to posit a subject. Needless to say, undermining the inference is possible only if there is a way to account for various phenomena that are seemingly in need of subjects without referring to the subject. And it is my purpose to do just that starting with the phenomenon that most obviously calls for a subject; viz. experiences.

1.2 *Ownerless experiences, the Impersonal Description Claim and Reducibility*

If the inference from an experience to the subject of an experience is to be undermined, one way to start is by making the concept of an ownerless experience tenable, and that concept is tenable just in case it is possible to describe experiences without reference to their alleged subjects. After Parfit (1984), I call this an impersonal description (ID) of experiences. Citing Lichtenberg, Parfit characterizes the impersonal description claim in the following way; “because we are not separately existing entities . . . we could fully describe our experiences, and the connections between them, without claiming that they are had by a subject of experiences” (p.225). Generally speaking, I will refer to an account or a description of x that does not refer to a subject or a self an impersonal account or description of x .

Now one may argue that even if some phenomenon that seems to call for a y can be described and explained without reference to y , it cannot therefore be concluded that y does not exist. In other words, the reducibility of y to x does not always provide good reasons to think that there are no such things as y s. For instance, the reducibility of water to H_2O molecules alone does not give us a good reason to think that water does not exist. Applying the same principle, one might say that even if those phenomena that seem to

call for a subject could be made sense of without referring to a subject, it might still be true that subjects exist.

I contend however that the implications of the reducibility of y to x for what does and does not exist depends on the kind of thing y is and one's knowledge about it separately from x . In general, the purpose of reducing y to x is not necessarily the elimination of y per se. Instead, it is to find out more about y , and to see whether there is anything about y that cannot be captured or cannot be provided by x is a way of finding out more about y . If y reduces to x , it should be possible to know everything about y by appeal to x . If for instance one rejected the reduction of events to substances, there would need to be something about events that is not captured by substances, e.g. time. That is, if time was an intrinsic part of events and it was impossible to derive any information about time or time as it relates to events from information about substances, then it could be argued that there is something to events that is not captured by substances, in which case events could not be reduced to substances. In other words, granted that y reduces to x , it does not immediately mean that x exists and y does not. It just means that there is nothing about y that cannot be captured by x , but whether or not y exists is another question and cannot be determined only on the grounds that y reduces to something else.

For instance, regarding the reducibility of water to H_2O molecules, there are well-defined theories of how the hydrogen atoms are bonded to the oxygen atom in a single molecule of water giving rise to its particular geometry and the unique polarity which in turn determine the particular types of intermolecular forces between H_2O molecules which in turn determine the properties of water as we know it, such as its boiling and

freezing points and its capacity to dissolve certain substances. More importantly, these latter properties can be observed and measured even without the knowledge of the reduction of water to H₂O molecules since information about water is readily available without the information about H₂O molecules. Scientific theories further allow us to see if the properties derived from hydrogen and oxygen atoms bonded in a particular manner fit the observed properties of water.

However, when it comes to the reducibility of subjects to experiences, it is a completely different matter. Selves or subjects are mysterious things. Our knowledge of them is nowhere close to our knowledge of, say, water. We seem to have no direct experience of them and we have no way of articulating their properties even when they are taken to be brains or Cartesian egos. We seem to turn to them only when we need to explain other things like the having of experiences, the unity of our mental lives, or even things like one's concern for one's future. So if it were possible to give an account of these phenomena without referring to subjects, there would be no reason to turn to subjects and rely on their existence. As such, the reducibility of subjects to experiences unlike the reducibility of other things, does give one a good reason to think there are no such things as subjects.

While Hume's (1739-40/1978) skepticism about selves is inspiring, it is risky to rely solely on the fact that one does not experience the self to undermine the existence of selves.²¹ It is also tricky to endorse Hume's atomistic view of perceptions and apply it to experiences in order to share Hume's suspicions about the existence of an entity that is

²¹ Compare for instance Hume's denial of experiencing a self to Trefil's (as cited in Strawson 2003, p.283) confidence in experiencing a self when he says "one single fact remains. . . I am aware of a self that looks out at the world from somewhere inside my skull."

the subject of experiences, because it is not obvious that the perceptions can exist independently of anything else. But if the subject's explanatory job is also undermined as a result of the reducibility of subjects to experiences, then, there are solid reasons, on top of the skepticism, to think there are no such things as subjects.

It is important to note then that while the reducibility of persons to experiences and bodies does not give us a good reason to deny the existence of persons, the reducibility of subjects to experiences, considering the mysterious nature of subjects, does give us a good reason to deny the existence of subjects. This is one reduction where Occam's razor should be effective and prevent us from unnecessarily overcrowding our ontology. There is no independent knowledge and no articulate understanding of subjects to start with so extra caution must be taken when they are posited. Besides, the reducibility of persons depends on the reducibility of subjects, since that is how one can say that the existence of persons consists in the existence of experiences and bodies and not in a further fact such as a self.

Given various seemingly subject-needing phenomena such as the unity of a mental life, personal identity or one's concern for one's future, it would be wise to start by understanding what an impersonal description of experiences consists in. To repeat, an impersonal description or an account, whether it be of experiences, or personal identity etc., is a description or an account that does not refer to subjects. At times I refer to it as a reductionist account too. The purpose is to render subjects explanatorily redundant so that they can be cut off from our ontologies. Furthermore, the description of experiences without reference to subjects would remove a major objection to reductionism, which is

that experiences are individuated by their owners. But before seeing if ID of experiences is tenable, some clarification is needed regarding another relation viz. the one between impersonal description of experiences and an atomistic view of experiences which is a major part of the Humean bundle theory.

1.2.1 Impersonal description of experiences and atomism

Before discussing how the Humean bundle theory and the impersonal description claim compare, I'd like to briefly note that Parfit (1984), who articulated reductionism in great detail in *Reasons and Persons* and suggested an impersonal description of experiences as a part of his reductionism, rejected it in his later work (1998) and took it to be dispensable for reductionism. Ever since, the impersonal description claim as a part of reductionism has been unduly ignored and abandoned. Only recently, Behrendt (2003), tracing the debate between Cassam's (1989) neo-Kantian argument and Parfit's (1984, 1998) reductionist argument, showed that Parfit's (1998) rejection of the impersonal description claim is not as substantial as it may at first seem. I will come back to her reasons and the neo-Kantian discussion involved in chapter 3 and right now I will address another criticism of the importance of impersonal description.

Once again though it should not be forgotten that the reduction of persons (as subjects) to experiences is not a paradigm case of reduction like the reduction of water to H₂O molecules where there are well-defined theories of the reducer and the phenomenon being reduced. We don't know how experiences are related to each other and to the persons that they in part constitute. Even though theories that relate brain states to experiences are getting better, they too are nowhere close to being fully articulated. As

such, impersonal description is an important methodology of reductionism, since it constitutes a way to undermine the explanatory role of subjects and hence suggest an eliminative approach towards the further fact.

In a paradigm case of reduction like the reduction of water to H₂O molecules, scientific theories allow us to be confident about the existence of hydrogen and oxygen atoms and their independent properties – independent in the sense that they don't have to form water molecules to have those properties – such as their atomic masses or electron configurations. If there was an inventory of material objects of the world, it is conceivable that hydrogen and oxygen atoms would be on the list but not water while, given scientific theories, water's existence could be derived from what is on the list. However, it is very difficult to think of experiences in the same atomistic manner as we think about hydrogen and oxygen atoms and apply what little theory we may have.

It is important to note the difference between atomism about experiences and an impersonal description of experiences. As a result of Shoemaker's (1985) criticism of the impersonal description claim, atomism about experiences and the impersonal description of experiences are usually run together even though they are separate issues. Shoemaker (p. 446) notes that in so far as experiences can be described impersonally, their status reminds one of the status of Humean perceptions, characterized as being "different, and distinguishable, and separable from each other, and may be separately consider'd, and may exist separately, and have no need of any thing to support their existence" (Hume, 1739-40/1978, p. 252). Shoemaker contests this notion of experiences and says that experiencing involves its subjects "as intimately as a branch-bending involves a branch"

(1986, p.107). For this reason he rejects ID of experiences.

However, that experiences can be described impersonally does not entail that they have no need of anything beyond them to exist. To that extent, ID does not entail that experiences are ontologically just like Humean perceptions. Experiences need not be atomistic, and may presuppose other experiences. But whether or not they presuppose subjects is another question. Endorsing impersonal description of experiences does not imply atomism about experiences. One could think that a relational matrix of experiences does not entail the existence of a subject and at the same time deny the possibility of there being a single isolated experience. Similarly, one can think that there could not be an experience without a subject but think that there could be a subject with only one experience.

I endorse the former view, deny atomism and agree with Lowe (1996) who says mental states are “only conceivable as elements within relatively well-integrated mental economies” (p. 28-9). I also contend that Hume’s despair would have been relieved had he not been so strongly committed to the atomistic view of perceptions. That is to say, he could have argued for a less independent existence of perceptions according to which mental states require other mental states, and still denied the dependence of perceptions on a subject. An experience is the experience it is because of the functional role it plays among a network of other experiences. There would be nothing about an experience that would make it a belief rather than a desire if it is the only experience there is. For instance a representational state might be a belief, if it is the effect of states of the sensory system, and a desire if it gives rise to further states in the motor system. Taken in

isolation, the state would not be a belief or a desire.

Hence, while it may be true that experiences presuppose other experiences, or maybe more primitive forms of mental states, it is not obvious that as a result they presuppose subjects too, unless maybe what is meant by subjects just is the relational matrix of experiences. Only in that sense could it be said that the description of experiences conceptually entails subjects, but that would amount to saying that the description of experiences entails other experiences, hence is not threat to ID of experiences.

Nevertheless there are two issues that need to be addressed here. Firstly the reason I've given to deny atomism regarding experiences may not be good enough, because it may be argued that a relational matrix of experiences is good for only determining the *kind* of experience an experience is – whether it is a belief or a desire or a thought, etc. – and not whether there is going to be an experience. On this argument, an experience may exist independently in the Humean sense and only when we are interested in determining what type of experience it is that a relational matrix of experiences is required.

But it is essential to an experience that it is the type of experience it is, since the typing of experiences constitutes a way to individuate experiences, and if we could not tell desires from thoughts, or beliefs from pains etc., experiences would be individuated by the timing of their occurrence and the particular body they would occur in. Now while such a description of experiences may have its uses elsewhere – for instance in establishing a reliable, presuppositionless criterion of personal identity, as discussed in chapter 2 – it is not very useful when it comes to understanding what experiences are.

Besides, if it is not essential to an experience what kind of an experience it is, and if an experience is individuated by its owner instead of by time and space, then there would be no reason why a person's whole life cannot be considered as one single experience.

While that might be a viable position to take, it brings a whole new meaning to what we understand of an experience, one which I and I suspect most other people do not endorse.

So I take it that an experience does presuppose a relational matrix of other experiences (and / or more primitive forms of mental states) at least within the confines of a somewhat orthodox way of thinking about experiences.

The second issue that needs to be addressed also has to do with relational matrix of experiences. When I say that an experience presupposes a relational matrix of experiences, it might still be true that the concept of experience presupposes relations to other experiences sufficient to constitute a subject or that experiences presuppose a relational matrix of experiences but that's what a subject is. So the description of experience does conceptually entail subjects. I find this objection very much related to a reductionist account of subjectivity so that's where I'll take it up.

Coming back to the relation of impersonal description of experiences to reductionism, then, I take ID of experiences to be an important, in fact an indispensable part of reductionism. Even though the ultimate purpose of a Parfitian (1984) reduction is an ontological / constitutive reduction of persons to experiences, unlike theories explaining the constitution of water, there are no articulate theories of the constitution of persons. Neither are there well-defined theories of the ontological relation between persons and experiences partly because it is difficult to think of experiences, which are

along with bodies the constituents of persons, in the same atomistic manner as we think about the constituents of water. Therefore the success of the reduction can best be judged by way of whether or not some apparently subject-needing phenomena can be explained without reference to a subject. In other words, we need impersonal descriptions of apparently various subject-needing phenomena and at their core, of course we need an impersonal description of experiences.

So after denying the view that an impersonal description of experiences calls for an atomistic view of experiences and arguing that the success of the reduction of persons can only be judged by seeing whether it is possible to describe seemingly subject-needing phenomena without referring to subjects, I now turn to see if a very important aspect of experiences can be described without referring to the subject; viz. subjectivity.

1.2.2 Subjectivity

One of the reasons why impersonal description is an ostensibly abandoned aspect of reductionism is that if experiences can be accounted for without appeal to their subjects, without even the assumption that they are the kinds of things that are had by subjects, this gives way to the possibility of ownerless experiences, like pains with no sufferers and the notion of ownerless experiences is highly controversial. James (1890/1950) argues that the elementary psychic fact is not just this or that thought, but someone's thought, every thought being owned (p.226).²² As mentioned before Shoemaker (1986) argues that experiencing involves its subjects "as intimately as a branch-bending involves a branch" (p.107). The possibility of ownerless experiences has

²² The word 'thought' here extends to all types of mental states; beliefs, desires, pains, perceptions, etc.

also been rejected on the grounds that if there were pains without sufferers, then there'd be no reason to ease the pain (Strawson, 2003, p.281). But perhaps it is part of a pain-state that it calls for a relief, in which case the need to ease a pain would not necessarily be dependent on there being something suffering from the pain, but on the state itself.²³

There are others who are skeptical about the inference from thoughts to thinkers. For instance Carnap quotes Nietzsche who says "It is merely a formulation of our grammatical habits that there must always be something that thinks when there is thinking and that there must always be a doer when there is a deed" (as cited in Bercic, 2004, p. 299).

One reason why it is so tempting to immediately infer subjects of experiences from experiences is that subjectivity is an essential part of experiences. Without the subject's viewpoint as a part of the concept of an experience, thermometers can be said to experience heat and become hot or trees can be said to experience the presence of water or lack thereof and become thirsty. But we typically do not think there is any subjectivity involved in these cases. The tree is not thirsty and the thermometer is not hot. To borrow Nagel's (1974) famous expression, there is no what-it-is-likeness for the tree or for the thermometer.

The concept of an experience, however, essentially includes what-it-is-likeness. And what-it-is-likeness involves a subject for which it is going to be like whatever it is like. So to make the impersonal description claim tenable, one must give an account of subjectivity without referring to the subject. In other words, one should resist treating the subjective character of an experience as an unanalyzable property of the subject. I believe

²³ On a related note, see Carruthers (2004) for an understanding of suffering without subjectivity.

that one does not have to look far to see how this might be possible because it is something that is already suggested – but most of the time overlooked – by our very own concept of experiences.

Typically we attribute experiences to beings with a sufficient level of biological complexity, e.g. to those that have well-developed nervous systems.²⁴ That is, we don't think machines have experiences, but persons do. We don't think unicellular organisms have experiences, but dogs and cats might. We don't think flowers and trees have experiences but dolphins and chimpanzees probably do. This suggests that there might be a way to understand subjectivity in virtue of the interactions between simpler – i.e. simpler than a full-blown experience – states which give rise to experiences when complex enough, and in turn between experiences themselves. If there is such a possibility, maybe we can understand the what-it-is-likeness of experiences without the what-it-is-like-for. In other words, although the subjective character of experiences is typically anchored in the subject's viewpoint, there might be another way to make sense of the phenomenal character of an experience that does not necessarily call for the subject.

This way of thinking about subjectivity requires denial of atomism about experiences. As stressed before, an impersonal description of experiences does not entail an atomistic view of experiences and I assume that experiences presuppose relations to other experiences. Subjectivity can then be said to consist not in the subject's viewpoint but in the fact that each experience is associated with a unique relational matrix of other

²⁴ See Damasio (1999) for a comprehensive account of what level of complexity gives rise to subjectivity as experienced by an organism.

experiences. In other words, an experience is the experience it is and has the phenomenal character it has as a consequence of the particular relational matrix of experiences it presupposes.

Consider a belief that it will rain at time t_n . Consider further that this belief arises as a result of an observation of the sky at time t_{n-1} . A belief that it will rain at t_n presupposes a perception of the sky at time t_{n-1} along with many other psychophysical states. Among these are past sky-observations, some of which were followed by confirmations, i.e., observations that it did rain, and by disconfirmations, observations that it did not rain which eventually lead to the belief that it will rain at t_n only when the sky-observation is in a particular way at t_{n-1} .

Likewise consider a present concern about a dentist visit that will take place in the future. It is a consequence of prior dentist visits, stories about dentist visits, a particular tolerance of pain – which of course has the threshold it has as a consequence of prior pains – knowledge of the particular procedure that is to be applied during the dentist visit or lack thereof that there is a concern in the first place and has the particular phenomenal character that it has.²⁵

It would be interesting to study what types of interactions are at work in the relational matrix. However, I do not attempt to articulate the specifics of the relational matrix here. I think it is an important question and requires a comprehensive study of mental states and consciousness, which I cannot undertake here. Perhaps the interactions are of a causal nature or maybe they are deterministic. It is important to note though that

²⁵ If at this point there are those who are skeptical about a reductionist account of one's concern for one's future, please see chapter 4.

my discussions in the remaining parts do not rely on the interactions being deterministic. This is not to say that the interactions in the relational matrix can't be deterministic, but it is important that reductionism does not depend on the truth of determinism.

Given the intentional character of psychological states, the examples above suggest that each experience about x is associated with a unique relational matrix of x -related experiences (or more primitive forms of mental states). It is possible then that subjectivity of an experience consists in the very uniqueness of the relational matrix. That is to say no two experiences of two different persons are the same because the relational matrix each one is associated with is different from the other. I give a detailed account of the uniqueness of each person's bundle of experiences in chapter 2. For the moment it should suffice to say that the uniqueness consists in the particular spatio-temporal path of a person's life.

Once impersonal subjectivity is made tenable, I return to the objection aforementioned viz. even if the fact that an experience presupposes other experiences, it might still be true that the concept of experience presupposes relations to other experiences sufficient to constitute a subject or that experiences presuppose a relational matrix of experiences but that's what a subject just is, viz. a relational matrix of experiences. I call this view, that is to say the view that there are subjects but they are nothing but a bundle of experiences so can be reduced to experiences, the non-substantial subject thesis. If the non-substantial subject thesis is true, supposedly, the description of experiences would conceptually entail subjects.

1.2.3 Criticism of the non-substantial subjects

I have a couple of things to say regarding this particular objection to the impersonal description of experiences. Firstly, if an experience presupposes a relational matrix of experiences, and as a result its description entails a relational matrix of experiences, then yes, I guess it could be said that a description of an experience entails a subject if one insists on calling a relational matrix of experiences a subject. But I don't see why that would be an objection to the impersonal description of experiences because it would still be possible to describe an experience without reference to a subject.

Secondly if a subject is a relational matrix, there would be one for a belief that it will rain and things associated with it, another for a concern about dentist visits and things associated with it, etc. That would amount to a large number of subjects during the course of a person's life. While that may sound odd, there is no reason why subjects cannot be these odd kinds of things, so that in itself may not be a problem.

There is a related problem, though. By agreeing to multiple subjects during the course of one's life, one faces the question of how the subjects of a single person are unified as the subjects of one person. This was a question concerning experiences and the non-reductionist answer was that they belonged to the same subject, the kind of subject that is over and above experiences.²⁶ Non-substantial subjects on the other hand, are not over and above experiences. They are just bunches of related experiences. What is important here is that we have the exact same questions after introducing subjects as relational matrixes of experiences and that suggests that this non-substantial, supposedly

²⁶ And the reductionist answer is given in chapter 2.

reduction-friendly conception of subjects does not serve much purpose.

The defender of the non-substantial subject and hence the opponent of the impersonal description of experiences might reply to this by saying that each experience presupposes not just a relational matrix of experiences that is related to what the experience is about but a relational matrix of all experiences. So a belief that it will rain will presuppose not just different observations of the sky when it rained in the past but also say a childhood memory of stealing apples. Now charitably understood, there is a sense in which this might be true, viz. if determinism is endorsed, however I do not think that the opponent of the impersonal description would want to rely on the truth of determinism to refute the possibility of an impersonal description of experiences, especially when, as a consequence of determinism, an experience would need to presuppose not just a relational matrix of experiences but a whole state of the universe. This consequence helps only in blurring the unity of experiences of a single person.

Now one should not turn one's back on the big issue here by getting preoccupied with multiple subjects. The big question here is undoubtedly one of mereological composition, what it is that makes a bunch of *x*s compose yet another object *y*.²⁷ While I don't have the space here to address this question, I'd like to point to some of the absurdities that result from thinking that subjects are composed in this way, that is to say, thinking that a group of related experiences compose yet another thing which is the subject.

In the beginning of this chapter I discussed the difficulties with the conception of

²⁷ See for instance Simons (2000), also van Cleve (2008) and Markosian (2008) for a discussion of related topics.

a subject that is over and above experiences. It seems to me that given the difficulties with the substantial subject, the substantial nature is removed from the subject and replaced by the claim that the subject is nothing but a relational matrix of experiences as a final resort to save the subject as the thing that has mental states, as the thing that thinks, desires and fears, etc. I believe that this is because the concept of a subject is highly invested in its being a 'haver' of experiences. It is built into our notion of experience that it must be had by something, which is precisely what I have been trying to undermine. But if one insisted on holding onto the relation between experiences and things that have experiences as a part of his understanding of an experience, but wanted to leave out the further fact about the subject (because that seems to lead to Cartesian souls or brains), it seems that he could take refuge in saying that subjects are nothing but a bunch of experiences because somehow there seems to be no contradiction in endorsing the following two claims at the same time: One, a subject just is a bunch of related experiences and nothing more and two, subjects have experiences, or experiences are the kinds of things that are had by subjects.

However, if one is tempted even slightly towards a restricted understanding of a mereological composition like van Inwagen's (1990), and frowned upon stipulations and theory-building compositions, he must agree that the subject is one such stipulative composition. I for one cannot see any reason why one would introduce subjects in one's terminology if he genuinely believed that it is nothing but a bunch of experiences unless he wanted to include the relation of having, i.e. the subject has the experiences, and experiences are the kinds of things that are had by subjects. Now it might have been

possible to be more tolerant towards introducing subjects in this manner if they had made the explanation of some related phenomena like the unity of a single person's mental life easier, but one is left with exactly the same types of questions about unity – subjects or no subjects – so there is not much room left for tolerance.

Besides, some relational matrixes of experiences are much simpler than others e.g. sensory perceptions at a very early age compared to New Year resolutions at a later age. Should these relational matrixes be considered equally subject-constituting? There is nothing wrong with saying that some bundles of experiences are more complex than others. But to take bundles of varying complexity to compose the same kind of thing, viz. the subject, does need a further argument.

Undoubtedly those who take the relation of ‘being had by something’ to be very important if not essential to their understanding of an experience will disagree. And given that my purpose is to undermine that very relation, at this time I can only go back to my attempts to provide a reductionist (or an impersonal) account, viz. explanation without reference to a subject or a self, of various phenomena and in the end suggest that the relation is neither important nor essential.

Therefore this might be a good time to start laying the groundwork for a reductionist account of the unity of experiences of a single person. To do that I will first articulate what the unity problem is.

1.3 *The unity problem*

One reason why subjects – substantial or non-substantial – are inferred from experiences has to do with the way experiences of a single individual are unified, “the

way the person's mental life is unified, around a single point of view" (Crane, 2003, p.248) which was also the motivation behind Crane's formulation of persons as mental unities. Without the person as the subject of experience, the obvious difficulty is to find a criterion by which one puts two co-personal experiences that are separated by time in the same group, like a memory from thirty years ago and a belief that it will rain today. It is possible to go wrong even with experiences that closely follow one another in time and content, like a desire for a cup of tea, and an intention to make tea, when the desire belongs to a master and the intention to the master's loyal servant.²⁸ It seems that with ownerless experiences the master's desire and the servant's intention could be wrongly united as the experiences of a single person.

The explanation of how experiences of a single person are unified consists in finding a criterion by which co-personal experiences can be bundled together that does not presuppose the identity of the person. So it is in effect a question of personal identity and actually a standing problem. One can trace it back to Butler's (1736/2008) and Reid's (1785/2008) objections to Locke's (1694/2008) account of personal identity. For Locke, what unifies psychological states that are separated in time is consciousness (p.45). In so far as one's consciousness extends back to some experiences in the past, they are the experiences of the same person. That is to say if S_1 remembers S_2 's thoughts and experiences with the same consciousness, S_1 and S_2 are one and the same person. Butler argues that Locke's account of personal identity presupposes one's identity since it is only our own experiences that we can remember (p.100).²⁹ Reid's objection on the other

²⁸ This example is from Sorabji (2006, p. 272).

²⁹ To address this objection, one might introduce the concept of quasi memory: where one may quasi-

hand shows how we retain our identity even though our memories disappear in time. He gives the famous example of a retired general who remembers himself as a brave officer but has no memory of himself as a young boy stealing apples (p.114-5). However the brave officer at the time remembers himself as a young boy stealing apples. Hence by transitivity, the retired general and the young boy should be the same person, but presumably, according to Locke's account of identity they are not because the retired general has no direct memory of the young boy and what he did.

Contemporary theories of personal identity tackle these problems by making use of psychological and / or physical continuity between the psychological and / or physical states of a person.³⁰ The psychological continuity criterion of identity for instance consists in there being a sufficient number of strong overlaps and connections between psychological states. The connectedness or the overlap supposedly arises from the kind of relation that exists between a desire and an intention that follows the desire, a memory and the experience it is a memory of, etc. On this view, the retired general and the young boy are the same person because even though the retired general has no direct memories of the young boy, he is psychologically connected to the brave officer who is in turn psychologically connected to the young boy.

But there are different forms of psychological connectedness. Both the retired general's and the young boy's psychological states could be connected with other people's psychological states and just as strongly. As Korsgaard (1989) notes too, "I may do something because I intended it; but so may my lawyer, secretary, executor, or slave"

remember things that never happened. See Shoemaker (1970).

³⁰ See Noonan (1989) and Perry (2008) for an overview.

(p.107), not to mention other forms of psychological connectedness as a result of which I may do something because *I think* I want it whereas it is actually my mother who desires it.

Another difficulty is the *lack* of unity between some experiences of a single person. For instance, there is almost no reason to think that a person's belief that it will rain tomorrow and the same person's childhood memory of sneaking out for a game of soccer should hang together in any special way. On the other hand, there is every reason to think that two different persons' psychological states such as a father's will to quit smoking and his child's desire for him to do so, hang together in a special way.

As such there needs to be a criterion or a set of criteria by which the psychological connectedness between one's own states can be judged to be different from the connectedness between his states and others. This criterion or the set of criteria should not presuppose the identity of the person, nor should it presuppose a separately existing soul or a Cartesian ego unless we want to give up the reductionist view. To go back to the example of the master's desire for tea and the servant's intention to make tea, there needs to be a way to tell the difference between the master's desire and the master's intention and the master's desire and the servant's intention.

In criticizing the reductionist point of view Campbell (1994) says;

Ordinarily I think of my experiences as mine. So how am I to achieve an impersonal way of thinking and talking about them? It is often supposed at this point that the reductionist must hold that experiences can in principle exist unowned, so that it is simply a contingency that all my experiences are in fact

bound together as the experiences of a single individual. (p.160)

The reductionist however does not have to agree that it is only by mere chance that the experiences of a single individual are bound together in the way they are. I think it is important for the reductionist to explain the unity between the experiences of a single individual. This explanation effectively consists in finding a criterion by which the experiences of a person are unified without presupposing the identity of the person, hence addresses the old, standing problems of personal identity. I suggest a way to do this in the next chapter.

2 THE BUNDLING PROBLEM AND SPATIO-TEMPORAL CONTINUITY

Introduction

Most of our psychological states are closely related to the belief that we remain the same person throughout our lives. Yet, what it is that makes anyone the same person over time is far less clear compared to the certainty of the belief. Philosophers have tackled the question of personal identity over time in a variety of ways but have often overlooked the dual nature of the question. I contend that there are in fact two separate questions intertwined into the question of identity over time. Firstly, there is the question of what any particular person essentially is, and secondly the question of what establishes the continuity of this thing that one essentially is. It is possible to phrase the former question as: what makes someone the person he or she is, and it should be noted that this is a separate question from what personhood essentially is.

Regarding the question of what one essentially is, Hume's (1739-40/1978, p.252) famous denial of experiencing anything other than a bundle of perceptions when he looks "most intimately inside" himself has put pressure on theories of persons and personal identity that presuppose or refer to a self, i.e. to the existence of something over and above a group of related physical / psychological events.³¹ Undoubtedly the existence of a self can neither be confirmed nor refuted by there being an experience of it or lack thereof. Nevertheless, the controversy over having such an experience, not to mention the

³¹ It may be noted here that there is also no direct experience of things like the contractions of the heart or the functioning of kidneys, but that does not seem to put pressure on theories that explain how our bodies function. The difference of course is that there is scientific evidence for these phenomena, and it is possible to have an indirect experience of them.

controversy over both the nature of the experience granted we have it, and the nature of the self granted there is such a thing, does give us a good reason to see if we can explain personal identity without possibly overcrowding our ontology. That is, without referring to a self. Besides, I have offered arguments in chapter 1 that are inspired by but do not have to rely solely on Humean skepticism regarding the self.

One may raise the objection that having to refer to a self in giving an account of personal identity over time does not confirm the existence of a self. The self can be seen merely as a formal condition in the Kantian sense. Hence there would be no problem of overcrowding in our ontology in the first place. While this is a fair objection, it can also be argued that taking something as a logical requirement for the explanation of some phenomenon pretty much commits one to its existence, at least in the particular explanatory framework at hand. Either way, we'd be in a better position if we could explain personal identity without referring to a self, assuming that we are concerned about Occam's razor. So I take the inquiry to be worthwhile nevertheless and I address the idea of a self as a formal condition and the neo-Kantian arguments involved in chapter 3.6.³²

An account of personal identity over time that does not refer to the existence of anything other than a series of related physical / psychological events is reductionist. In what follows I will argue that if a reductionist account of personal identity is endorsed, one must appeal to physical continuity to provide a criterion of identity over time.

³² Especially that of Cassam's (1989).

2.1 The Bundling Problem

In order to understand why physical continuity is needed, one should first articulate what the problem with the reductionist account of personal identity would be if it did not endorse physical continuity as a criterion of identity over time. It is thus appropriate at this point to take note of Hume's (1739-40/1978) concerns. Although Hume denies observing anything but a bundle of perceptions when he looks inside himself (p.252), he is also extremely dissatisfied with his account of how these several perceptions are bound together, allowing us to ascribe an identity to them (p.636). To address Hume's concern properly one must also take note of the fact that Hume already assumes too much by taking a bundle as given, for it is not clear what particular type of binding when applied to multi-personal perceptions – that is perceptions that belong to more than one person – would actually give rise to causally discrete bundles, i.e., to different persons. As mentioned before, my desires can lead to my intentions, forming a mini-bundle, but they can give rise to others' intentions as well, resulting in another bundle.³³ It is also conceivable that all psychological states are bundled under one mental being.

The bundling problem for the reductionist view of personal identity lies in the difficulty of finding a criterion by which the psychological states of a single person are brought together without referring to or otherwise assuming a self. One obvious way of bundling is to take into account what the states are about. And this is something most psychological continuity theories of identity rely on, viz., the overlap and continuity in

³³ Such objections can be found in Korsgaard (1989, p.107) and also in Sorabji (2006, p.272).

content, including Parfit's (1984). For instance, if I have a desire for a piece of cake, I may form the intention of cutting a slice and eating it. Or if I have a memory of a day at the beach, I am likely to have had an experience of being at that beach on that day.

But without taking into account their respective subjects, psychological states cannot be grouped properly by merely checking what they are about. I may form the intention of doing something because I want it, but my friend can also intend to do it because I want it. My desire and my friend's intention would then wrongly be bundled together as constituting the states of one person. Hence proper bundling could not be achieved on the basis of what psychological states are about and their overlap and continuity in content. So we need something more than intentionality, or to put it more accurately, something other than intentionality.

I'm suggesting that one's resources in making use of psychological states in determining personal identity are not restricted to what those states are about, or whom they belong to. I further suggest that what seems to be the least informative set of facts about psychological states, which is where and when they occur, can actually turn out to be the most helpful as far as the bundling problem is concerned because spatio-temporal properties are not hostage to the problems that the reductionist faces. I will now briefly sketch a way to make use of spatio-temporal properties to address the bundling problem.

2.2 Preliminary Assumptions

I start out with the following assumption;

- 1) (The occurrence of) all psychological states are rooted in physical events.

It should immediately be noted that (1) is not the same as physicalism. (1) is necessary

for physicalism but not sufficient. Even though I take the occurrence of a psychological state to be rooted in physical events, I intend (1) to be neutral as to whether or not every aspect of a psychological state can be given a physical description. Physical events might conceivably have physical and nonphysical properties, so a psychological event can be a physical event without implying physicalism.

(1) does not claim for instance that phenomenal properties can be described physically, nor does it claim that they cannot be. Consider for instance the perception of a red chair. This state is necessarily rooted in physical events that involve the wavelength of the light reflected from the object perceived and the interaction between this light and retinal cells. Whether or not the phenomenal properties of seeing a red chair can be described physically in a way that would allow the super-scientist Mary to know what it is like to see red in her black and white room is beside the issue here.³⁴

Similarly, (1) does not claim that the physical description of psychological states captures what those psychological states are about, nor does it claim that such a description is incapable of capturing intentionality. To repeat, (1) is neutral as to such claims, hence is not the same as physicalism.

Consider a state of fear which has physical manifestations such as the dilating of pupils or tightening of muscles. Fear is an emotional response which is believed to be caused by an increased activity in the amygdala and it has been observed that people with damage to their amygdalae are incapable of fear, cannot recognize an expression of fear, and are far less cautious in their social interactions compared to physically normal

³⁴ The example of Mary the super-scientist belongs to Jackson (1986). Mary is a scientist with access to all the physical information there is, but she's been born and raised in a black and white room and has not seen a colored object in her life.

individuals.³⁵ It is difficult to think that fear – or any other emotion for that matter – has the capacity to cause bodily changes, while it is not rooted in some physical phenomenon. But again whether its phenomenology or what the fear is about can be given a distinctive physical characterization is beside the point here. Hence I take the rejection of (1) to imply acceptance of Cartesian dualism, according to which there are two realms that are obviously related without there being any obvious way of explaining how they might be related.

Based on (1), the following assumptions with respect to psychological states follow;

- 2) Psychological states endure during a particular time interval.
- 3) Psychological states occupy a certain region of space.

Because of my neutral stance towards physicalism, I cannot claim that the exact coordinates of psychological states can be given. And empirically speaking, such information is not available anyway even if physicalism were true. It is impossible to localize even physical events precisely, let alone psychological ones. For instance, nobody doubts that the collision of two automobiles is physical, but exactly what are its coordinates?³⁶ The center of mass of the two cars at the moment of collision? The regions of space occupied by all the post-collision fragments? Every region of space somehow affected by the sound waves emanating from the collision? Hence, the region of space in which psychological states occur can only be approximated.

For all practical purposes, this region can be restricted to brains, or to the brain

³⁵ For a related discussion see Damasio (1999, p. 64-5).

³⁶ I owe this example to Michael Levin.

and the spinal cord, or some combination of the parts of the body that are thought to be fundamental to the constitution of a nervous system.³⁷ To that extent, psychological states may be taken to be the states of nervous systems, again without implying physicalism. Now it should be noted that (3) can be also be interpreted as the consequence of another fact viz. the fact that states are necessarily states of something, therefore occupy the region occupied by the things that they are the states of. In this case, psychological states would occupy the region occupied by the nervous systems of which they are the states.

Whether (3) follows from (1) or from the fact stated above, viz., states occupy the region occupied by the things they are states of, may seem to be unimportant. However it is not. I take (3) as a consequence of (1) and not as a consequence of the fact that psychological states are states of a brain or a nervous system. While I do not deny that psychological states are states of a nervous system, if (3) depends primarily on this fact, one might be forced into endorsing physicalism and a more traditional physical continuity theory of identity. And a traditional physical continuity theory of identity runs the risk of presupposing identity by taking a particular brain as the basis of continuity. Besides, as I will discuss later in 2.3, I do take the psychological features to be constitutive of one's identity, but I contend that the significance of a certain aspect of psychological features of identity is overlooked, which is the spatio-temporal continuity of psychological features. That is why I resist the traditional distinction between physical and psychological continuity theories.

In short, my understanding of (3) follows from (1). And *then*, for all practical

³⁷ Either way, common sense tells us that this region of space occupied by the psychological states can be taken to be bound by the particular body of the organism, or minimally by the particular brain.

purposes I take the approximated region of space occupied by the occurrence of psychological states to be bound by nervous systems.

While (2) and (3) are rather uncontroversial, they do give rise to some skepticism about deeply embedded intuitions regarding what would happen to a person whose psychological states are transferred to another body, as exemplified in Locke's (1694/2008, p.44) famous discussion of the prince and the cobbler, not to mention of the numerous thought experiments in contemporary discussions of personal identity. According to such thought experiments, psychological states are free to move around like passengers on a hop on hop off bus and if the psychological states of person *A* are transferred to body *B* while the psychological states of *B* are transferred to body *A*, *A* is taken to continue to exist in *B*'s body and vice versa. As such, when whatever it is that captures the prince's consciousness occupies the body of the cobbler, the prince continues to exist and what would be the cobbler to an outsider's eye is now the prince.

It should be noted that there *are* objections raised against these thought experiments,³⁸ but these objections most of the time do not address the *possibility* of psychological states' following a different spatio-temporal path than the one they originated in. Rather, they address the issues that arise given the possibility of the transfer, that is to say issues like split brains and fissions etc. I, on the other hand, am more interested in the actual possibility of the transfer itself. I will now explain why (2) and (3) are at odds with such thought experiments and undermine the possibility, usually granted, of transfer.

³⁸ See for instance Williams (2008) and Nagel (2008).

2.3 Spatio-temporal properties and the possibility of the transfer of psychological states

According to (2) and (3), given that psychological states endure during a particular time interval and occupy a certain region of space, it naturally follows that they will be in a particular spatio-temporal relation to the environment. And insofar as these regions are approximated to nervous systems or brains and since nervous systems are supported by organisms that interact with the environment from a particular orientation, some psychological states will inevitably be related to each other in the way they are as a result of the specific spatio-temporal location. For instance, an intention to get something from the shelf occurring in a particular body is in a particular spatio-temporal relation to the environment which influences other psychological states that figure in how far one reaches or whether an additional intention to get a stepping stool is formed. The importance of the coherency between psychological states as a consequence of their spatio-temporal relations with the environment within an existing body should not be overlooked when considering the conceivability of the transfer of psychological states from one body to the next.

Since the occurrence of a psychological state is dependent on a body – and if one denies that, as mentioned before, one is a dualist to start with – the possibility of the transfer of psychological states depends at least on some medium of a bodily nature that can sustain these psychological states, unless there is a reason to believe that psychological states, after their occurrences, transcend their bodily nature and are stored in some non-physical accommodation – which again is essentially a dualist picture. That

is why on one hand the process of transferring psychological states resembles a brain transplant during which the brain is kept alive when it is outside both bodies, and then placed into the recipient. On the other hand, it needs to be emphasized that psychological states are not literally things or entities, like brains, that can exist between nervous systems. A state is a state of something and it doesn't have an existence outside of the thing it is a state of. The question is of course whether the "it" a psychological state is a state of, is a self or an organism, or both. While this is an important question, my main purpose is restricted to seeing if personal identity can be explained when psychological states are taken to be states of an organism and organism only, and nothing further, insofar as they are taken to be the states of something.

Thought experiments that test our intuitions about personal identity seem to acknowledge the simple fact that a state is a state of something, because the scenarios in these thought experiments most of the time do involve a brain transplant. Unfortunately not enough attention has been given to exactly how the psychological states are supposed to remain intact *after* they are transferred, which suggests that they are after all taken as literally things or entities like little brains rather than states of the nervous system. As long as the transfer is logically conceivable other significant issues are overlooked, especially regarding what might happen after the transfer – which is usually the most important part of any transplant.³⁹ A heart transplant would not be considered successful if the heart did not start functioning properly in the new body. By analogy, what happens

³⁹ Here, I am not referring to the many manipulative scenarios in various thought experiments about what might happen to the person in the future. There are plenty of those considerations. See Parfit (1984), or Williams (2008) for various discussions. I mean what will happen to the psychological states that are transferred.

after the transfer of psychological states should have some significance, especially if there is a good chance that the psychological states that are supposedly transferred are inconsistent with the new body, in a sense of “inconsistency” I will try to explain.

Once the psychological states of the prince are transferred to the body of the cobbler, it is because we believe them to persist that we take the prince to continue to exist in the body of the cobbler. We believe the prince will endure because we believe that his psychological states remain intact. But how can we be sure that they persist, or that they remain intact, especially when certain states are related to each other in the way they are because of the particular spatio-temporal relation they have to the environment? That they can be transferred does not imply that they remain intact or even continue to exist. And this is where the flaw of thought experiments that are considered to support psychological continuity lies, that is, in the over-ambitious inference from the mere conceivability of a transfer to the unproblematic persistence of the states that are transferred.

Suppose that my psychological states are transferred to another body. Among those psychological states are certain beliefs like that I am short and that I am a female. Suppose further that the new body is that of a very tall man. Obviously my beliefs about my height and gender, and the states they normally give rise to can no longer be sustained in this new body since the particular relation these psychological states have to the body that they are embodied in is now completely different, in fact not only different but inconsistent. Imagine I’m in the same house that I’ve been living in for the past ten years – so I’m very familiar with where things are – and I need to get something from the top

shelf in the kitchen in this new tall body. My prior psychological states that figure in how far I reach, or the immediate intention to get a stepping stool before even attempting to reach, are going to be, at least momentarily, inconsistent with the way this new body needs to act. Surely it won't take long to get used to things, but that's precisely the point. That is, getting used to things consists in the termination of such *prior* psychological states, viz. states that supposedly remained intact after they were transferred.

Likewise, if a young person's psychological states are transferred to an old body, beliefs about how much weight he can lift, the types of things he can eat etc. will be incoherent with what he can actually do in the new body. Eventually, those psychological states that are transferred will be overridden by others. Now in an attempt to undermine the significance of this incompatibility, it may be argued that this is what happens over time in a single body anyway. But such changes in a singular body take place very gradually, keeping the coherency between psychological states significantly more intact compared to the coherency between states when one wakes up in another body.

The main idea here is that since psychological states stand in a certain relation to the body of the organism that they occur in and the surrounding environment and since the change in the body is too big and sudden for this relation to remain compatible, the transferred states will be replaced by other psychological states, and such replacement casts a shadow over the claim that these states continue to exist. Unless by transferring psychological states one means a momentary success in sustaining the existence of psychological states, and I'm pretty sure that is not what is intended, there is a lot at stake as far as the continuity of psychological states is concerned which undermines the

supposed success of the case of the prince and the cobbler.

Perhaps the defenders of transferability can restrict thought experiments to those cases where the transfer takes place between bodies of the same gender, around the same age, having approximately the same health etc. But given that this list can go on and on if one is picky enough about it, and that these constraints are all body-related, the very purpose of defending the priority of psychological continuity over physical continuity as a criterion of identity would be undermined if not defeated.

One may raise the objection that there are some mental states, non-sensory ones in particular, that are possibly immune to bodily changes, e.g. my belief that the sky is blue, that $2 + 2 = 4$, or that my mother's name is Selmin. Given that there is a large number of mental states of this kind that could survive an exchange of bodies, one's identity would be sustained despite the changes in or termination of other mental states.

One possible answer to this objection is that, due to the interaction of states, even the belief that my mother's name is Selmin would be experienced differently if there was an exchange of bodies. For instance, if I transferred into the body of a Frenchman, and I tapped into the French words for most things, there would be an experienced discordance with having a mother with a Turkish name.⁴⁰ Or if I transferred into the body of an Englishman living in London, while there may not be a substantial change in my belief that the sky is blue, there would certainly be a change in the frequency of my blue-sky-experience which is an important part of my belief. And it may take a few unpleasant sunburns to finally take note of the discordance between how sensitive this new pale skin is to sunlight and my prior, far less cautious experience with the sun.

⁴⁰ I owe this example to Michael Levin.

However, the same type of discordances would also be present if I, remaining in my body, learned French and moved to France or in the latter case if I became sick with some disease that also caused unexpected sensitivity to sun. None of this would have caused any difficulty as far as my identity is concerned. No one expects that a person will have precisely the same characteristics, same beliefs, same dispositions etc. throughout one's life, so no one considers such homogeneity a condition of one's identity. One might then wonder why such changes should constitute a problem in the case of exchanging bodies, when they are considered to be perfectly plausible changes in one's life with no threat to one's identity.

The problem with this criticism is that it misses the main claim that thought experiments involving an exchange of bodies rest on. One of the reasons for endorsing psychological continuity due to the alleged possibility of exchanging bodies without losing one's identity rests on the claim that psychological states remain intact and thus survive an exchange of bodies. Hence changes in one's dispositions, beliefs etc. matter for the possibility of exchanging bodies in a way that don't when there is a single body. While this may sound question-begging, it is not, because strict homogeneity is not a condition of one's identity but it is a condition for the success of thought experiments involving an exchange of bodies.

A proponent of the Lockean view might agree that given very weak physicalism, psychological states cannot be transferred between bodies. But he would add that if they *could* be transferred, then the recipient of those states would be the old person. If, somehow, the prince's states could be made to survive the transfer to the cobbler's body,

it would be obvious that the prince was now in the cobbler's body. So still, conceptually, one would be using a psychological and not a bodily criterion of identity.⁴¹ If the prince is not in the cobbler's body, it is because his psychological states cannot be transferred to the cobbler's body.

I do agree with the proponent of the Lockean view. It should be noted that my purpose is not to undermine the idea that psychological features of a person are constitutive of one's identity, or rather constitutive of what makes a person the person he or she is but to point out that psychological continuity is rooted in physical continuity. Hence I disagree with the view that psychological states of a person – especially those that are related to one's identity – can trace a different spatio-temporal path from the one a person's body does. To that extent I contend that the traditional distinction between psychological continuity theories of identity and physical continuity theories of identity is misguided. If anything, they should be taken as complimentary views and not opposing ones.

If the problem of identity is to be properly addressed, two questions need to be answered; firstly what should be taken as the identity-forming constituents? Like other psychological continuity theorists, I take the answer to be psychological states. So I agree that if the psychological states of *A* were transferred into *B*'s body and vice versa, and if they really survived intact in their new bodily surroundings, one would find person *A* in *B*'s body and person *B* in *A*'s body.

Coming back to the examples of mental states that seem to be immune to bodily changes, it is still not clear how one's thought that $2 + 2 = 4$ would be any different, let

⁴¹ I thank Michael Levin for making me realize the significance of this point.

alone significantly different, in another body. It is worth mentioning here that some psychological states are more relevant to one's identity than others. As mentioned before, one's identity over time incorporates not only the continuity of the person that one is but also the psychological features that make someone the person that he or she is. It is likely that thoughts like ' $2 + 2 = 4$ ', or 'a square has four sides' are much more commonly shared and less significant when compared to other mental states that make someone the person that one is. Hence, while there does not seem to be a good reason why the thought $2 + 2 = 4$ would be inconsistent with other mental states in a new body, there is equally no good reason to take such a thought to be significant to one's identity. I therefore contend that the possibly unproblematic endurance of such mental states can be ignored.

If a substantial percentage of the transferred states disappear in time because they are inconsistent with the potential psychological states of the new body, the fact that there has been a transfer is no longer significant. At most it supports the existence of a metastable state. A metastable state is one where a system is not at its equilibrium point, or at its lowest possible energy under a particular set of conditions, but nevertheless remains somewhat stable for a certain amount of time. A supersaturated solution would be an example for a metastable state where for a time interval a solution contains more than the solvent can dissolve. But surely the purpose of prince-cobbler type thought experiments is not defending the possibility of metastable continued identity. Their purpose is to argue that the prince continues to exist – not that the prince momentarily exists at a metastable equilibrium and then ceases to exist – and as a result argue in favor of a psychological criterion of identity.

Beliefs, desires, mental dispositions, character traits, memories are the kinds of things that make a person the person that he or she is. However, as mentioned before, the psychological features that make someone the person he or she is are modifiable in a substantial way, if not eliminated, due to the possible inconsistencies between prior features and the new body. Hence psychological states should not be considered as logically distinct from the body that they are associated with.

A substantial portion of our psychological states are acquired from a particular orientation to the environment, rendering the locus of our being to be an indispensable part of our mental lives. As such my view differs from other psychological continuity theorists because of the answer I give to the second question of regarding personal identity, i.e., of what properties of identity-forming constituents provide an objective, non-identity-presupposing criterion of identity. I take the spatio-temporal properties of psychological states to provide such a criterion. To that extent, the view that I endorse does not consist in denying the significance of psychological features for personal identity. Rather it involves an important addition that seems to be overlooked, which is the significance of the continuity of the spatio-temporal path of psychological states for personal identity.

Either way, despite the problems with thought experiments of psychological continuity aforementioned, one fact remains. During a hypothetical brain transplant, following a certain region of space in time, psychological states trace a continuous path during the transfer from one body into the next. And this is what I want to come back to in order to address Hume's bundling problem. That is, putting the issue of the transfer

aside, I want to come back to the spatio-temporal properties of psychological states and how these properties can help with the bundling problem.

2.4 The Spatio-temporal criterion of personal identity

I will now introduce some definitions that will be useful in stating precisely a spatio-temporal criterion of personal identity. Given an experience X ,

$B(X)$ is the moment X begins.

$E(X)$ is the moment X ends.

$D(X)$: $[B(X), E(X)]$ is the time interval during which X endures.

It should be noted that the time interval $D(X)$ is a vector; it is not a scalar quantity like $E(X) - B(X)$. This is important because two separate experiences may both last three minutes but may be taking place at different times. And in order to tell one from the other despite their equal duration, one needs to take into account the interval of time and not just the duration. Next;

$A(X, t)$: the region of space occupied by X at a moment t .

The region of space-time occupied by X consists of all the points of space time (x, t) such that t is in the interval $D(X)$, and x is in $A(X, t)$.

Before saying anything further, a few words on the individuation of a psychological state X are needed. Recall that in the framework endorsed here, psychological states cannot be individuated by their content, so distinguishing a

psychological state from another one by referring to the fact that one is a *belief that it will rain* and another one is *seeing a house* does not work here. Otherwise, psychological states of different persons may be wrongly bundled together as the states of one person.

Therefore, our intuitions regarding what constitutes a change in a psychological state such that it can be individuated as a psychological state X, need to be challenged. Firstly, in this content-free framework, the timing of psychological states constitutes a change by definition and thereby a way to individuate psychological states. For the sake of argument, consider a person whose only psychological state is a belief that it will rain, and who has this belief for fifteen seconds. Typically we think that there is one psychological state that lasts for fifteen seconds, viz., that of the belief that it will rain and we think in this way by referring to what type of state there is and what the state is about. But in the content-free framework, such individuation is not possible. So in order to individuate what in our traditional way of talking about experiences is a *belief that it will rain*, firstly one needs the moment that the experience begins and the moment that it ends so that it can be distinguished from other experiences that last for fifteen seconds. Secondly one needs the region of space occupied by the experience so that it can be distinguished from other experiences that occur during the same time interval.

However, appealing to the region of space occupied by states or to the time intervals they belong to, may not at times be sufficient to accommodate certain other traditional ways of individuating experiences, especially when it comes to simultaneously occurring multiple experiences. Again, for the sake of argument, consider a person watching the sea, having an emotion and then he starts to watch a dog, while the emotion

continues. Typically we think there are three states in total: the emotion, the visual perception of the sea and the visual perception of the dog. And the emotion accompanies both perceptions. In the content free framework however, one cannot appeal to the differences between watching *the sea*, watching *the dog*, or having an *emotion*, so three states may collapse into one state with a longer duration.

Still, there might be a way to handle the differences in the content-free framework. Even with simultaneous mental states of the same person, appealing to the region of space might be enough, since simultaneous mental states like seeing something and having an emotion are instantiated in different regions of the brain. For instance, the visual image would be located in the visual cortex and the emotion would be located in the medial frontal cortex. And if simultaneous mental states of the same duration are states of different persons, then it is possible to tell one from the other since they are instantiated in different brains.

Now with respect to the differences between consecutive mental states of the same person, like watching the sea and watching the dog, things might be a bit more difficult, since such states are likely to occupy approximately the same region of the brain; viz., the visual cortex. However, I contend that by making use of spatial properties, it may be possible to think of them as two different states. It cannot be denied that the position of the state with respect to the external environment has changed when the focus is on the dog. Just as a vector has components, the occurrence of psychological states can too be considered to have similar components. Focal length and perspective can constitute such components. For instance the change in this particular case viz., the

change from the state of watching the sea to watching the dog is likely to be constituted by a change in the focal length rather than the perspective, granted that the sea and the dog fall in the same perspective. Even though the terms *focal length* and *perspective* may sound as if they are only visual-related components of states, they are not necessarily specific to the sense of sight. The perspective component can be defined over the angle from which the body is interacting with the environment and the focal length component can be defined over whatever it is at the center of that perspective. When the state is that of hearing, the perspective component can be said to determine the angle from which things can be heard, and the focal length can be said to determine which particular auditory input is at the center.

So, even though the physical occurrence of psychological states does not necessarily tell us what the states are about, and thereby makes it difficult if not impossible to tell e.g. D(X: visual perception of the sea) from D(Y: visual perception of the dog), when the perspective and the focal length are taken into consideration, the changes in the psychological states can be accounted for, eliminating the possibility of considering X and Y as one single state with a longer duration viz. D(X+Y). While this may prove to be a useful way to handle sensory states, things are obviously more problematic with non-sensory states.

I will not try to find a way to handle non-sensory states here. The discussion above is intended for those who might not be happy with the content-free framework and its possible failure in distinguishing certain psychological states from others, especially simultaneously occurring ones, while such states can usually be easily distinguished by

the traditional method of considering their content. However, it is worth noting that for some, simultaneously occurring mental states of the same person should be treated as one single state. For instance, according to Tye's (2007) conception of mental states, simultaneously occurring states such as watching the sea and feeling the breeze constitute only one state by definition.

Nevertheless, this failure, if one wishes to call it that, arises from individuating psychological states by the time interval during which they endure and the region of space that they occupy during that time interval. My purpose was to suggest that especially with growing scientific knowledge of the brain, it could be possible to appeal to the different regions of the brain or take into account perspective and focal length measures to accommodate the differences between certain psychological states of the same person. So ripping psychological states off their content does not necessarily leave us in the dark when it comes to individuating them. There are other tools that one can make use of and possibly achieve a form of individuating the results of which might resemble the results of individuating by content.

However more importantly, it should be repeated that my main concern is to give a *criterion of identity* which is based on physical / spatio-temporal continuity. Hence it is only natural to make use of psychological states in a more technical sense that does not immediately translate into a traditional conception of psychological states. But the technical sense; i.e., the sense according to which psychological states are individuated by their spatio-temporal location, is not meant to be exclusive of the traditional conception of experiences, neither is it meant to contradict it. It is true that where there

are two psychological states according to the traditional conception, there may be one according to the technical conception, but I don't take this as an inconsistency. Rather we have two different ways of describing some mental phenomenon, and in giving a criterion of identity that is based on spatio-temporal continuity, one needs to stick with the technical conception.

One way of handling psychological states such that they can be used as a criterion of identity is by way of defining an experience as the total state of the nervous system, which occupies a region of space; and then any change within the nervous system would define the time at which the experience begins and the time at which it ends. Alternatively, one might commit to a standard unit of time for the duration of experiences, such as three seconds which, according to Strawson (1997, p.414) who cites Pöppel, is an approximate measure of one's sense of *nowness*, or the traditional "specious present"⁴² and assume that every experience lasts exactly that long. But that might seem implausibly restrictive. If on the other hand, one measures time intervals by taking into account the changes of the nervous system, one does not have to assume that all experiences last exactly the same amount of time. One should also realize that if the criterion is not sensitive to change, that is, if the changes in the nervous system or the brain are not properly accounted for, one's mental life can be said to consist of a single state.

Since psychological states are defined not only by the time interval during which they occur, but also the space they occupy, it is important to see what can be said with regard to two states that occupy different regions of space during the same time interval.

⁴² For a definition of the specious present, see James (1890/1950) p.609.

Given my assumption that the exact coordinates of psychological states cannot be given, despite the advances made in brain studies, I am going to bring restrictions to the approximated region of space in which psychological states are instantiated and take that region to be the whole of the nervous system or the brain. Hence, even though simultaneous psychological states that occupy different regions can either be states of the same person, occupying different regions of the same brain, or of different persons, occupying different brains all together, henceforth I take such states to be the states of two different persons by definition, given that two brains cannot occupy the same space at the same time.⁴³

Coming back to the spatio-temporal continuity criterion, recall that the region of space-time occupied by X consists of all the points of space time (x,t) such that t is in the interval $D(X)$, and x is in $A(X,t)$. So, an experience over the next minute, assuming it is a single experience, is the sum of the space-time points that the brain, in which the experience is instantiated, occupies over the next minute. It can then be said that experience Y is connected to experience X if they share a region, or if they share a point at the same time. Consequently one can conclude that a maximal chain of these connected experiences form an individual or a Humean bundle and constitute the identity of a single person over time.

The question then is what does it mean for any two experiences to be connected, or for them to share a region or a point? For expository reasons, let me consider the case of the master's thirst and the servant's intention to get him water vs. the master's thirst

⁴³ If however one wishes to distinguish between the simultaneously occurring but different mental states of the same person like an emotion and a visual perception, then the different regions of the same brain or things like focal length and perspective measures must be taken into consideration.

and his intention to get water for himself. What is needed is some spatio-temporal overlap between the master's thirst and his own intention that does not exist between the master's thirst and the servant's intention. It is expected that the state of thirst and the state of intention are defined by different activities. Let X stand for the state of thirst, and Y stand for the intention. I'm saying *thirst* and *intention* for expository reasons only and strictly speaking we don't actually need to know that X is a state of *thirst*, and Y is a state of *intention*. All we need to know is that there is some activity of some nervous system that starts at a certain moment t_n and ends at t_{n+a} and another activity that starts at t_{n+b} and ends at t_{n+c} . Let

D(X) stand for the time interval $[t_n, t_{n+a}]$

D(Y) stand for the time interval $[t_{n+b}, t_{n+c}]$

One way there is going to be a spatio-temporal overlap between X and Y is if the following conditions hold;

- i) $b \leq a$, in other words if experience Y starts before experience X ends, or at exactly the same time X ends.
- ii) For all a, b such that $b \leq a$ there exists a time period $\Delta t \neq \emptyset$ such that $\Delta t \in [t_{n+b}, t_{n+a}]$ and a non-empty region $C \in (x, t)$ for all $t \in \Delta t$ such that C is contained in both $A(X, t)$ and $A(Y, t)$.

Since no two nervous systems or brains can occupy the same region during Δt , it can be concluded that if X and Y share the same region C for Δt , they share the same nervous system for Δt . In other words, X and Y are connected, meaning they belong to the same bundle.

The advantage of the spatio-temporal criterion for a reductionist view of identity is that it is not hostage to various types of psychological connectedness between different persons' mental states since it appeals to a spatio-temporal overlap which is basically a relation of time and space, and it does not require the existence of a further fact such as the existence of an enduring self.

Undoubtedly there are some objections that may be raised against the spatio-temporal overlap or continuity criterion. In the remaining few paragraphs, I will address some of these issues. Firstly the critic of the spatio-temporal criterion may have some reservations about the reliability of this criterion because of the complications involved with the timing of various psychological states. For instance as a result of deep sleep or coma, there may be gaps between consecutive states of the same person which prevent overlap. That is, if Y is an experience after deep sleep, and X is an experience before the sleep, $B(Y)$ falls neither within $D(X)$, nor at $E(X)$. What kind of a temporal criterion is to be applied with respect to such states?

Similarly, consider a case where A after having a thought at 10:00 in his armchair is knocked unconscious, and moved away to another city. The next psychological state of A after the thought at 10:00 in the chair, takes place far away. In the meanwhile, imagine that B comes into the same room where A was, sits in the same chair and has an emotion. The question is how is it possible for A 's psychological state after she regains consciousness to be connected to her last mental state before she lost consciousness? By the same token, how is it possible to prevent B 's psychological state from being connected to A 's last mental state?

This suggests the need to trace the physical medium subserving psychological states – which for all practical purposes has been taken to be the nervous system or the brain – in the absence of psychological states, which gives rise to a third condition;

- iii) In the absence of mental states, the spatio-temporal continuity criterion must appeal to the mere physical goings-on of the medium that subserves mental states – which for all practical purposes is the nervous system or the brain.

In other words, for any psychological state X, the nervous system or the brain which corresponds to the approximated region of space X occupies, must also be physically traced and made use of in case experience X is not connected to any experience Y in the forward light cone of one of its points. In fact, this is how we would be able to find the prince in the cobbler's body – at least momentarily – if the prince's psychological states were transferred to the cobbler's body after a successful brain transplant. It should be noted that tracing the medium that subserves the mental states does not just consist in tracing a lump of matter, but includes tracing an on-going minimal activity as well. This way one would not have to say that the mental life of an individual ends when he or she is in deep sleep or under anesthesia. Instead, it would end when even the minimal activity of the nervous system or the brain stops as a consequence of the organism's death.

In the light of the discussions above, one might wonder whether teleportation is possible on the given account of personal identity. In one sense, teleportation resembles the case in which a person is knocked down unconscious and is moved to another place. I have suggested condition (iii) for such cases; i.e., the possibility of tracing the brain or

the nervous system. The same suggestion goes for teleportation; hence, the answer to whether teleportation is possible given the spatio-temporal continuity criterion, depends on what exactly is going on during the teleportation. As a result, it would require a precise, scientific account of the teleportation. And it should be so, since in philosophical thought experiments such cases are simply taken for granted with no proper account of or explanation for the suggested extreme cases. When information is missing, such thought experiments run the risk of being more misleading than helpful.

All in all, personal identity over time can be accounted for in two ways; either by assuming that there is nothing to persons over and above a group of related physical / psychological events and spelling out the relation between these events that provides a criterion of identity over time, or by referring to the endurance of something over and above a group of related physical / psychological events, viz. by referring to a self. For reasons described in chapter 1, it is worthwhile to see if personal identity can be explained without referring to a self. I argue that it can insofar as psychological states are not considered to be logically distinct from the bodies that they are associated with. While psychological states are important because they are constitutive of what makes a person the person he or she is, it is the spatio-temporal overlap between psychological states and the spatio-temporal continuity of the physical medium that subserves the psychological states that provides a reliable criterion of personal identity over time.

3 THE POSSIBILITY OF CONSCIOUS MENTAL STATES WITHOUT SELF-CONSCIOUSNESS

Introduction

As mentioned in the beginning of chapter 1, our understanding of beliefs, desires, memories, emotions, perceptions etc., ordinarily assumes something that believes, desires, remembers, feels, perceives etc. So far, I have been trying to undermine this assumption by suggesting ways to account for various aspects of experiences such as subjectivity and unity without referring to a subject or a self. However, it may be argued that persons need to have a conception of themselves as the subjects of their experiences in order to have experiences in the first place. In other words, some form of self-consciousness may be required for consciousness.

There is a sense in which it would be absurd to think that one needs to be aware of himself as the subject of his experience in order to have an experience. For instance, one does not need to be aware of oneself as reading a book in order to have the experience of reading a book. So the kind of self-consciousness that is thought to be required for consciousness – if there is indeed such a requirement – is actually something much subtler than an attentive awareness of oneself engaged in some mental activity. Hence one needs to first articulate what this subtle form of self-consciousness might be before concluding that it poses a threat to reductionism.

There is another perhaps less interesting reason why the phenomenon of self-consciousness does not pose an immediate threat to reductionism, and that is that

consciousness of x never implies the existence of x since consciousness has to do with the way things appear and not necessarily with the way things are. Unfortunately, this criticism does not explain the phenomenon of self-consciousness, but explains it away and is not very conducive to developing an account of it, let alone a reductionist one. But it does mean that there can be an experience of self-consciousness without there being a self that is the object of that experience.

The question of whether self-consciousness is needed for consciousness ultimately turns on a question that is concerned with conscious mental states and how they differ from non-conscious mental states. This is a huge question, but most of the issues involved are not relevant here. My purpose is restricted to seeing if a conception of oneself as the subject of one's experiences is a necessary condition of conscious mental states. Let me also add here that I take experiences to be conscious mental states.⁴⁴

Whatever this subtle form of self-consciousness is, the reductionist can either deny that consciousness requires self-consciousness, or argue that the phenomenon of self-consciousness does not really call for a self or a conception of oneself as the subject of one's experiences. I intend to do both for two reasons. Firstly, I contend that there is no such thing as a subtle form of self-consciousness, hence I deny the claim that a subtle form of self-consciousness is needed for consciousness. Secondly, while I do think that there are some mental states like regret, Frankfurtian second-order desires, or self-related concerns, that seem to call for an explicit sense of self-consciousness, I do not think that the explicit sense of self-consciousness that they call for actually involves becoming

⁴⁴ This is not obviously so because for some a conscious experience is not a redundancy. For a discussion see Byrne (2004).

aware of a self.

In what follows, I first articulate what this subtle form of self-consciousness is and explain why the argument for it is not tenable. Then I discuss two approaches to consciousness, viz., the state principle and the individual principle, that attempt to explain what conscious mental states are. I then argue that if one endorses the individual principle, it is possible to account for conscious mental states without relying on self-consciousness. In the following two sections, I critically examine Kriegel's (2004) arguments for the dependence of consciousness on self-consciousness and discuss the difficulties involved to suggest once again that self-consciousness is not required for consciousness. Next, I consider the neo-Kantian argument that a conception of oneself as the subject of one's experiences may be needed in order to have the concept of enduring objects, according to which subjects are required not as separately existing entities but as a formal condition (Cassam, 1989). Finally I critically examine Parfit's (1998) suggestion that there could be intelligent beings with no conception of themselves as the subjects of their experiences, but have experiences similar to humans.

3.1 Pre-reflective self-consciousness

It is possible to trace the idea of no consciousness without self-consciousness – although not in so many words – as far back as to Aristotle (*On the Soul* 3.2) and his view that it is impossible to perceive something and not be aware that one is perceiving it, to which Sorabji refers as Aristotle's “most Cartesian remark.”⁴⁵ The idea can also be found in Locke's (1694/1975) saying “thinking consists in being conscious that one thinks”

⁴⁵ See Caston (2002) for a discussion of this. Caston cites Sorabji's remark on p. 759.

(p.115). Gallagher and Zahavi's (2008) survey of the phenomenological literature reveal considerable consensus that consciousness calls for a subtle form of self-consciousness (p.46-7).⁴⁶ It is agreed that the kind of self-consciousness that is required for consciousness cannot be the attentive kind where a person voluntarily directs his attention towards himself, because there are innumerable experiences involving no such attentive awareness of oneself engaged in the experience in question. The kind of self-consciousness that is considered to be required for consciousness is;

an intrinsic feature of primary experience. . . it is not thematic or attentive or voluntarily brought about; rather it is tacit, and very importantly, thoroughly non-observational (that is, it is not a kind of introspective observation of myself) and non-objectifying (that is, it does not turn my experience into a perceived or observed object). (p.46, their parenthesis)

It is called pre-reflective self-consciousness (p.46) and it is this form of self-consciousness that is "a constitutive feature and integral part of phenomenal consciousness" (p.50).

Unlike what the name might suggest, pre-reflective *self*-consciousness is not the awareness of a self as a mere being either, whatever that might be. It does not make sense to say that a mental state is conscious in virtue one's being aware of oneself. Rather, it is the awareness of the self in action as it were. That is to say, it is an implicit awareness of oneself as the subject of the mental state in question, in the absence of which the mental

⁴⁶ Gallagher and Zahavi (2008) cite Husserl who says that consciousness always involves a self-appearance, Heidegger who says every consciousness is also self-consciousness, and Henry who claims that experience is always self-manifesting (p.46). They also cite Sartre who says "This self-consciousness we ought to consider not as a new consciousness, but as the only mode of existence which is possible for a consciousness of something" (p.47).

state supposedly could not be conscious.

There are similar remarks in the analytical tradition too. Goldman (1970, p.96) says that the process of thinking about something carries with it a non-reflective self-awareness. Flanagan speaks of a “low-level self-consciousness involved in experiencing my experiences as *mine*” (as cited in Gallagher and Zahavi, 2008, p.194). So there has been some growing consensus with respect to pre-reflective self-consciousness.

As such, pre-reflective self-consciousness causes some tension for the reductionist. The reason for the tension is that on one hand pre-reflective self-consciousness is usually described as involving an implicit understanding of oneself as the subject of one's experiences, and on the other hand, it is considered to be an integral part, or a structural feature of the experience, hence not separate from the experience. Therefore, on one hand pre-reflective self-consciousness seems to assume the very relation between a subject and its experience according to which the former is the possessor of the latter, and on the other hand, the emphasis on the givenness of this conception in the experience undermines that very assumption and hence appears reduction-friendly. Actually this dual nature of pre-reflective self-consciousness is not just a difficulty for the reductionist, but it is a difficulty in itself since it suggests a contradiction. So it is worth considering what precisely the argument for pre-reflective self-consciousness is.

Gallagher and Zahavi (2008) note that there is no direct argument for pre-reflective self-consciousness. Instead, there is “an indirect argument by elimination and consists in a rejection of the two obvious alternatives” (p.55). The reasoning is basically

as follows. Firstly, phenomenologists deny that one could consciously experience something “without in some way being aware of or being acquainted with the experience in question” (p.55). While this may sound reasonable at first, one must note that what would really be reasonable is that experiencing something requires being aware of or acquainted with something, but not necessarily the experience.

Secondly, there is the idea that this awareness or acquaintance cannot involve an attentive way of becoming aware of oneself having an experience, because we have tons of experiences where we do not have to pay attention to our awareness of having an experience to have the experience. As mentioned before, I don't need to be aware of myself as reading a book in order to have the experience of reading the book. In fact, when one is intensely engaged in some activity, we often say “he's lost himself in it”.

The argument then is, if some awareness of the experience is needed and it cannot be of an attentive kind, then it needs to be an acquaintance with the experience, involving a conception of oneself as the subject of one's experience without objectifying the having of experience. It would be the kind of thing that is present even in infants who clearly have no explicit conception of themselves as the subjects of their experiences.⁴⁷ So, even though pre-reflective self-consciousness is a kind of self-consciousness, it is not really about an attentive way of becoming aware of oneself having an experience. Instead it is an implicit awareness of oneself as the subject of one's experience.

The problem with this indirect argument is that given its implicit premises, there

⁴⁷ For studies that suggest otherwise by arguing that infants have the ability to develop a sense of self even before mirror self-recognition tasks and false-belief tasks, see Gallagher and Zahavi (2008, p.206-208). But it should be noted that while such studies indicate some abilities, to think that it is necessary to develop a sense of oneself as a self to have these abilities is a far-fetched idea and remains to be argued for.

is a far more immediate and a plausible conclusion to be drawn than the one that is actually drawn. Consider the following premises;

- 1) No experience without an awareness of it.
- 2) Tons of experiences but rare awareness of oneself as having an experience.

The only reason why (1) and (2) are not flat out in contradiction is that there is an equivocal usage of 'awareness' in (1) and (2) in relation to the thing of which there is an awareness that is exploited in drawing the conclusion 'no experience without an *implicit* awareness of oneself as having an experience'. (1) refers to an awareness of an *experience*, and (2) to an awareness of oneself as the subject of the experience, or awareness of oneself as having the experience, or awareness of oneself as the owner of the experience. In this context they all mean the same thing and I will refer to them as awareness of I-in-the-experience for short.

This suggests that the argument for pre-reflective self-consciousness requires an additional premise to link the awareness of an experience to awareness of I-in-the-experience. And to be fair, Gallagher and Zahavi (2008, p. 55) do mention an additional premise. According to this premise;

- 3) First-personal awareness of one's experiences is a form of self-consciousness.

It can safely be said at this point that first-personal awareness of one's experiences is the awareness mentioned in (1), i.e., awareness of an experience, and self-consciousness is the awareness mentioned in (2), i.e., awareness of I-in-the-experience. The conclusion then follows that awareness of an experience, although it is a form of awareness of I-in-

the-experience, cannot be of the kind mentioned in (2) which is the explicit kind, so it should be of a different kind, viz., an *implicit* awareness of I-in-the-experience.

However, it is not obvious why first-personal awareness of an experience is a form of self-consciousness. It is important to note that using the phrase *first-personal* to qualify the awareness of an experience does not give one a reason to accept that it is a form of self-consciousness, at least not an immediate one. I contend that the significance of *first-personal* awareness of an experience consists in its being a *non-inferential* awareness. And it is definitely not obvious that non-inferential awareness of an experience calls for a conception of oneself as the subject of the experience, so it is not obviously a form of self-consciousness. It would need to be argued for.

If one assumes that experiences, and mental states for that matter, are the kinds of things that are had by subjects, then one might suggest that an awareness of an experience implies or is a form of an awareness of oneself as the subject of the experience one is having, since the subject is supposedly given with the experience as the thing that has the experience. And I say one might only suggest it, because even if the subject thesis were true, one could not immediately infer an I-in-the-experience awareness from experience awareness, because there is no logical connection between how things are and how they appear, and awareness of things has to do with how things appear. Still, the possibility of deriving I-in-the-experience awareness from experience awareness would be very much supported by the subject thesis.

Once you let go of the assumption that experiences are the kinds of things that are had by subjects – and I’ve already given reasons to reject this assumption and to think

there are no such things as subjects in chapter 1 and 2 – then there is no immediate reason to think that an awareness of an experience implies awareness of I-in-the-experience. So when the following two premises are reconsidered, it should be obvious that nothing follows from them.

1) No experiences without an awareness of the experience.

2) Tons of experiences and rare awareness of oneself as having an experience

(1) merely asserts that conscious mental states differ from non-conscious mental states in virtue of there being an acquaintance with the mental state in question. And (2) merely asserts that the phenomenon of explicit self-consciousness is rare. The aim of the reductionist then is two-fold; firstly to argue that the kind of acquaintance with the mental state needed for that mental state to be conscious – hence to be an experience – does not necessarily involve conceiving oneself as a subject. And secondly to argue that the rare phenomenon of explicit self-consciousness does not involve becoming conscious of a self. To do the first, I need to articulate the nature of the kind of awareness referred to in (1) in more detail.

3.2 *Individual principle vs. state principle*

Before articulating what an awareness of an experience is, a few preliminaries are called for. As mentioned before, I take experiences to be conscious mental states. I also assume that not all mental states are conscious and I'm using the terms 'awareness' and 'consciousness' interchangeably unless noted otherwise. Given these preliminaries, it is important to take note of the two different ways the term 'conscious' can be used. One refers to the *individual*, a person or an organism, becoming conscious *of* something, and

the other refers to a *mental state's being* conscious. This distinction is noted and explained in detail by Rosenthal (2005, p.4) and he calls the former transitive consciousness and the latter state consciousness. There is also creature consciousness which basically amounts to an organism's being awake and mentally responsive (p.41).

So (1) – no experience without an awareness of it – can be understood in the following way; no mental state is conscious unless an individual is aware of it in some way. In other words, it is in virtue of an individual's becoming conscious of a mental state that the state in question is considered to be conscious. Rosenthal (2005) calls this the transitivity principle(p.4).

Now compare the transitivity principle to another view according to which “the concept of consciousness-of presupposes the concept of consciousness” and “a person cannot be conscious . . . when none of her mental states is conscious” (Kriegel, 2004, p.184-5). According to this view an individual is conscious of something in virtue of being in a conscious mental state. According to transitivity principle however, “. . . being a conscious creature does not entail being in conscious states” (Rosenthal, 2005, p.4) and “A creature can be in mental states without being in conscious mental states, and can be conscious of external or bodily events without also being aware of its own mental states” (p.41). To emphasize the contrast between the two views I will henceforth call the transitivity principle the 'individual' principle and the other view the 'state' principle because there is a sense in which the former approach to consciousness attributes a priority to the individual's consciousness whereas the latter to the mental state's consciousness.

If one endorses the individual principle, the question of consciousness turns on a question of how an awake and mentally responsive organism gets to be aware of something, or rather how it gets to be aware of its mental states to be precise and it would call for a cognitive mechanism by which this process of awareness can be explained. As Rosenthal (2005, p.22, 46-7) notes too, this way of thinking allows the explanation of conscious mental states to be given in terms of non-conscious mental states, and then possibly mental states in terms of non-mental states, rendering a more fruitful line of explanation all together.

On the other hand, if one accepts the state principle, consciousness becomes an intrinsic feature of conscious mental states since transitive and creature consciousness both depend on it. And if one accepts that an individual is aware of things, or even just is mentally responsive and awake, in virtue of being in conscious mental states, it is not clear on what grounds one can make room for non-conscious mental states. However, given the existence of non-conscious mental states, the defender of the state principle is in need of either a property in the absence of which a mental state would not be conscious, or some other form of consciousness on which state consciousness would depend. Kriegel (2004) offers both. He argues that state consciousness in turn depends on another form of consciousness, a form of self-consciousness to be precise. And with regards to the property in the absence of which a mental state would not be conscious, Kriegel says that the “surface feature that guides us in classifying [a mental state] as conscious is . . . the property of being *first-person knowable*.” and that “... all and only conscious states are first-person knowable” (p.197).

It is interesting to note that these two claims, viz., the dependence of consciousness on a form of self-consciousness and the mark of a conscious state as first-person knowability, are very similar to the additional premise used in the argument for pre-reflective consciousness, viz., that first-personal awareness of an experience is a form of self-consciousness. I have already suggested that it is not obvious that a first-personal awareness of an experience is a form of self-consciousness, and that it remains to be argued for. Kriegel (2004) does meticulously argue for the view that an awareness of an experience implies awareness of I-in-the-experience and I will return to his arguments in 3.4 and 3.5.

But for the moment, I should note that Kriegel (2004) accepts that first-person knowledge is not what “defines the conscious, or captures its essence. Rather . . . we use it in order to home in on the phenomenon we want to talk about. It is what fixes the reference of 'conscious' and its cognates” (p.198). Gallagher and Zahavi (2008, p.53) also remark that pre-reflective self-consciousness is not a phenomenon which, when added to an otherwise non-conscious mental state, makes it conscious. Rather it is an intrinsic feature of conscious mental states rendering the explanation of state consciousness somewhat obscure with regards to why some mental states carry with them an awareness of I-in-the-experience.

I mentioned earlier that the argument for pre-reflective self-consciousness assumes that experiences are the kinds of things that are had by subjects, since without that assumption the claim that awareness of an experience is a form of self-consciousness would not be tenable. The defender of the state principle in suggesting that the mark of

conscious mental states is that they are first-person knowable, seems to rely on the subject thesis too. Indeed Kriegel (2004) notes that the only experiences that we can have first-person knowledge of are those that we have self-consciously, in other words, those that are accompanied by a conception of ourselves as the subjects of our experiences.

It should be noted that the reliance of both the state principle and pre-reflective self-consciousness on the subject thesis does not constitute an independent argument in favor of the subject thesis. To support this view, it may be worthwhile at this point to look at another approach to consciousness, one that does not seem to be in immediate need of the subject thesis. I take the individual principle to involve an approach to consciousness that is I-in-the-experience awareness free. As mentioned before, when the individual principle is accepted, the explanation for consciousness is fundamentally an explanation of how a creature gets to be aware of something. I contend that such an explanation does not rely on an I-in-the-experience awareness, and will explain why in the next section.

3.3 *Individual principle and the subject thesis*

It is worthwhile to start by repeating that according to the individual principle, it is the creature's awareness of *its mental states* and not of other things, and not of others' mental states, that is supposed to provide an explanation for state consciousness, and that it is possible for the individual to be conscious of things without necessarily being in the relevant conscious mental states.⁴⁸ As such, one should not be too hasty in suggesting that

⁴⁸ Rosenthal (2005) says "We are conscious not only of people and other objects, but of our own states of mind. I can be conscious that I want something to eat . . . And that is just what happens when these states of mind are conscious – when, that is, my wanting something or seeing or thinking about

this particular approach to consciousness does not rely on the subject thesis.

To first understand how one can be conscious of things without being in conscious mental states, consider Rosenthal's (2005) example of a person who is angry but is not aware of it. I will call her Matilda. Matilda's friend notices that Matilda is angry from the way she behaves and tells her. Matilda, counting on her friend's observation, may infer that she, herself, is angry without being in a conscious mental state of anger.⁴⁹ So while there is no experience of the anger, there is an inferential awareness of it, and it should be noted that this kind of awareness, viz., an inferential one, does not necessarily give rise to the actual experience.⁵⁰

This might be a good place to digress briefly and explain why for some, the phrase 'conscious experience' is not redundant.⁵¹ In the example above, Matilda is not experiencing anger but she is angry and acts like a person who is angry. That's how her friend can infer that Matilda is angry, anyway. But apparently, if Matilda is asked why she is acting in the way she does, she will probably be puzzled and say, "What do you mean? I'm not acting in any particular way!" And some might say that's because Matilda's experience of anger is not conscious. For all practical purposes, I have chosen not to refer to Matilda's anger as an experience, but I can see why some might prefer to refer to Matilda's anger as a non-conscious experience. The bottom line is that Matilda is not aware that she is angry but she acts like an angry person would.

At any rate, the individual's inferential awareness of her mental state is neither

somebody are conscious states" (p.310).

⁴⁹ In other words if someone asked Matilda how she is feeling she will not say "I'm angry".

⁵⁰ And it doesn't thereby mean that the inferential type of awareness requires Matilda to have a conception of herself as the subject of her experience.

⁵¹ For a related discussion see Byrne (2004).

necessary nor sufficient for that mental state to be conscious. It is for the same reason that the phrase 'first-personal' was used to qualify the awareness of an experience before, and why it was claimed by Kriegel (2004) that all and only conscious states are first-person knowable. Prima facie, the need to explain how one becomes non-inferentially conscious of one's own mental states for those states to be experiences may seem to support the requirement of pre-reflective self-consciousness for experiences, since becoming non-inferentially conscious of *one's own* mental states seems to require an understanding of those mental states to be one's own. And what better way is there to have this sense of ownership than to conceive oneself as the subject of one's experiences? However, as mentioned before, I contend that becoming conscious of one's own mental states in a non-inferential way does not in addition call for a conception of oneself as the subject of one's experiences. If according to the individual principle, the individual's awareness of its mental states in a non-inferential way is what state consciousness depends on, then explanatorily speaking, one needs to zoom in on the awareness of a particular class of things in a particular way. However, that hardly means that a conception of oneself as the subject of one's experiences is needed just because it is one's own mental states that are involved or just because the awareness is non-inferential. It is time to explain why.

There is a vast literature on how an organism gets to be aware⁵² of its mental states, and I obviously don't have the space here to address all issues involved.⁵³ Again, I am interested merely in seeing if such awareness is possible without the awareness of I-in-the-experience. According to some theories of consciousness, one way an organism

⁵² I will no longer qualify the awareness as a non-inferential awareness. It is to be understood that way unless noted otherwise.

⁵³ For a comprehensive anthology of related issues, see Block, Flanagan and Guzeldere (1999).

gets to be aware of its mental state x is by there being another mental state y about it. Mental state y could be a thought or a perception⁵⁴ and is usually described as a higher-order state because it is about another state. Whether the higher-order state is a thought or a perception is not relevant here. The important point is that it is only in virtue of there being a mental state about another mental state that the individual is aware of the latter. It can then be suggested that such an account is relatively free of a conception of oneself as the subject of one's experience.

I say relatively free because according to the higher-order thought theory, the higher-order state is described as a thought to the effect that I am in this (target) state (Rosenthal, 2005). The content of the higher-order state, "I am in that state," perhaps does imply the existence of the subject and an awareness of oneself as the subject of state so and so. But one should keep in mind that the dispensability of subjects is not really the concern of a consciousness theorist unless he has a strong commitment to some version a neo-Humean bundle view or a Parfitian reductionism, so he will not be concerned whether the content of the higher-order state can be put in a subject free manner. But here, it does matter and it is possible that the occurrence of the higher-order state with a certain content does not imply the existence of a subject or an awareness of I-in-the-experience. For instance, there is no reason to think that the higher-order state could not be an indexical kind of thought with the content "that state" rather than "I am in that state." The purpose of the indexical type of higher-order state may be to assign the target state to a group of other states that fall on the same spatio-temporal path including the higher-order state itself. Hence, the subject-friendly content of the higher-order state

⁵⁴ See Rosenthal (2005) for the former and Lycan (1996) for the latter.

which is “I’m in so and so state” may be replaced by the following subject-free content of the higher-order state: “this target state goes with this body and this series of psychophysical states.”

But I haven’t still offered a way to answer the question of what allows the organism to be aware of its mental state x . According to higher-order theories, when x is accompanied by another mental state y , the individual is aware of x . But then we’re left with the question of why some mental states are accompanied by higher-order states while others are not. I have endorsed a non-atomistic view of experiences in chapter 1 – that each experience presupposes a relational matrix of other mental states (conscious and non-conscious) – and I now suggest that it is the interactions in a relational matrix of mental states that a mental state x is associated with which determines whether or not the organism will be aware of x . Consider for instance a belief that it will rain at time t . Consider further that this belief arises as a result of an observation of the sky at an earlier time. A belief that it will rain at t presupposes a perception of the sky at an earlier time along with many other psychophysical states. Among these are past sky-observations, some of which were followed by confirmations, i.e., observations that it did rain, and by disconfirmations, observations that it did not rain which eventually lead to the belief that it will rain at t only when the sky-observation is in a particular way. As mentioned before in chapter 1, these various psychological states constitute the relational matrix of mental states that the belief that it will rain is associated with.

I suggest that it is the interactions in a relational matrix of mental states that a mental state x is associated with which determines whether or not the organism will be

aware of *x*. It is possible that every time a creature is (non-inferentially) conscious of its mental state, there is a higher-order state. The higher-order state can be seen as a byproduct once the interactions reach a certain threshold. The higher-order state could be a by-product of the activity of the relational matrix and when I say activity, it might be possible one day to think of the activity in quantitative terms. There are after all some quantitative accounts of other aspects of our mental lives such as the binding of inter-modal experiences. So there is no reason to think that a similar account of the activity of the relational matrix that would cause the organism's attention to be directed at its mental state is impossible. The important point here is that by endorsing the transitivity principle it is possible to account for consciousness without relying on an awareness of I-in-the-experience.

One may argue that this approach is too broad, that it does not explain how an organism gets to be aware of its *mental states*, that it would cover how an organism gets to be conscious of other things as well, like rocks and chairs, rendering the explanation of state consciousness somewhat uninformative. But, there is also no reason to think that becoming aware of one's mental states should be fundamentally different from becoming aware of things in general. Without a doubt, the general explanation of becoming aware of things would need fine tuning and some way to distinguish inferential processes from non-inferential ones would be called for to make the explanation more specific, but in the end it could be based on some neurological articulation of how the organism's attention mechanisms work, and not on the organism's seeing itself as this mysterious thing called the subject.

Recall the case of the absent-minded long-distance truck driver (Armstrong, 1968). The reason why he is capable of driving without causing any accidents but in a distracted state of mind ultimately depends on the familiarity of road, how long he's been a truck driver, whether he is traveling at night or daytime, etc. It is much more difficult to establish a link between all this input and one's conception of oneself as the subject of his driving than between this input and the relational matrix of mental states that is associated with his driving which practically consists of the kinds of things that are in the input.

The differences in the phenomenology of similar experiences of different persons are also better accommodated when the relational matrix is taken into account. It is unlikely that the reason why a painter is conscious of things that I am not when we look at the same painting, is that he has a conception of himself as the subject of his experience and I don't. This is even true for the differences in the phenomenology of repeated experiences of the same person.⁵⁵ If I read books on art history, visit museums and examine paintings, my experience of looking at a painting is likely to differ from looking at it long time ago. Again it is absurd to attribute the changes in what I am conscious of today when I look at the painting to a lack of awareness of I-in-the-experience in the past when I looked at the same painting.

In chapter 2, in order to explain the unity of experiences of a single person, I appealed to the fact that a substantial number of a person's psychological states are acquired from a particular orientation to the environment, rendering the locus of our

⁵⁵ For a related discussion see how Rosenthal (2005, p.187-8) applies his higher-order thought theory to explain cases like wine-tasting.

being an indispensable part of persons' mental lives. So now I contend that what the real “intrinsic feature of primary experience” (Gallagher and Zahavi, 2008, p.46) or the real “constitutive feature and integral part of phenomenal consciousness” (p.50) is that locus and its continuity, and not pre-reflective self-consciousness.

Kriegel (2004), I think as a result of endorsing the state principle, ends up arguing that state consciousness depends on a certain form of self-consciousness. I have suggested a way of understanding the individual principle according to which a conception of oneself as the subject of one's experience may not be needed for consciousness, and appealed to it as an alternative to the state principle. However, one must take into consideration Kriegel's arguments and his distinctions between different forms of consciousness in detail to deny the dependence of consciousness on self-consciousness and inevitably this will involve a criticism of the state principle too. And that is what I come back to now, viz., the distinctions Kriegel draws between different forms of consciousness and self-consciousness. Although the different forms of self-consciousness mimic the different forms of consciousness, they are not related to each other in the same way that the different forms of consciousness are. My purpose is to show the difficulties involved with these distinctions in order to deny what might have otherwise been positive and independent reasons to think that consciousness depends on self-consciousness.⁵⁶

⁵⁶ One of Kriegel's (2004) arguments against eliminativism towards intransitive self consciousness, or peripheral self-awareness, is the fact that we accept the presence of peripheral awareness for everything else, so self-awareness should not be fundamentally different.

3.4 *Four forms of consciousness*

Kriegel (2004) uses four qualifying terms to talk about different kinds of consciousness. They are creature, state, transitive and intransitive. And he always uses a pair of these terms to specify which type of consciousness he is talking about. Hence there are transitive creature consciousness, intransitive creature consciousness, transitive state consciousness and intransitive state consciousness. In relation to the different forms of consciousness introduced earlier, viz., state and transitive consciousness, one can say that intransitive state consciousness is the same as state consciousness, which consists in the mental state's being conscious. Transitive creature consciousness is the same as transitive consciousness, which consists in the creature's being conscious of something. Intransitive creature consciousness is the same as creature consciousness which consists in the organism's being awake and mentally responsive. That leaves us with transitive state consciousness, which has not been referred to earlier.

It is relatively easy to grasp what it means for a creature, an individual, a person or an animal to be conscious of something. But what does it mean for a mental state to be conscious of something? Kriegel (2004, p.183) too notes that a sentence like "Smith's thought is conscious of Smith's new car" is ungrammatical but he adds that this only has to do with the way the phrase 'conscious of' is used and does not thereby imply that there is no such thing as transitive state consciousness. He describes it in the following way;

the occurrence of transitive creature consciousness implies the occurrence of a mental state. When Smith is conscious of her new car, she is in a certain mental state, and it is in virtue of being in this mental state that she is conscious of the car

(as opposed to being conscious of a table, or not being conscious of anything). We may label the property the mental state in question has transitive state consciousness. A mental state exhibits transitive state consciousness iff it is a mental state in virtue of which its subject exemplifies transitive creature consciousness. Thus, Smith's thought of her new car is transitively state-conscious, because in virtue of having it, Smith is conscious of her new car. (p.183-4)

Given Kriegel's emphasis on how Smith would be conscious of a *car* as opposed to something else, or as opposed to nothing, I believe that the transitive nature of state consciousness consists really in the mental state's being *about* something. It is because the mental state is about the car that Smith is conscious of a car (and not of anything else) in virtue of being in that mental state. Kriegel (2004) defines transitive state consciousness as the *property* a mental state M has when M's subject is conscious of something in virtue of being in M (p.185). I believe that by *property* Kriegel here means something other than the property of consciousness. Instead it seems to be the property of being about something. Being about something is the property a mental state has when the possessor of the state is conscious of something in virtue of being in that mental state.

Kriegel (2004) claims that transitive state consciousness in turn depends on intransitive state consciousness. In other words, it is only if the mental state is conscious, can the individual be conscious of something in virtue of that state, depending of course on what the mental state is about. Kriegel then gives the following reasons as to why transitive state consciousness depends on intransitive state consciousness. And it should

be obvious that these reasons can be interpreted as the reasons for the state principle, because as stated above, the dependence of transitive state consciousness on intransitive consciousness consists in the individual's becoming conscious of something in virtue of being in a conscious mental state (about that thing). I will now consider whether these reasons can be used against the individual principle.

First Kriegel (2004, p.185) uses the example of a tacit belief that 17.3 is greater than 13.9 and says that we would not want to say that someone is conscious of the fact that 17.3 is greater than 13.9 just because he has a tacit belief to that effect. Hence first the mental state involving this mathematical proposition should be conscious so that one can talk about the individual being aware of such a proposition. However, this example is a bit biased. It is much more likely that a person's mental states regarding *mathematical propositions* are conscious when he is doing calculations and is aware of how the numbers relate to one another. But one cannot thereby infer that this is how it'll be for all mental states across the board. In other words one cannot infer that for everything that an individual is aware of, there will be an accompanying relevant conscious mental state. Consider the example of the long-distance truck driver again. As he is driving on the highway, not causing any fatal accidents, he must be conscious of some things, but at the same time it is obvious that he is not in as many conscious mental states as his awareness of the road would call for if the state principle were true.

The second reason Kriegel (2004) gives for why transitive creature consciousness depends on intransitive state consciousness is that "as far as the normal usage of the terms is concerned, it is all but incoherent to describe a person as unconsciously

conscious of the fact that p” (p.185). In other words, since a person can be conscious of something only in virtue of being in a conscious mental state, one cannot both be in a non-conscious mental state about x and be conscious of x, i.e., one cannot be unconsciously conscious of something. However, the incoherency in the phrase cannot be used as a reason why one cannot be at the same time conscious of x and not be in a conscious mental state about x, especially when the word ‘conscious’ is clearly being used in two different ways; the former qualifying term ‘unconsciously’ represents state consciousness and the latter ‘conscious’ or rather ‘conscious of’ represents a person’s being conscious of something, so transitive consciousness. This is an *ad-hoc* argument for the state principle since being unconsciously conscious of something amounts to a person’s being aware of something but not being in the relevant conscious mental state. And this is not immediately incoherent unless one endorses the state principle.

The third reason Kriegel (2004) gives for the state principle is that persons are conscious of only a few things at a time and that if persons were conscious of things in virtue of being in non-conscious mental states, they’d be conscious of innumerable many things (p.186). I contend that this argument reveals a misunderstanding of the individual principle as an alternative to the state principle. The individual principle says that those mental states that an individual gets to be aware of are classified as conscious mental states. An individual’s being conscious of things (other than mental states) does not imply the presence of conscious mental states. But it does not thereby claim that an individual gets to be aware of things in virtue of being in non-conscious mental states. It just endorses the possibility of transitive consciousness without state consciousness. That

does not mean every non-conscious mental state one is in will make him conscious of something.

After discussing Kriegel's (2004) reasons for the state principle, I'd like to now turn to his arguments for why consciousness depends on self-consciousness, or rather why intransitive state consciousness in turn depends on intransitive state self-consciousness, rendering all forms of consciousness to be dependent on intransitive state self-consciousness. To do that, I will first consider Kriegel's four forms of self-consciousness.

3.5 *Four forms of self-consciousness*

Kriegel (2004) uses the same terms as earlier, and four pairs of them, to qualify self-consciousness. They are intransitive creature self-consciousness, transitive creature self-consciousness, transitive state self-consciousness, and intransitive state self-consciousness. Intransitive creature self-consciousness is exemplified by a statement like 'Smith is self-conscious'. This is supposed to be different from regular creature consciousness in the sense that that Smith is not just awake and mentally responsive but also in a kind of deep and introspective state of mind – whatever that may be. It is difficult to conceive how a person can be in this inward-looking state of mind without looking inwards to some mental state or other, so it is not clear where its intransitive nature comes from. Any kind of deep, introspective, contemplative state of mind seems to require contemplative thinking, hence thinking about something. So unless, intransitive creature self-consciousness is meant be some kind of a character trait, and Kriegel does mention this possibility, it is hard for me to understand the intransitive nature of what he

calls intransitive creature self-consciousness. However, it is not going to be very relevant to the remainder of the discussion so I'm not going to dwell on this.

Next is transitive creature self-consciousness, exemplified by a statement like 'Smith is self-conscious of her new car.' So consider Kriegel's (2004) following examples;

4) Smith is conscious of her new car. (p.183)

5) Smith is self-conscious of her new car. (p.187)

(4) exemplifies transitive creature consciousness and (5) is supposed to exemplify transitive creature self-consciousness. Since (5) is ungrammatical, Kriegel suggests the following in place of (5) (p.187);

6) Smith is self-conscious of thinking that her car is new.

or

7) Smith is conscious of her thinking that her car is new or Smith is consciously thinking that her car is new.

In explaining why (5) is ungrammatical, Kriegel says that self-consciousness cannot take a car or, for that matter, any other external object as its object (p.187). So (5) is modified in a way that allows the transitive creature self-consciousness to be directed at something internal and not external. Also, it is directed not just at the thought as the internal state, but at the having of the thought, in other words, at oneself *and* the thought, since only then is there awareness of oneself as well which is supposed to be the appropriate object of transitive creature *self*-consciousness. So it is possible to say that transitive creature self-consciousness is not really a different form of consciousness in the way transitive

consciousness is different from state consciousness. Instead, as Kriegel notes too, it is more like a special form of transitive creature consciousness. It “is the property a creature has when she is conscious of herself and her mental states” (Kriegel, 2004, p.187).

What makes transitive creature self-consciousness a special form of transitive creature consciousness is not only that the object of the former cannot be an external thing, but that it cannot just be a mental state either. It has to be a mental state and oneself, or put together, oneself in the mental state. Kriegel (2004) says that this state of self-consciousness is distinct from the thought about the car (p.188), so it's not just an integral part of the thought about the car, which means it is not the kind of implicit awareness of I-in-the-experience mentioned earlier. This one is the reflective, thematic, explicit self-awareness. Prima facie, transitive creature self-consciousness or the explicit type of self-awareness, seems to assume the subject thesis, just like pre-reflective self-consciousness did. The idea is only if the subject thesis is true, is it possible for there to be awareness of oneself (as the subject of one's experience) when there is awareness of an experience. But technically speaking, as mentioned before, how things are does not necessarily translate into how individuals are aware of them. The subject thesis does not have to be true for me to be aware of myself as the subject of my experiences. Put in another way, the possibility that an awareness of a thought is most naturally described as being accompanied by an awareness of the thinker does not imply that thoughts are the kinds of things that are had by thinkers, but again would be supported by it.⁵⁷

Still, how is the awareness of oneself as the thinker of a thought any different from the awareness of the thought? When there is awareness of oneself as a subject

⁵⁷ I think this is true regardless of one's position in relation to the state or the individual principle.

having pain, in addition to the awareness of the pain, does it hurt more? Does it hurt in any different way? One may argue that it does not make sense to ask this question, because such awareness does not involve a sensation like pain and hence lacks phenomenal properties, having no effect on how much it hurts. While this may be a fair objection, it brings with it the question of what possible effect it may have whatsoever. And there seems to be no answer to that either. One might say that it is because I am conscious of my having pain that I am capable of seeking relief. But not necessarily so. When there is pain, it comes with the whole package; the suffering, the unpleasant feeling, the desire to avert it etc. And if you start taking away from the package, then you no longer have pain.

I contend that transitive creature self-consciousness is an abstraction from rather complicated mental states such as regret, or Frankfurtian type second-order desires, or one's concern for one's future where it seems that we must be aware of ourselves as the subjects of our experiences to have regret or a more noble desire or a fundamental concern for one's future. But all that is required for such mental states is the kind of continuity that binds the experiences of a single person together, as discussed in detail in chapter 2. And memories alone provide good means for the awareness of the continuity involved in a single person's mental life without relying on a separate conception of oneself as this mysterious thing that is the subject. I will come back to this explicit sense of self-consciousness incorporated into certain psychological states at the end of this chapter and in chapter 4.

Next on the list is intransitive state self-consciousness, which mimics intransitive

state consciousness. Consider the following where (8) is an example of intransitive state consciousness and (9) of intransitive state self-consciousness (Kriegel, 2004).

8) Smith's thought that her car is new is conscious. (p.183)

9) Smith's thought that her car is new is self-conscious. (p.187)

According to Kriegel (p.187-8), (9) represents a certain *way* the experience is had by the subject. It is the same as saying,

10) Smith is *self-consciously* thinking that her car is new.

Notice the contrast between (10) and (6), viz., 'Smith is *consciously* thinking that her car is new'. This is the contrast between transitive creature self-consciousness and intransitive state self-consciousness, both having to do with the *way* an individual is in a mental state M. According to Kriegel (p.189), there are two ways a person can be aware of being in mental state M; either peripherally or focally, the former referring to intransitive state self-consciousness and the latter transitive creature self-consciousness. As mentioned earlier, in sentences like (6), the thought about the car and the state of self-consciousness are two distinct states (p.188). In (9) and (10) however, the thought about the car and the state of self-consciousness are not two distinct states (p.188). Kriegel notes that when one is aware of a mental state in a self-conscious kind of way, the attention is directed at whatever the external object of the mental state may be.

Kriegel's (2004) description of the self-conscious kind of way of having an experience, i.e., the way in (9) or (10), is very similar to the description of pre-reflective self-consciousness in the phenomenological tradition. He first cites Goldman (1970, p.96) who says;

In the process of thinking about x there is already an implicit awareness that one is thinking about x. There is no need for reflection here, for taking a step back from thinking about x in order to examine it . . . When we are thinking about x, the mind is focused on consciousness, not on our thinking of x. Nevertheless, the process of thinking about x carries with it a non-reflective self-awareness.

Referring to Goldman's distinction between reflective and non-reflective self-awareness, Kriegel (2004) adds;

In what Goldman calls 'non-reflective self-awareness', the primary object of the subject's mental state is not herself or her mental state, but whatever the external object of these mental states may be. When one has a mental state self-consciously, one is aware primarily of the (external) object of the mental state in question. But there is also a more subtle, peripheral awareness of oneself implicit in that state. One is aware of oneself precisely as that state's *owner*, or *subject*. To say that Smith is thinking that her car is new in a self-conscious sort of way, then, is to say that Smith is implicitly, or peripherally, aware of her *having* the thought, or of the thought being *her own*. (p.189)

Consider the case of the long-distance truck driver again and imagine that at some point he witnesses an accident. Surely there is a difference between the truck driver's awareness of the accident and his awareness of the road before the accident. The problem is, given Kriegel's (2004) infrastructure of consciousness, it is difficult to accommodate this difference. Earlier, the example of the long-distance truck driver was used to support the view that an individual can be conscious of things without being in conscious mental

states. Since Kriegel denies that one can be “unconsciously conscious” of things (p.185), I'd expect that he takes the truck driver's mental states such as the perception of the road etc. to be conscious. Therefore, when the awareness of the accident and the awareness of the road before the accident are compared, there should be no significant difference because if one gets to be aware of things in virtue of being in conscious mental states, then both the driver's perception of the road before the accident and his perception of the accident need to be conscious on Kriegel's account, unless the driver is taken to be unaware of the road. But the driver must be aware of the road or he would be in an accident rather than witness it.

It is interesting that Kriegel (2004) uses the example of the long-distance truck driver too to support the view that there is such a thing as the peripheral awareness of things (p.191). But I believe Kriegel's appeal to the long-distance driver in relation to peripheral awareness is more confusing than helpful because his main purpose in using the example is to articulate what it means to be peripherally aware of things in general as opposed to being focally aware, so it amounts to the same type of consideration as those involved in peripheral vision. But in his theory of consciousness the qualifying terms “peripheral” and “focal” are used in relation to transitive creature self-consciousness and intransitive state self-consciousness, so they're really about one's awareness of oneself-in-the-experience. Given the frequent usage of the example of the truck driver in theories of consciousness, one can easily be misled into thinking that the truck driver is focally aware of the accident, meaning it involves transitive creature self-consciousness, and he is only peripherally aware of the road (before the accident) meaning it involves

intransitive state self-consciousness. But neither is the case.

The only difference between the awareness of the road before the accident and awareness of the accident on Kriegel's account would have to be about the way the driver is aware of being in these mental states. I guess using Kriegel's terms, it would be safe to say that when the driver is aware of the accident, he is *more* self-consciously aware of his conscious perception of the accident and when the driver is aware of the road before the accident, he is *less* self-consciously aware of his conscious perception of the road. This would be supported by Kriegel's (2004) comment that peripheral awareness and focal awareness need not be two exclusive states and that there is "probably is a wide spectrum of awareness, ranging from the very focal to the very peripheral" (p. 202, fn 12). While this may seem to be supportive of how awareness of the road compares to awareness of the accident, it contradicts Kriegel's treatment of transitive creature self-consciousness and intransitive state self-consciousness as two separate forms of consciousness, because the 'very focal' awareness would consist in transitive creature self-consciousness, meaning that there'd be two states involved, one the mental state – whatever it is about – and two the state of self-consciousness. But in peripheral case, there is only one state, which is the mental state – whatever it is about.

One may defend Kriegel by saying that the wide spectrum of awareness ranging from the very focal to the very peripheral is not about one's awareness of oneself in the experience. As mentioned before, it's about the awareness of things in general like peripheral vision vs focal vision. But one of Kriegel's (2004) arguments for peripheral I-in-the-experience awareness is that it would be absurd to deny peripheral vs. focal self-

awareness when the distinction between the peripheral and the focal awareness is accepted in all other fields (p.195). That means the wide spectrum should be applicable to I-in-the-experience awareness as well, rendering the criticism above tenable.

Of course there is one reason to deny the distinction between peripheral and focal self-awareness and accept the distinction between peripheral and focal awareness in all other fields and that is because there may be no such thing as self-awareness to which the distinction can be applied. But the objection that Kriegel considers is that there may be no such thing as peripheral awareness of things in general. He is concerned about a methodological difficulty here and says even that if we want to convince the denier of peripheral vision “whenever we want to direct our attention to the periphery of our visual field, the latter thereby becomes the focal center” (Kriegel, 2004, p.193) I believe this methodological difficulty should pose no problem when it comes to peripheral self-awareness, because as Hume (1739-40/1978, p.252) had famously noted too, no matter how intensely one tries to direct one's attention to himself, nothing seems to change. That also suggests that peripheral self-awareness and peripheral awareness of things are not as similar as Kriegel might think.

All in all, there are inconsistencies in Kriegel's (2004) account and his arguments for the dependence of consciousness on self-consciousness. In fact, I contend that Kriegel's arguments for the dependence of consciousness on a form of self-consciousness, viz., intransitive state self-consciousness, undermines his endorsement of the state principle in favor over the individual principle. To see why, consider once again the description of intransitive state self-consciousness. It is about the way the *individual*

is aware of being in some mental state. So basically it is about the individual's awareness of a mental state. Initially he says that an individual is aware of things only in virtue of being in conscious mental states and then he claims that no state consciousness is possible unless the individual is (peripherally) aware of having the mental state.

According to Kriegel (2004), the dependence of consciousness on self-consciousness comes down to the following argument; “conscious states are first-person knowable; first-person knowable mental states must be intransitively self-conscious; therefore conscious states are intransitively self-conscious” (p.198). I have already addressed the difficulties with this argument in 3.3. To repeat, the problem is with the premise that first-person knowable states, in other words, those states that one gets to be non-inferentially aware of, must involve an awareness of I-in-the-experience.

Undoubtedly the premise relies heavily on the subject thesis, and there are good reasons to reject the subject thesis as I discussed in chapter 1. However, I have also suggested independent reasons why non-inferential awareness of an experience does not have to imply self-awareness.

Kriegel (2004) argues that our awareness of our conscious states is an awareness of ourselves as the subjects of those mental states because it is impossible to think about a particular mental state without thinking about the subject whose mental state it is (p.199). I contend however that it is impossible to think about a particular mental state in isolation because I deny an atomistic view of experiences and not because I cannot think about a mental state without a subject. More importantly I'm suggesting that there is nothing to be gained by thinking about an experience as something that is had by this mysterious thing

that is the subject, when there is already at hand a spatio-temporally continuous sum of psychophysical states and a body.

Kriegel (2004) does consider the objection that it *may* be possible to think about a particular mental state without thinking about the subject whose mental state it is, and he agrees that it may be possible, but then he adds that it would be abnormal for at least a human thinker to do so (p.199). Supposedly then it is in fact possible to think in a subject-free manner but this possibility does not apply to human thinkers. I cannot conceive any reason other than an extreme loyalty to our linguistic habits and a conceptual scheme arising from those habits why humans can't think in this way especially when the possibility of subjectless thinking is granted, and I take this reason to be philosophically insignificant.

Nevertheless, there may be other reasons for thinking that subjects are more than just grammatically generated illusions while denying that they are separately existing entities. One prime example of this consists in the Kantian view of subjects according to which the ascription of experiences to a numerically identical subject is necessary to make sense of our experience of the objective world.⁵⁸ That is to say to be able to conceive of diverse perceptions as perceptions of an enduring object. Hence if the experiences of humans were to be fundamentally different, for instance if they could not have the concept of persisting objects because they did not think of themselves as the persisting subjects of their experiences, then obviously not thinking of oneself as a subject would have substantial significance rendering the explanatory role of subjects indispensable.

⁵⁸ For a discussion of this and other related topics see Rosenberg (2005).

Cassam's (1989) neo-Kantian argument explains why a numerically identical subject to which experiences can be ascribed and who is aware of those experiences as *his* experiences is needed to have an experience of persisting objects. Consequently Cassam challenges Parfitian reductionism and the impersonal description claim. In what follows I will see if the neo-Kantian argument can be used to refute reductionism and the impersonal description claim.

3.6 The neo-Kantian argument and reductionism

One difficulty with Parfit's reductionism has to do with the following two claims that he endorses simultaneously. On one hand, Parfit (1984, p.211) says that;

- 11) "A person's existence just consists in the existence of a brain and body and the occurrence of a series of interrelated physical and mental events."

In relation to this claim, he suggests that reality can be re-described without claiming that persons exist (p.251). It should be noted here that Parfit sometimes uses 'persons' and 'subjects of experiences' interchangeably for reasons that will become clearer as I proceed. Throughout this essay however, they are treated as separate things. As mentioned before, while I argue that there are no good reasons to think that subjects of experiences exist, I contend that there are good reasons to think that there are such things as persons just as there are good reasons to think there are mountains, chairs and cats.

Parfit (1984) also says that;

- 12) A person is an entity that is *distinct* from a brain and body and a series of interrelated physical and mental events, but it is not a separately existing entity (p.211, his italics)

Here, Parfit thinks of persons as thinkers, believers, etc. – as opposed to just thoughts, and beliefs, etc. – in other words, he thinks of persons as subjects of experiences, hence uses the terms ‘persons’ and ‘subjects’ interchangeably. He says “A Reductionist can admit that . . . a person is *what has* experiences, or the *subject of experiences*” (Parfit, 1984, p.223). But then he goes on to say that “this is true because of the way in which we talk” (p.223) and emphasizes that what a reductionist denies are subjects as separately existing entities.

One important question is whether or not (11) and (12) can be reconciled. Cassam (1989) thinks not. Neither do I, but for completely different reasons. I take the reduction of persons to bodies and experiences to imply an eliminative approach to subjects and my reasons for not being eliminative with respect to persons have nothing to do with their being thinkers, believers, etc., in other words, subjects of experiences.

According to Cassam (1989), although Parfit does not want to be considered as an error theorist and he does not want to be eliminative with respect to subjects, his endorsement of (11) and consequently the impersonal description claim makes it very difficult to support this view. It seems that Parfit can only appeal to the way in which we talk to support the existence of subjects – again not as separately existing entities – and not be eliminative with respect to them. However, Cassam argues that appealing to surface grammar is not going to cut it, since a strict denial of subjects as separately existing entities, when coupled with the impersonal description claim, renders subjects explanatorily redundant, leaving no sufficient grounds to think of subjects as distinct entities hence no reason why one should not be eliminative about subjects.

Consequently, if one wishes to endorse the existence of subjects, while denying they are separately existing entities, one needs a better reason than linguistic habits. And Cassam (1989) believes that the Kantian view that a numerically identical subject is required to think of diverse perceptions as the perceptions of one and the same enduring object, provides a good way to argue for subjects that are more than linguistically generated illusions and at the same time not separately existing entities since Kant does not consider subjects as separately existing entities. He says “The identity of the consciousness of myself at different times is therefore only a formal condition of my thoughts and their coherence, and in no way proves the numerical identity of my subject” (A363)⁵⁹. But as Cassam notes too, when the Kantian view is endorsed, it is difficult to argue for (11) and the impersonal description claim.

Parfit (1984) initially takes the Kantian view to be incompatible with reductionism because it renders the impersonal description claim false. But later Parfit (1998) revokes the impersonal description claim. However the reason is not that he is incapable of explaining the unity of consciousness without referring to subjects – another phenomenon which Kant explains by referring to the numerical identity of subjects. Parfit (1984) appeals to co-consciousness (p.250-1), for unity of consciousness at a time and R-relatedness (p.215), for unity of consciousness at different times which is a relation of psychological continuity to explain the unity of consciousness, and this is not something that he gives up. In chapter 2, I argued that psychological continuity is not enough to explain the unity that holds between the experiences of a single person, without

⁵⁹ References are to Kant (1781-7/1929), viz. to the *Critique of Pure Reason*, translated by Norman Kemp Smith.

presupposing the identity of the person and I suggested an alternative account, one that is based on spatio-temporal continuity that did not appeal to subjects.

The reason why Parfit (1998) gives up on the idea of the impersonal description claim is Cassam's (1992) objection that in order to describe the content of a self-ascriptive thought, one needs to appeal to the thing it is a thought of, viz. the subject. So as it turns out, one cannot re-describe reality without claiming persons exist if one wants to include those self-ascriptive thoughts in the description.

Cassam (1989) argues that reductionism takes self-ascriptive thoughts for granted and treats them as a brute fact, and does not explain the purpose of such thinking. He considers this as a weakness of reductionism and believes that the Kantian argument provides a way to improve reductionism, without contradicting it since the Kantian notion does not take subjects of experiences to be separately existing entities either (A107). According to Kant, there is an intimate connection between the idea of experience of independent objects and unity of consciousness and Cassam (1989) believes that according to Kant, "for intuitions to be conceptualizable as representations of an objective world, they must be self-ascribable by a subject capable of consciousness of his own identity as the single subject of the representations" (p.86-7). However it is not clear how this thesis can be defended for it is not clear why in considering that diverse perceptions are perceptions of the same thing, one must also take these diverse perceptions to be his.

According to Cassam (1989), the explanation of the Kantian view, i.e. discussing the reasons why a numerically identical subject is needed for there to be an experience of

the objective world, may provide a meaningful way to reconcile (11) and (12) according to which subjects would constitute a formal condition without being ontologically distinct. In this way, subjects are no longer explanatorily redundant so they are significant in some sense; i.e., more significant than a linguistic by-product, and at the same time not a separately existing entity. Cassam (1989) says,

Successive perceptions are conceptualizable as perceptions of the very same thing precisely because of the possibility of ascribing those perceptions to a numerically identical subject whose route through the world anchors those perceptions to one particular object . . . It is the numerical identity of the subject of diverse experiences which provides the anchoring. And for the numerical identity of the subject to play this part, those experiences must be thought as having a subject. (p.92)

In order to defend and explain Kant's view, Cassam (1989) talks about the distinction between having successive perceptions of one and the same thing, and successively perceiving qualitatively similar but numerically distinct objects. According to his discussion, one way of conceptualizing diverse perceptions as the perceptions of an enduring object consists in the spatio-temporal interactions between what is perceived and the perceiver and relevant intertemporal restrictions⁶⁰ consistent with the physical laws of this world. Cassam argues that in order to make use of this spatio-temporal interaction in conceptualizing diverse perceptions as the conceptions of an enduring object, the numerical identity of the perceiver must first be established.

I do agree that the numerical identity of *something* is a necessary condition to

⁶⁰ Cassam (1989, p.95) borrows the term 'intertemporal restrictions' from Peacocke. See fn 32.

conceive objects in the world as enduring entities. Since perceptions arise in the way they do as a result of unique spatio-temporal interactions between the perceiver and the perceived, given a perception, and the numerical identity of the perceiver, it should be possible to arrive at the identity of the perceived object. Likewise, given the perception and the perceived object, it should be possible to individuate the perceiver too. And by individuating the perceiver, I am talking about capturing the particular orientation of the perceiver with respect to the environment, which must coincide with the perceiver's location in some unambiguous way, giving rise to a perception in a non-arbitrary manner. That is to say, if I am standing at the southwest corner of 5th avenue and 34th street and looking up at the Empire State building, which side of the building is perceived or its slope from where it is being perceived are not going to be arbitrary.

Cassam's (1989) argument does not contradict this. However, he contends that I would not be thinking of the building as one and the same building that I passed by several times, unless I also considered the several perceptions of the building as *my* perceptions and considered myself to be the singular persisting subject of those perceptions. The question is why this additional requirement is called for to think of the building as an enduring object in the world. Imagine *s* is the perceiver, *p* is the perception, and *o* is the object perceived. Given *p* and *s*, it should be possible to establish the identity of *o*. Likewise, given *p* and *o*, it should be possible to establish the identity of *s*. In the latter case though, it is obvious that *o* cannot think of itself in a certain way. Still it helps to locate *s*. Then why should *s* think of itself in a particular manner to establish the identity of *o*?

One obvious answer to this question is that *o* is used merely as a tool to scientifically locate *s*, as an enduring, numerically identical thing. Likewise, given *p*, *s* can be used as a tool to locate *o* (as an enduring, numerically identical thing) regardless of the thinking abilities that go with *s*. Note that *s*'s being used as a tool to determine the numerical identity of *o* here consists in its physical nature and its tracing a physical route in space. In other words, *s* is the body of a person acting as an experience-generator. This idea of the body of a person as an experience generator was also present in chapter 2, according to which the occurrence of all psychological events were taken to be physical in the sense that they all occurred in some body. It is in the body of a person that experiences are generated in virtue of the body's interaction with the environment and in accordance with some suitable theory of consciousness. What matters for the numerical identity of objects then is the location of this experience generator, with respect to the things there is an experience of.

So on one hand there is the *scientific* account of establishing an object such as the Empire State building as an enduring and numerically identical thing from multiple perceptions and a perceiver, which does not require the perceiver to see herself as the perceiver of these several perceptions. This is true especially when one accepts that, also, the numerical identity of the perceiver can be established on the basis of the object perceived which clearly does not think, let alone think of itself as the thing perceived. Hence in the scientific account, if the intertemporal restrictions and the object are given, it is possible to establish the numerical identity of the subject from the perception. Likewise if the intertemporal restrictions and the subject are given, it is possible to

establish the numerical identity of the object from the perception. For instance if you feed into a computer program various snapshots of the Empire State building, the times that the snapshots were taken, plus the spatio-temporal path of the camera that was used to take those shots within that time interval, and the physical laws of our universe, I see no reason why the computer would not be able to predict the numerical singularity of the building.

As opposed to the scientific account, there is what I call the experiential account, or in Cassam's (1989, p.103) terms an account "from a standpoint within a given mental life" according to which the requirements of the scientific account are not sufficient for a subject to *conceive* of the diverse perceptions of the Empire State building as the perceptions of one and the same building. One needs to think of himself as the persisting, singular subject of these diverse perceptions in order to think of the building as one and the same enduring building.

To bring out the difference between the scientific and the experiential account, it is worth considering what Cassam (1989) says in his account, which I take to be the experiential account: "if the perceiver has the conception of himself as appropriately located with respect to O_1 , and not to any other O-type object, then he has the resources to conceptualize his successive perceptions *as* successive perceptions of the single object O_1 " (p.89). I contend however that in relation to the scientific account, it would be enough that the perceiver *is* appropriately located with respect to O_1 . Additional thoughts such as the conception of oneself as the subject who is appropriately located would not matter.

Supposedly then there is something about the experiential account that is not captured by the scientific account. But here is the tricky part. Even if I were to think of myself as the singular enduring subject of my experiences, I would not necessarily be convinced that my diverse perceptions of the Empire State building are in fact perceptions of one and the same building. For instance I may believe that a dozen of these buildings have been built and every night they are being switched one with another. So obviously my having a conception of myself as the subject of my experiences is not a sufficient condition. But more importantly, it seems that the best way to disconfirm various hypotheses that would prevent me from thinking that the building is one and the same, is by appealing to a scientific account.

Still there is the possibility that awareness of I-in-the-experience is a *necessary* condition of having the concept of enduring objects. It is true that the numerical singularity of the subject can be used as a tool to scientifically establish the singularity and the enduring property of an object. I contend that the scientific account may also help a skeptic with his thinking about diverse perceptions as perceptions of one and the same enduring object. Clearly the scientific account is not in need of awareness of I-in-the-experience. But more importantly, awareness of I-in-the-experience is not a significant part of the experiential account either because whatever reasons there are for being skeptical about the numerical identity of an object, they seem to be addressed by a scientific account and not by awareness of I-in-the-experience.

One may argue here that being skeptical about the numerical identity of an enduring house already presupposes the idea of the house as an enduring entity and the

diverse perceptions of the house as the perceptions of one and the same house. Hence one might wonder where the concept of persisting objects is coming from for the scientific account to make use of when one is skeptical about interpreting diverse perceptions as the perceptions of a persisting object. If indeed the answer to this question is awareness of I-in-the-experience, in other words, if it is indeed the necessary condition, then the defender of the Kantian view needs to show that the “I” of which there is supposedly an awareness, is itself persisting, or that one is aware of it as persisting. However, unless one appeals to the body of the perceiver, there is no reason to think that the “I” there is an awareness of in the experience, is one and the same for various experiences. I have appealed to this criticism again and again under the heading of ‘multiple subjects’. Here too, the possibility of multiple subjects; i.e., the possibility of there being one subject for each experience – whichever way one individuates experiences, and individuating experiences by their subjects in this context would be circular – takes the defender of the Kantian view back to square one.

There is another way the problem of multiple subjects arises which the Kantian view faces. Rosenberg (2005, p.124) says that Kant is perfectly aware that in having the following thoughts I think [X] and I think [Y] and I think [Z], one is not entitled to thinking that the I who thinks X, and the I who thinks Y, and the I who thinks Z are one and the same I.⁶¹ It is only in virtue of having the thought I think [X+Y+Z] that the identity of the subject is possible. Kant says “Only in so far . . . as I can unite a manifold of given representations in *one consciousness*, is it possible for me to represent myself

⁶¹ Kant says “For the empirical consciousness, which accompanies different representations, is in itself diverse and without relation to the identity of the subject.” (B133)

the *identity of the consciousness in [i.e. throughout] these representations*” (B133).

However, it is possible that while there is the thought $[X+Y+Z]$, there may yet be another thought $[A+B+C]$. And unless one is capable of having a thought that involves the conjunction of all the mental states during the course of one’s life, the diversity of the subjects remains a possibility. Again the defender of the Kantian view is back to square one.

It is possible that a scientific account of the kind mentioned above, i.e., the spatio-temporal continuity of the body as an experience generator might help in replacing multiple subjects with one, but then the possibility of the concept of enduring, independent objects would essentially depend on the scientific account and not on awareness of an enduring I-in-the-experience.

Once the idea of a person as a body and its “personal route”⁶² in the physical world is granted, the desire to still cling to subjects is puzzling to me especially when the idea of a full-blown subject as a separately existing, ontologically distinct entity is already denied. When we think of the person as this experience generator, and we locate him by making use of his body, it seems like we have everything we need to have knowledge of the objective world, that is to say knowledge of the world as including independently existing enduring objects. And why we need the relation of ascription of experiences to subjects and the subject’s awareness of itself as a subject on top of the singular personal route is not clear to me. I contend that this middle position between denying subjects of experiences as separately existing entities and assigning to them an

⁶² I borrow the term ‘personal route’ from Cassam (1989, p.92) who appeals to P.F. Strawson’s (1966, p.164) term ‘experiential route’.

importance that surpasses the linguistic conventions is neither tenable nor useful.

I believe that the apparent use of the ascription of experiences to a single subject on top of the actual singularity of the experience generator feeds again off linguistic habits. Subject-predicate propositions in our grammar may cause one to speak of himself as the thinker of a thought rather than a thought. But insofar as these are merely issues at the surface grammar level, with no metaphysical implications, the idea of a subject as a formal condition of certain types of experiences, becomes the idea of a subject as a formal condition of the language we use to talk about these experiences. This way of thinking about positing an x as a formal condition makes more sense because it is not clear how one can think of x as a formal condition of some phenomenon y and not be committed to x 's existence unless x is a *linguistic* formal condition or unless one is not concerned with Occam's razor, and is happy with a jungle ontology.

3.7 Parfit's selfless community

Behrendt (2003), tracing the debate between Cassam's (1989) neo-Kantian argument and Parfit's (1984, 1998) reductionist argument, argues that Parfit's (1998) rejection of the impersonal description claim is not as substantial as it may at first seem. I contend that the discussion above also supports this thesis. One of the reasons why Behrendt takes Parfit's rejection to be insignificant has to do with Parfit's suggestion (1998) that there could be imaginary intelligent beings who have no concept of themselves as the subjects of their experiences.⁶³ This suggestion is similar to Kriegel's

⁶³ This suggestion was intended as a reply to another objection that comes from McDowell (1997), who

view that it *may* be possible to think about a particular mental state without thinking about the subject whose mental state it is, but that it would be abnormal for at least a human thinker to do so. Parfit claims that these beings are no worse than us, humans, both scientifically and metaphysically, that they think like us, have experiences like us – only without a conception of themselves as the subjects of their experiences.⁶⁴

According to Parfit (1998), these beings – I’ll call them the members of the selfless community – have the concepts of persisting objects, of a sequence of thoughts, a sequence of experiences and acts, and they are capable of thinking that a particular sequence of thoughts occurs in a persisting body (p.228). Instead of thinking about what is involved in “seeing something” or “feeling something”, they think of what is involved in “something being seen”, or a “seeing of something”, and instead of giving names to people, the members of the selfless community give names to particular sequences (p.228-29). Hence where we say “Tenzing climbed Everest”, they say “in Tenzing there was a climbing of Everest” (p.229).

The members of the selfless community are also capable of making decisions, and are aware of their decisions, but they don’t think of themselves as making those decisions. Parfit (1998) says that the members of the selfless community conceive events like decision-making processes and the resulting acts as “another kind of happening, distinctive only in the way in which these events are the product of practical reasoning,

argues that reductionist accounts feed on our understanding of the persisting existence of persons and thereby are circular. For a detailed discussion of Parfit’s reply, see Parfit (1998, p.227-238).

⁶⁴ See Parfit (1998) p.221. This is one of the reasons why Behrendt believes that the debate between Cassam’s neo-Kantian argument and Parfit’s Reductionist argument is not resolved. Cassam (1989) argues that in order to have the concept of objects as things that endure independently of our minds, we need to first have a concept of ourselves as persisting subjects of experiences. Parfit on the other hand argues that these imagined being who have no concept of themselves as the subjects of their experiences are nevertheless capable of having concepts of persisting objects.

or, in simpler cases, of beliefs and desires” (p.229). As a result, a mountaineer from the selfless community has thoughts like; “Should *this* include a crossing of that ridge of ice? Yes it should. And, unless that crossing starts now, it will be too late. So let the ascent begin!” (Parfit, 1998, p.230)

According to the Kantian view, one major criticism of the possibility of a selfless community would have been that the members of the selfless community are not similar enough to persons, that they don’t have experiences like persons, e.g. they don’t have the concept of persisting objects. However in the light of the discussions above, I contend that an awareness of I-in-the-experience is not a necessary condition of having the concept of persisting objects. Still one may wonder whether Parfit's imaginary beings are similar enough to humans or whether such a community can be coherently imagined. There are various self-ascriptive thoughts in the absence of which the experiential life of the members of the selfless community might be significantly impoverished and not be similar enough to the experiential life that humans have, even if they did have the concept of persisting objects. The question then is whether or not the reductionist can make sense of psychological states of the self-ascriptive kind that enrich human lives, like having hopes for oneself, regretting a past deed, wanting to be different and better, etc.

It would not be efficient to try to give a separate reductionist account for each and every type of psychological state that is of a self-ascriptive kind. Reductionist accounts for states of regret, for states of hope, or for Frankfurtian types of second-order volitions are likely to have some common base. What is methodologically important then is to determine a suitable psychological feature that is the common factor of such self-

ascriptive states, and provide a reductionist account of that. I contend that this psychological feature consists in one's concern for oneself or one's future. Consequently, if a reductionist account of one's concern for oneself can be given, then a reductionist account for these self-ascriptive types of psychological states will also be tenable and I argue for the reductionist account of one's concern for one's future in chapter 4, and 4.4 in particular.

A few words on why I take the aforementioned psychological states to be forms of concern for oneself, and why I take these three cases, viz., regret, hope, and second-order volitions of the Frankfurtian kind to be providing a good sample of self-ascriptive states, are called for. Frankfurt's second-order volitions constitute a characteristic that is supposedly uniquely human. It seems that among other sentient beings, humans are the only ones that are "capable of wanting to be different, in their preferences and purposes, from what they are" (Frankfurt, 1971, p.7). In other words, they are capable of being concerned about themselves. Hence Frankfurt's second-order volitions provide a good test to see if Parfit's imaginary beings, viz., the members of the selfless community, are similar enough to persons, because if a reductionist account for this allegedly uniquely human attribute can be given, Parfit's imaginary beings will succeed in fulfilling one important condition for being considered to be similar enough to persons. This is the incentive behind appealing to Frankfurt's second-order volitions. As for the psychological states of hope and regret, I take them to be two different ways of expressing one's wish to be different in one's preferences and purposes from what one is. In regret the wish is retrospective and in hope it is prospective. Hence in a manner of

speaking, regrets express retrospective Frankfurtian second-order volitions, and hopes express prospective second-order volitions.

Some of Frankfurt's remarks are very similar to the criticisms brought up against a reductionist view of persons and personal identity. Frankfurt (1982) in drawing a distinction between caring and having (first-order) desires and beliefs says,

Desires and beliefs can occur in a life which consists merely of a succession of separate moments, none of which the subject recognizes – either when it occurs or in anticipation or in memory – as an element integrated with others in his own continuing history. . . . The moments in the life of a person who cares about something, however, are not merely linked inherently by formal relations of sequentiality. The person necessarily binds them together The outlook of a person who cares about something is inherently perspective; that is, he necessarily considers himself as having a future. On the other hand, it is possible for a creature to have desires and beliefs without taking any account at all of the fact that he may continue to exist. (p.83, my italics)

Frankfurt's (1988) criticism of formal relations of sequentiality as a sufficient condition of caring⁶⁵ is very similar to Schechtman's (1996) criticism of the reductionist view of personal identity as an identity theory that can justify some of the very basic features of human experience like self-interested concern. Schechtman offers an alternative approach which is the narrativity thesis to explain the four basic features of human experience which are compensation, moral responsibility, self-interested concern,

⁶⁵ For my purposes, caring just is having second-order volitions. There may be certain differences between the two that matter in other contexts, but not here.

and survival. Her narrativity thesis consists precisely in what Frankfurt holds as a condition for having the experience of caring, which is an understanding of one's life as having a past and future. I address Schechtman's arguments and her criticism of the reductionist account in chapter 4, but for the moment let me just say that those formal relations of sequentiality that both Schechtman and Frankfurt seem to belittle do more than what meets the eye and should not be oversimplified. I come back to this in chapter 4, but before articulating the importance of relations of sequence, and arguing for a reductionist account of self-concern, I'd like to briefly suggest a first-order way of thinking about second-order volitions.

In being able to see that some of our desires are more important than others, thus pursue them instead of others, it is not so clear whether one needs to formulate this ability in terms of seeing oneself as the subject of multiple desires and then forming a *second-order* desire over the first-order ones. Instead, given contradictory desires, one can articulate the process of acting in accordance with one desire and ignoring the other in terms of recognizing something about the former desire, e.g. that it is a good desire, or a better desire, without positing second-order desires, in pretty much the same manner as noticing one cup to be full of water and another one empty and reaching out for the full one when one is thirsty.

For instance, consider that I care about being successful at work so I stay late at my office on a Friday night instead of going out even though I want to go out and have fun. When thirsty, my perception of the two cups is pretty much enough for me to reach out for the full cup. However the decision between going out or staying at the office is

going to be influenced by a number of factors such as the amount of work completed in that week, the level of fatigue, the level of inefficiency due to fatigue, the particular plan for that night, the people involved, my relations to them etc. As such the decision is not immediate and it is not obvious.

Nevertheless, the complexity of the situation is not a reason – or at least not an obvious reason – for assuming that a fundamentally different perspective that involves a conception of oneself as the subject of one's desires is required for experiences of these kinds to be possible. I am not saying that such a perspective would fail to explain the phenomenon of deciding to stay late at the office. On the contrary a psychological account involving such a perspective, viz. the perspective of oneself as the subject of one's experiences, is likely to succeed in explaining self-ascriptive states. But I am suggesting that this success comes with a cost, that is the cost of a metaphysical illusion. And it should be noted that my purpose here is not to suggest a full psychological account of self-ascriptive states but to offer a way to accommodate and understand them without positing formal conditions and things that we can't stand behind. In other words, I'm suggesting a way to understand them with respect to a certain metaphysical view of persons and personal identity, which is the reductionist view.

4 A REDUCTIONIST ACCOUNT OF ONE'S CONCERN FOR ONE'S FUTURE

Introduction

We have strong intuitions about the importance of identity, and many of our daily practices are based on them. One's concern for one's future constitutes one such case. It's been claimed that one's concern for one's future – which is called the special concern – cannot be justified if one endorses the reductionist view of personal identity.⁶⁶ Parfit (1984, p. 307) calls this the extreme claim. The extreme claim is one of the reasons why reductionism as a thesis with respect to persons is not very welcomed. As a result, although reductionism is taken seriously since it is after all a strong alternative to theories that defend Cartesian egos in disguise, it's hardly been endorsed as a thesis that can be made compatible with the types of concerns and features that lie at the very core of human experience such as the special concern. Instead, when reductionism is accepted, it is treated as a highly abstract metaphysical theory of identity that leaves common human experiences to be accounted for by other disciplines and theories.⁶⁷

It is not clear why the extreme claim should hold, and even if it does hold, why on those grounds reductionism should be refuted. As such, the problem with the extreme claim is three-fold. On one hand there is the question of whether metaphysical theories of personal identity should have implications on matters like the special concern. On the other hand, granted that it is reasonable to expect theories of identity to have some

⁶⁶ Such objections date as back as Butler's objection to Locke's account. See Kind (2004) for a discussion. Also see Schechtman (1996), Langsam (2001), Whiting (1986), and Wolf (1986).

⁶⁷ See Wolf (1986) for a discussion of why reductionism as a metaphysical thesis has no implications for her interest in persons.

bearing on attitudes like the special concern, there is the question of whether reductionism is compatible with and can explain the special concern. Last but not least, granted that the extreme claim is correct, there is the question of whether reductionism should be refuted as a consequence of the extreme claim or whether the special concern should be taken as an illusion.

I believe that not enough attention has been paid to how in fact the reductionist account of personal identity can *explain* if not justify the special concern. It is thus my purpose to undermine the implications of the extreme claim by providing a reductionist account of the special concern, and suggesting that a non-reductionist view is no better than a reductionist one when it comes to *justifying* the special concern. I assume that identity theories should provide some explanation for issues arising from the importance of identity such as the special concern. With this assumption, I first look at the main motivation behind the extreme claim, which is that continuity between psychological states does not give one a good reason to be concerned about one's future, and since the reductionist view of identity is based on continuity, it cannot explain the special concern. I then consider other personal identity theories that are not based on continuity. I argue that while they may succeed in explaining the special concern, they fail as identity theories. Given that my purpose is to see if an identity theory can provide an explanation for the special concern, it is not possible to fall back on an account that explains the special concern but fails as an identity theory. I then give an argument for why continuity alone is in fact important for one's concern for one's future.

4.1 *Is explaining the special concern a responsibility of identity theories?*

As mentioned above, the problem with the extreme claim is three-fold. I start by considering the first, i.e., whether identity theories should have any bearing on psychological attitudes like the special concern. This question applies to all metaphysical theories of identity, reductionist and non-reductionist alike. The extreme claim may be refuted on the grounds that *concern* for one's future is not the kind of thing that would be implied (or can be justified) by a metaphysical theory of identity.⁶⁸ According to this line of thinking, a theory of personal identity may be necessary but not sufficient for psychological attitudes like concern. Supposedly such attitudes can and should be justified on other grounds with a series of additional assumptions. As a result, even if reductionism fails to give an account of the special concern, being a metaphysical thesis, it is supposed to escape the criticism.

This view results from the belief that questions of personal identity can be dissociated from questions about why identity matters and is not a defense of the reductionist account per se. Instead it is just a criticism of the way identity theories are used in relation to issues regarding why identity matters. As a result, the criticism can be used in exactly the same way to defend the non-reductionist account if a non-reductionist account of identity was accused of not providing good reasons for one's concern for one's future.

There's been more emphasis in the literature on showing that the extreme claim cannot be used to reject reductionism since metaphysical theories with respect to persons

⁶⁸ See for instance Wolf (1986).

don't need to have any bearing on *real-life* phenomenon such as one's concern for one's future, and only a few have actually attempted to reject the extreme claim itself by arguing that the reductionist account is capable of justifying the special concern.⁶⁹

Although I agree that no *logical* inference can be made from a theory of personal identity to justifying a particular attitude concerning one's identity such as one's concern for one's future, I also believe that an unbridgeable gap between theories of personal identity and questions of why identity matters is not favorable, especially when the theory is of a reductionist kind. That is why a theorist of personal identity must at least attempt at explaining why there would be such feelings of concern.

Intuitively speaking, one of the most immediate consequences of the belief that we remain the same person over time is one's concern for one's future. So it is only natural to expect, at least pre-theoretically, that a theory of what makes anyone the same person over time to be compatible with issues arising from the importance of identity like the special concern. This expectation is of special importance for the reductionist account since the reductionist account takes the additional step of denying itself reference to anything that may be unnecessary for the explanation of a variety of phenomena including personal identity even though such reference – especially one to a self – may make explanations a lot easier than otherwise. Therefore there is more pressure on the reductionist in terms of how much he can explain. As such it would be significantly to the reductionist's advantage if his personal identity account is compatible with the importance of identity and provides at least an explanation for the special concern and not just be a highly abstract metaphysical theory resting behind the claim that identity

⁶⁹ See for instance Wolf (1986) for the former view and Whiting (1986) for the latter.

theories are separate from theories of why identity matters.

It should be noted however that this is not an argument but a preference. Hence it is my belief that questions of personal identity should not be dissociated from questions about why identity matters and my assumption that identity theories should provide some explanation for psychological attitudes regarding one's identity like the special concern. Hence, although it may not be necessary for the special concern to be *justified* on an account of personal identity for that account to be successful, I will nevertheless provide a reductionist account of the special concern in my attempt to undermine the implications of the extreme claim. Since the non-reductionist is not in a better position when it comes to *justifying* the special concern for reasons that I will discuss in 4.2; henceforth, my criticism of the extreme claim is focused not on the reductionist's inability to *justify* the special concern, but on its alleged inability to *explain* the special concern.⁷⁰

4.2 What is the main reason for the extreme claim?

Before attempting to undermine the extreme claim, a better articulation of the special concern in relation to the extreme claim is called for.⁷¹ I suggest three possibilities and should note that these possibilities are not meant to be exclusive of one another.

- 1) What the reductionist cannot explain is the *special character* of one's concern for oneself as opposed to one's concern for another. One's concern for one's future is special in the sense that e.g. while I would be worried about there being an accident and people getting injured, my worries and thoughts would be very

⁷⁰ See Johansson (2007) for a related discussion of why non-reductionism is not the better thesis when it comes to justifying the special concern. The reasons that I give in 4.2 however are different from his.

⁷¹ See Whiting for a discussion of the different ways the extreme claim is interpreted, p.549 fn 3 in particular.

different if I knew that *I* would be in that accident and the reductionist supposedly fails in accommodating this particular sense of the concern. This sense of the extreme claim can be traced back to Sidgwick's criticism of Hume. Sidgwick (1907, p.419) argues that there is no reason why a *bundle* of perceptions would be concerned in a special way about another *bundle*. One could also argue that there is no reason why one element of a bundle would be concern for another element, or one element would be concern for the whole bundle.

- 2) If one's concern for one's future is derived from one's present or immediate concern for one's *self* there is no way to accommodate the special concern on the reductionist account since on the reductionist account there is no self to which the concern can be attributed to.
- 3) Alternatively, one can assume that one's concern for one's self is accommodated in some sense but the reductionist account of personal identity cannot accomodate one's concern for one's *future* because it does not provide a plausible candidate in the future for which there will be a concern today.

These three conceptions of the extreme claim share the same rationale. Reductionist theories of identity have to do with the holding of certain continuities and it has been argued that the holding of continuities does not explain why we are concerned about our futures.⁷² To use Langsam's (2001) example, if I am worried about going to the dentist tomorrow, it cannot be merely because the person who will be sitting in the dentist's chair tomorrow has some memories, beliefs, desires etc. that are continuous with me. It doesn't make sense to say that I'm concerned about going to the dentist tomorrow

⁷² See Langsam (2001) and Schechtman (1996) for a discussion of related issues.

because the person that will be at the dentist tomorrow has the same childhood memories as I do. I will refer to this reason for the extreme claim as the irrelevance of continuity. If I were told that I would gradually forget all my past by the time I sat in the dentist's chair, I would see no reason why I would be less worried about going to the dentist. The only reason why I might be less worried about the dentist visit is because I'd have more problems to think about.

I will have more to say on whether or not the irrelevance of continuity is a just criticism of the reductionist account in relation to the special concern. However, for the moment my purpose is to better understand the extreme claim. According to the first conception of the extreme claim, there is a difference between one's concern for one's future and one's concern for another's future. As mentioned before, even if I forgot all about my past, I would still be worried about going to the dentist or being in pain. So even if there is nothing left of continuity, there is still the concern. Hence it cannot be the holding of some continuity that explains one's concern for one's future. In comparison, if I completely forgot a shared past with a friend, he or she would become no more than a stranger to me and indeed I would no longer be concerned about my friend, at least not any more concerned than I would be about a complete stranger.

Removing the continuity between the experiences of x and y removes the relation between them, leaving x in a complete state of ignorance with respect to y . But removing the continuity between x 's experiences apparently does not eliminate x 's concern for itself; hence the special character of one's concern for oneself and the reductionist's failure in accommodating it. In short, while the continuity (in the sense of sharing

experiences or a past) may be necessary for one's concern for another, it is not so for one's concern for one's future. Since continuity is what the reductionist account of personal identity is based on, and since continuity is allegedly not necessary for one's concern for one's future, the reductionist account either fails in explaining the concern or the concern itself is an illusion. And this is what the first conception of the extreme claim is based on.

Continuity is the main problem for the second and the third conceptions of the extreme claim as well. According to the second conception, there needs to be an immediately present self to which the concern can be attributed, and according to the third conception, there needs to be a self in the future to which the concern can be attributed. In other words, in both cases, there is a need for a separately existing entity to be the appropriate subject of the concern and since a personal identity theory that is based on continuity does not give this entity, the special concern is not justified.

However, it is not clear how there being a self would justify the special concern, and in fact there is no well-articulated non-reductionist account of the special concern in the literature, except perhaps Langsam's (2001). But his account is based on a very different criticism of the reductionist account of personal identity. Langsam argues that the reductionist's failure in justifying the special concern lies in its general failure in justifying subjectivity, or the what-it-is-likeness of experiences. He claims that the reason why I would be worried about going to the dentist tomorrow is basically that it is going to hurt and according to Langsam without positing a persisting subject of experience, it is not possible to make sense of phenomenal consciousness. Since a reductionist account

does not posit a subject, let alone a persisting one, supposedly consciousness cannot be understood, including what it is like to have pain at the dentist. Hence one is left with no reason to be concerned about the dentist appointment.

I have addressed these issues in several places. I have argued that consciousness is possible without a conception of oneself as a subject of one's experiences in chapter 3. Regarding what-it-is-likeness of experiences I suggested that there is a way to understand subjectivity without appealing to the subject in chapter 1. The main difficulty in interpreting subjectivity in this way was in the difficulty of bringing together the experiences of a single subject without appealing to the subject and that I addressed in chapter 2.

It has just been assumed that the problem is with the reductionist because of the irrelevance of the continuity and then assumed that the non-reductionist would not face these problems since a non-reductionist view of identity is not based on continuity. Instead, it is based on there being something over and above physical and psychological states, separately existing and enduring through time. The fact that it endures through time makes it an appropriate subject of concern. But if the self is something that exists separately and is not influenced by the changes in psychophysical states, since it is over and above them, then how exactly does the self qualify as an appropriate subject of concern? It is not clear what it is about this self that one would be concerned about. It is easy to see how my thought that there is going to be an accident can give rise to a state of fear, but I don't understand how my alleged self which is neither my thought nor my fear is affected by all this.

The immediate response here goes back to Reid (1785/2008, p.109) who suggests that the self is neither the thought, nor the fear. It is what thinks and fears. But as I tried to argue in chapter 1, I agree with Lichtenberg and am very skeptical about the inference from a thought to a thinker. But more importantly in this context, even if there is something that fears going to the dentist, it is not clear why this justifies apprehension about the dentist in a special way. Just because there is something that will have the pain, why does that make the pain worse than occurrence of the pain? The reply that the pain would not exist unless someone has it begs the question, and actually contradicts the assumption behind the extreme claim, namely that the sheer existence of a sequence of experiences, including pain, would not be enough to justify the special concern. If the possibility of such a sequence is rejected a priori, there remains nothing that the argument for the extreme claim can depend on.

There is a further difficulty; even if it was possible to infer from the thought that there is a thinker and from the fear that there is something that fears, it is not obvious that the thinker and the fearer are the same. There would then be a need for a criterion of identity for the self, supposedly the thing that thinks, fears, has desires etc. and is supposed to remain one and the same throughout one's life. Just because it is a separately existing entity and it is over and above psychophysical states, it does not mean that it is going to be the same self for each mental state.⁷³ As such there would be a need for an identity theory for the self in which case we go back to square one, face the exact same questions and with an additional posit in hand to take care of.

Given that the extreme claim is based on the irrelevance of continuity, in what

⁷³ For related discussions see Strawson (1997, 2003) and also Rosenthal (2002).

follows I consider other identity theories that are not based on continuity, but are not as a result non-reductionist either, at least not in the traditional sense. My purpose is two-fold; I consider whether these theories succeed as identity theories *and* as theories that accommodate the special concern.

4.3 Do identity theories that are not based on continuity succeed in explaining the special concern?

One type of personal identity theory that is not based on continuity, nor is *prima facie* non-reductionist, is an identity theory that is based on a particular psychological attitude.⁷⁴ Whiting (1986) for instance argues that one's concern for oneself is *constitutive* of personal identity just as one's concern for another is constitutive of friendship. We are not concerned about our friends because they are our friends. Instead, some people become our friends because we are concerned about them in a certain way. Whiting suggests a similar reasoning for the way we are concerned about ourselves. She says “. . . we make our future selves by coming to care about them in much the same way that we make friends by coming to care about others” (p.566). As such while x 's concern for x is necessary for x 's identity, x 's concern for y is not necessary for y 's identity (but is necessary for the friendship between them).⁷⁵

Whether a particular psychological attitude such as concern for oneself can be

⁷⁴ Although Locke's theory of identity is usually interpreted as the start of psychological continuity theories of identity, I think it is also an example of an identity theory that is based on a particular psychological attitude where the attitude is that of remembering.

⁷⁵ This resembles the first conception of the extreme claim insofar as it suggests a difference between one's concern for one's future and one's concern for another's. But it should be clear that the difference on Whiting's (1986) account has nothing to do with the reason for the extreme claim. On Whiting's account, one's concern for oneself is constitutive of one's identity, hence is fundamentally different from one's concern for another.

taken as a criterion of personal identity is an interesting question. While Whiting (1986) does argue that concern for oneself establishes the future selves of a person to be concerned about, she does not aim at articulating a particular personal identity theory based on this idea. Her main purpose is to address the extreme claim by using friendship as a model to account for the type of concern one has for oneself. As such the question of whether or not a particular psychological attitude such as concern for oneself can be taken as a criterion of personal identity is not fully discussed. But there is another well-articulated personal identity theory that is based on another psychological attitude and it is fully developed by Schechtman (1996). Her personal identity theory is based on having a narrative outlook and she argues that one's concern for one's future along with other issues arising from the importance of identity such as moral responsibility and compensation can be justified only by way of having this narrative outlook.

In what follows I will evaluate Schechtman's (1996) narrativity thesis as a personal identity theory that succeeds in justifying one's concern for one's future.

4.3.1 Schechtman's narrativity thesis

Schechtman (1996) argues that the importance of identity⁷⁶ cannot be captured by identity theorists who ask the reidentification question, which is a “question of what makes a person at time t_2 the same person as a person at time t_1 ” (p.1). Instead the characterization question (CQ), which is a “question of which beliefs, values, desires, and other psychological features make someone the person she is” should be asked (p.2)

According to Schechtman (1996), it is possible to answer CQ only by way of

⁷⁶ Schechtman (1996) suggests four basic features of personal existence arising from the importance of identity which are survival, moral responsibility, self-interested concern, and compensation.

having “an implicit understanding of one’s history as unfolding according to the logic of the story of a person’s life” (p.113-4) and coming to think of oneself as a persisting subject who has had experiences in the past and will continue to have experiences in the future (p.94). By way of having this narrative outlook on one’s life, one can not only distinguish essential features that make one the person that one is from accidental events in one’s life, but also acknowledge the influence of the different times of one’s existence on one’s present.⁷⁷ For instance a surgeon by having a narrative outlook will know that liking strawberries is accidental to who he is but being a surgeon isn’t.⁷⁸ The surgeon will also know to what extent certain events in his life have influenced his present and by extrapolation he will have an idea of how his present can further influence his future. This gives rise to a meaningful concern for both the past and the future and according to Schechtman, it is not possible without the narrative outlook.

CQ asks whether a certain experience, thought, action or a mental state can be attributed to a particular person. It aims at finding those “characteristics that are *truly* those of some person as opposed to ... those which are his as a result of hypnosis, brainwashing, or some other form of coercion” (Schechtman, 1996, p.73). It tries to capture the distinction between states that merely occur in one’s history and states that make someone the person that one is and thus constitute personal identity.

Such a distinction is crucial especially with respect to our moral judgments. Consider for instance two men who stole some money. They are both caught, and a just

⁷⁷ Schechtman (1996) says, “the past is reinterpreted and experienced in a new light in virtue of the present; the expectation of the future gives a different taste to current experience; and future experiences will have their character within the context of the whole” (p. 144).

⁷⁸ Or maybe it too is accidental to who he is.

punishment is being decided upon. It may be the case that one of these men has been a thief for a long time, he doesn't like to work and has made stealing a way of getting on in his life, whereas the other man has worked hard for years, had never stolen before, but due to some extreme circumstances, he stole only this once. By asking whether the crime is attributable to the person we're going to punish, the characterization question brings out the differences between the two men. As a result, the jury would probably think that the act of stealing is not really attributable to the latter man, that he may deserve some punishment since he did steal, but perhaps not a severe one, at least not as severe as the other would deserve.

On CQ, the question "Is this the person who committed the crime?" becomes "Is this past state attributable to the person that we're going to punish?" (Schechtman, 1996, p.90). On RQ the question becomes "Is the person we plan to punish the same person as the person who committed the crime?" (p.90). The weakness of RQ, according to Schechtman, is twofold. Firstly the form of RQ does not admit of degrees, and calls for a yes or no answer, where neither would do justice to the one-time thief.⁷⁹ On the other hand, the answer to CQ admits of degrees, calls for qualifications, which reflect the differences between the two men, thus accommodating the differences in our moral judgments with respect to the two men. That's how we know that the person who stole only once in his life requires a different moral judgment than the one who has turned it into habit.⁸⁰

⁷⁹ If yes, due to lack of tolerance, if no due to too much tolerance.

⁸⁰ Likewise a kleptomaniac would require a totally different kind of moral judgment.

Secondly CQ establishes a relation between a person and a state that is his.⁸¹

Schechtman (1996, p.90) argues that with respect to RQ “we can only attribute the crime to the person receiving the punishment *indirectly* – it must first be attributed to a past subject and then attributed to the present subject only via its connection with the past one.” But with CQ one can attribute this past experience directly to the person. In this way we provide the person “with a strong enough relation to those actions and experiences to justify our judgments of moral responsibility, self-interested concern, compensation, and survival” (p. 90-91). A Relation is established between the person and the state, rather than the person now and the person then.

4.3.2 Criticism of the narrativity thesis

It should be clear by now that CQ has its uses. Remembering that there are two questions intertwined in the question of personal identity, viz. the question of what makes anyone the person that one is and the question of what establishes the continuity of the person that one is, CQ seems to be serving the first question with no bearing on the continuity question. However, in order to address the question of personal identity, one must address the continuity question as well. And even though CQ is related to personal identity because it addresses the question of what makes anyone the person that one is, I contend that it can do so only by relying on the reidentification question, hence on continuity.

Consider the man who stole only once. The reply to CQ – is this past state attributable to the person that we’re going to punish – is not possible without covering a

⁸¹ By saying that the state is his, Schechtman probably means that only the minimal requirement, viz., that it occurred in that person’s history, is met.

certain period of time during the man's life such as the times when he did not steal and the times when he was in the extraordinary circumstances that caused him to steal. All this information is relevant to answering CQ and can only be obtained from the reply given to the reidentification question and not from the reply given to other characterization questions. This shows that it is possible to reply CQ only by knowing what the answer to the relevant RQ is. In other words, numerical identity is a logical requirement for personal identity and it is possible to determine it only by providing an answer for RQ and not CQ.⁸² So while the characterization question undoubtedly has its uses and helps to answer the question of what makes anyone the person that one is, it does so only by presupposing numerical identity throughout one's life or parts of one's life that seem to be relevant to the characterization question. The narrative outlook or concern for oneself can only be based on and not replace the logical requirement of numerical identity.

Besides, there are further difficulties with CQ even in cases where it seems to be most useful, because at times the characterization question takes contradictory answers. Going back to the example of the two thieves, imagine that the man who stole for the first time later found stealing rather an easy way of solving his problems, and gradually turned it into a habit. It would be much more difficult to answer the question "Is this past state attributable to the person that we're going to punish?" in this case. The difficulty arises from the states associated with the first act of stealing, such as a kind of nervousness that is related to a sense of guilt, or a heavy feeling of conscience. It would no longer make sense to attribute such states to the type of person he's become, viz. the experienced thief.

⁸² And that is exactly why RQ calls for a yes or no answer because it determines numerical identity.

So while the first act of stealing is readily attributable to the person he has become, psychological states that are closely associated with it aren't.

Consider further that the man is brought to court sometime after his first theft, but before he turns into a genuine thief. He is not found guilty due to the circumstances. Even though terribly frightened at the time, having seen that he could so easily get away with this crime, he is tempted and then he steals another time. This time he is not even caught. So he carries on stealing. After say five felonious acts, he is caught and brought to court. Should that first time when he stole and was found not guilty be taken into consideration in this court? It was a consequence of CQ, we are assuming, that the first crime was not really attributable to him. But what now? Should the man be guilty of five or six crimes? These are interesting questions and suggest that the answers given to CQ are not as helpful as they are claimed to be.

4.3.3 Personhood and Personal Identity

According to Schechtman (1996), the significance of the characterization question and consequently her narrativity thesis seem to be supported by a different discussion viz. Frankfurt's (1971, 1988) discussion of second-order volitions. Schechtman appeals to Frankfurt (1988, p.84) who argues that first-order desires which are endorsed by second-order desires are more truly one's own. That would be one way of answering the question of which characteristics can be attributed to a person: those that are endorsed by one's second-order desires / volitions. However, Frankfurt's discussion is not directed to personal identity. Instead it is about personhood, hence directed to the question of what the distinguishing properties of persons are or how persons differ from other sentient

beings. Indeed Frankfurt never claims that his views constitute a personal identity theory.

I contend that both in Whiting's (1986) and in Schechtman's (1996) discussions, personal identity theories are confused with personhood theories. This is not surprising given Locke's influence on personal identity theories and his emphasis on how the idea of identity should be suited to the idea of the sort of thing it is applied to. Locke (1694/2008) argues that it is one thing to be the same substance, another the same man, and a third the same person; hence identity conditions should be different for all of them. For obvious reasons we do not want to say that a person who has lost his arm as a result of some accident is no longer the same person who has been in the accident, because personal identity conditions are not and should not be the same as the identity conditions for a mere lump of matter, or for an organized body, for that matter. While this is a perfectly plausible point, if misinterpreted or rather overemphasized, it holds the risk of putting too many restrictions on identity conditions. This is especially true given the fact that we do not have a good grip on what exactly it means to be a person.

That is why there is a need to keep the two questions, viz. the question of personhood and the question of personal identity apart to a certain degree. I also agree with Olson (1997), that in talking about problems of personal identity, it should not be granted that "whatever is once a person must be a person throughout its career" (p.24). It is possible to deny personhood to infants, or to Alzheimer's patients in their late stages or people in deep coma. But since infants become persons, and persons become Alzheimer's patients, one must acknowledge the fact that there are gradual changes involved with respect to personhood and I don't think it makes sense to give different criterion of

identity for the different stages throughout one's life depending on whether or not personhood requirements are met during those stages. On the contrary, a theory of personal identity that is compatible with the importance of identity should be based on a criterion that allows a connection between such stages. This is one of the reasons why I take spatio-temporal continuity to be a more reliable criterion for personal identity.

I do not wish to undermine Locke's (1694/2008) emphasis on the coherence between identity conditions and the sort of thing to which those conditions are applied. In developing the personal identity criterion that I endorsed in chapter 2, I have taken note of the fact that persons are essentially the kinds of things that are capable of psychological states and based the spatio-temporal continuity on this very fact. But to further specify a particular psychological state, or an attitude in the name of developing an identity theory turns that theory into a theory of personhood. Both Schechtman's (1996) and Whiting's (1986) supposedly identity theories take that additional step of specifying a particular psychological attitude. But, one neither has to have a narrative outlook with respect to one's life, nor a concern for one's future in order to continue to exist.⁸³

Consider Frankfurt's (1971, 1988) example of three drug addicts; the unwilling addict who wants to quit drugs, the willing addict who doesn't, and the wanton to whom it makes no difference whether or not he quits drugs. Frankfurt (1988) says "The unwilling addict identifies himself ... through the formation of a second-order volition ... with one rather than with the other of his conflicting first-order desires. He makes one of them more truly his own..." (p.18). This talk of identifying oneself with one of his

⁸³ For a related discussion against narrativity, see Strawson (2004).

desires, some desires being truly one's own as opposed to others, supports the idea that some psychological states are more essential to the person that one is. In this case, even though a desire for drugs can be attributed to all of the three men, there is a difference in their preference to act upon this first-order desire. Supposedly then, there is a difference in the person that each one of them is. This is reflected in our moral judgments; more sympathy for the unwilling addict, less tolerance for the willing addict, and frustration with the wanton.

Usually when we talk about psychological features in relation to a personal identity theory that is based on psychological continuity, we do so in a loose way since it is the continuity between them that matters. Identity theories based on continuity focus on the continuity between memories, desires, beliefs etc. and don't put any restraints on the psychological states between which continuity exists. For instance we don't judge some of these psychological states to be more important to one's identity than others, even though the question of personal identity is closely related to the question of what makes anyone the person that one is. We just acknowledge that it is psychological states that matter to what makes someone the person that one is and are not concerned about which psychological states matter more for one's identity. So suggesting specific psychological attitudes, such as Whiting's (1986) concern or Schechtman's (1996) narrative outlook seems to be filling in this gap.

While the motivation may seem to be an innocent one, I contend that an attempt to determine an essential psychological state in virtue of which the continuity between other states becomes relevant to personal identity, puts unnecessary restrictions on an

identity theory. When a certain psychological characteristic is taken as a condition of personhood, there is bound to be gaps during the existence of a person, because surely there will be times when the so-called essential psychological characteristic is not going to be met. But more importantly, when the psychological characteristic as a condition, is incorporated into the identity theory, then the theory in turn will fall short of providing a criterion of continuity with regards to the times when the psychological characteristic is not met. One may argue that taking a certain psychological characteristic as a condition of personhood does not imply that the particular characteristic has to be met at every single moment of the person's existence. I definitely agree, and that is precisely why it should not be incorporated into an *identity* theory.

Frankfurt uses the notion of second-order volitions to discuss what it means to be a person and not what establishes the continuity of the person that one is. He (1988) says;

If there is an unresolved conflict among someone's second-order desires, then he is in danger of having no second-order volition; for unless this conflict is resolved, he has no preference concerning which of his first-order desires is to be his will. This condition, if it is so severe that it prevents him from identifying himself in a sufficiently decisive way with any of his conflicting first-order desires, destroys him as a person. (p.21)

But what destroys one as a person does not necessarily destroy the continuity of the thing that one is unless one assumes that the kind of thing that one is has to remain the same in order for it to be continuous.⁸⁴ Frankfurt would certainly agree that a moment of

⁸⁴ I should repeat here that my purpose is not to refute Frankfurt but to criticize Schechtman's appeal to Frankfurt in relation to her narrative self constitution view as an identity theory.

indifference between two preferences, even a prolonged one, does not mark the cessation of personal identity.

I contend that the main purpose of an identity theory is to first and foremost establish the continuity of something. As Locke (1694/2008) argued originally, whether that thing is a person, a man or a substance surely matters to the way a reidentification question is answered; in other words, to the way the identity conditions are chosen. And that is the reason why almost all *personal* identity theories including the one endorsed in chapter 2 appeal to psychological states. But this influence should not be taken out of proportion for various reasons. Firstly we do not have a clear grip on what it means to be a person. On a related matter, whatever the consensus is on a theory of personhood, there are stages in one's life when it is not clear whether personhood requirements are met. For instance, it is obvious that neither embryos nor infants have second-order volitions, but it would not be right to consider them to be discontinuous with the persons that they become, since whatever it is that allows them to have the distinguishing characteristics of persons develop sometime in between.⁸⁵ Hence, in order to address personal identity properly, one has to address continuity regardless of the person's capacity to have a narrative outlook or a concern for its future or second-order volitions or any other psychological attitude for that matter. As such a spatio-temporal continuity criterion is the most suitable candidate for an identity theory that can establish the continuity between these stages. In short, psychological attitudes cannot and should not constitute criterion for personal identity.

It may be argued here that while RQ may be necessary for providing the criterion

⁸⁵ Possibly as a result of developing a fully functioning nervous system.

of personal identity, it does not capture the *importance* of identity and that's what Schechtman (1996) is arguing for, and that was the general problem with the reductionist account in the first place – one form of which is the extreme claim. In other words, the reductionist account may succeed as a personal identity theory but fail as a theory that justifies the special concern. It may be argued that a particular psychological attitude such as having a narrative outlook is required not for identity but for one's concern for oneself.

Firstly, it is not clear how having a narrative outlook can *justify* the special concern. As mentioned before justification is not the problem of only the reductionist account of personal identity. Even if a person did have a narrative outlook, it is still conceivable that he just does not care about what happens to him.⁸⁶ The narrativity thesis can however explain the special concern, but I contend, so can reductionism for reasons that I discuss in the next section. More importantly, my purpose was to see if a successful *identity theory* is capable of explaining the special concern so I am not interested in mere psychological accounts that can explain the special concern. Schechtman's (1996) theory seems to be an account that serves as an identity theory and explains the special concern, but as I argued it is not suitable as an identity theory. It is now time to face the problem squarely and see if the reductionist account can explain the special concern.

4.4 The reductionist account of one's concern for one's future

As mentioned before, the main reason behind the extreme claim was irrelevance of continuity. To repeat an example discussed before, according to the irrelevance of continuity, it does not make sense to say that I would be worried about going to the

⁸⁶ For a related discussion, see Strawson's (2004) criticism of the normative nature of Schechtman's narrativity thesis.

dentist tomorrow because the person who will be sitting in the dentist's chair tomorrow has the same childhood memories as the ones that I have (Langsam, 2001, p.250). While this may sound reasonable at first, there are other considerations that may make the continuity between memories, beliefs, dispositions etc. significant for my concern, especially the continuity between my concern and my prior dentist experiences. Whether or not I have been to a dentist, the number of times I have been to a dentist, the stories I have heard from other people about their dentist visits undoubtedly play a role in whether or not and to what extent I would be concerned about my own dentist visit. Or if it needs to be put in a strictly reductionist language, it is possible to say the following: it is likely that prior dentist visits, knowledge about other dentist visits, that fall on a spatio-temporally continuous bundle would be associated with the psychological state that is associated with the present dentist visit. If there is absolutely no information about dentists and dental procedures, there is not much that a mental state about a first dentist visit can be associated with other than the directions to the right address.

One reason why this is overlooked by the defenders of the extreme claim is due to the way one's concern for one's future is interpreted. Defenders of the extreme claim seem to forget the intentional character of the psychological state of concern; that one's concern is always about something. Surely there are cases when one has a disposition of being concerned or rather agitated without knowing what he is concerned about, but that is not the kind of concern implied by the extreme claim. One is concerned about one's health, financial resources, concerned about how the appointment with the dentist will go etc. We are not *just* concerned. We are concerned *about* something. And when we talk

about being concerned about one's future, it is a combination of things that one is concerned about in the future. In other words, one is not concerned about the future *per se*, but about something in the future. One may be concerned about getting old, not having enough money to take care of one's kids etc.

Since the phrase 'one's concern for one's future' is used in this careless way, without taking note of its intentional character, it does seem reasonable to think at first that the continuity in memories, beliefs, desires etc. does not have much to do with one's concern for the future. But when one takes note of the fact that concern is an attitude directed at something, call it *x*, the continuity between prior experiences involving *x* and one's concern regarding *x* becomes significant.

One's concern for (something in) one's future is basically a matter of anticipation. If I have never been to a dentist before, I am likely to be less worried compared to a person who's been to the dentist before. Or maybe I will be more worried because of fear caused by not knowing. Either way, the particular nature of my concern is explained because the person going to the dentist tomorrow is continuous with the person who has not been to the dentist before. Psychological states in the present including one's concern are the way they are as a result of the psychological states that precede the present, in other words, determined by the particular spatio-temporal path to which they belong. So, yes maybe when the fact that the person who will be at the dentist tomorrow and the person who is concerned about the dentist visit share the same childhood experiences is considered in isolation, such sharing may seem to be irrelevant to the concern. However the sharing of other experiences isn't, especially those that involve dentist visits which is

basically the foundation of why one would be concerned about going to the dentist in the first place. And since both the childhood memories and the prior dentist visits are on the same spatio-temporal path, the holding of continuity does become significant and it explains why one is worried about something in one's future.

The mistake of the defenders of the extreme claim is that they focus on the continuity between some psychological states that seem to be irrelevant to the concern and from that infer the irrelevance of continuity in general. But that would mean that the particular spatio-temporal path to which the concern belongs is irrelevant to the concern. This inference is untenable and it is a way of throwing away the baby with the bathwater.

One may argue here that the continuity matters for the concern only insofar as the person can remember the experiences between which the continuity holds. And this may require having a narrative outlook of the sort which Schechtman (1996) argues for. Well, firstly, there is no doubt that having a narrative outlook on one's life explains one's concern for one's future and this is not something that I deny. But for reasons discussed before, a particular psychological attitude such as having a narrative outlook cannot be a criterion for an identity theory. So while Schechtman's narrativity thesis may explain the special concern, it cannot do so as an identity theory and my purpose is to see if the concern can be accommodated by a theory of identity.

Secondly, I believe that attempts to base personal identity criteria on psychological attitudes rather than on continuity are in fact attempts to make sense of the non-reductionist account of the special concern. As mentioned before, there is no

argument for why the further fact would justify the special concern.⁸⁷ Since there seems to be nothing about the self to be worried about, adopting a psychological attitude may be an attempt to give the self some tangible psychological substance that it would make sense to be worried about it. The main difference between a non-reductionist account of the concern and a reductionist account of the concern is that the former posits some subject to be concerned about whereas the latter, assuming the intentional character of the psychological state of the concern, aims at explaining why and to what extent there would be a concern. As a result, the subject to be concerned about in the former case needs to be fleshed out but the fleshing out in turn needs to go beyond a series of random psychophysical states. Adopting a particular psychological attitude may be one way of doing that. On the narrativity thesis, the particular attitude would be having a narrative outlook. Consequently the self would be a subject that is capable of a narrative outlook. But as has been suggested several times before, it is not clear how the ‘haver’ of a psychological state is supposed to do more explanatory work than the psychological state itself, at least when it comes to personal identity and now the special concern.

As mentioned before, the reductionist’s explanation of why and to what extent there would be a concern based on the intentional character of the psychological state of the concern, does not really amount to the justification of the concern. It is one thing to explain why there would be a worry, and another thing to argue that given the circumstances, there *should* be a worry. And none of the arguments given above seem to

⁸⁷ For related discussions, see Shoemaker (1985), also Johansson (2007). As mentioned before, Langsam (2001) seems to be an exception but his account is based on a very different criticism of the reductionist account of personal identity, hence on a different account of why non-reductionism justifies the special concern which turns on an a particular understanding of consciousness. I have addressed these issues in 4.2.

justify the special concern, because they lack the normative aspect that might be called for for justification purposes. But it should be noted that this is a problem for both accounts, viz. the reductionist and the non-reductionist. And this goes back to a point I made earlier in 4.1, viz. the question of whether or not metaphysical theories of persons and personal identity have any implications on psychological attitudes like the feelings of worry.

I noted in 4.1 that it would be impossible to *justify* the special concern, both on the reductionist and the non-reductionist account. Nevertheless it was important to see if reductionism is compatible with such feelings of worry. So I tried to explain why according to the reductionist account there *would* be feelings of concern, but I do not thereby claim that my arguments involve a reason for why one *should* be worried about one's future. It was important to see if the reductionist account was compatible with the concern, since when reductionism is criticized with respect to the types of concerns and features that lie at the very core of human experience, several issues are run together, and what precisely the reductionist is accused of is never clear. Consequently the non-reductionist's inability to justify the concern is overlooked. However, given the discussions above, I contend that while both the reductionist and the non-reductionist fail in justifying the concern, the reductionist is in a better position of to explain the concern.

To conclude, irrelevance of continuity is not a just criticism of the reductionist account in relation to the special concern because continuity *is* relevant to one's concern once we accept the fact that one's concern is always about something. That is why I believe that the example given before regarding how I would still be worried about going

to the dentist even if I gradually forgot all my past is a mistaken example, especially when one takes note of the fact that in the process I would also forget my past regarding memories and information about dentists and dentist visits. Then, there is no reason to think that I would still be worried since given my amnesia, being told I'll be going to the dentist should not be very different from being told now that I'd be going to the dentist tomorrow to which I'd respond with a blank expression, inquiring what a dentist is.

4.5 Is the special concern an illusion?

As I approach to the end of this chapter, a few words on the third conception of the extreme claim are called for, i.e. granted that the extreme claim is correct, the question of whether reductionism should be denied as a consequence of the extreme claim or whether the special concern should be taken as an illusion. However the third conception of the extreme claim now needs to be clarified in the light of the discussions above. Firstly, reductionism cannot be refuted on the basis of the extreme claim because, as mentioned earlier, there is no logical inference from an identity theory to a particular attitude about one's identity. And this is true for non-reductionism as well.

There may be another sense of the extreme claim according to which one may suggest that reductionism should be denied if the extreme claim were correct. The traditional extreme claim says one's concern for one's future cannot be justified if one endorses reductionism. The new extreme claim may simply say that reductionism is incompatible with the special concern. The question then is, granted that the special concern cannot be explained or even made sense of within the reductionist framework, firstly, should reductionism be denied, and secondly, should the reductionist take the

special concern as an illusion?

Even though I agree that there is no logical inference from an identity theory to a particular attitude toward one's identity, I also suggested that a gap between the types of concerns and features that lie at the very core of human experience such as the special concern and reductionism is not favorable. So, if the new extreme claim were correct, in the sense that there is an unbridgeable gap, reductionism would be seriously undermined, because the reductionist's aim is to reduce persons to a series of related physical and psychological states and one's concern for one's future is one such state. By the same token, the second option, viz., explaining away the special concern as an illusion – granted that there is an unbridgeable gap – is not open to the reductionist. The important point is that false beliefs, and illusions do occur, and their occurrences are just as real as the occurrences of other psychological states. Therefore they must be explained, and should even, at best, be shown to be plausible-seeming. The reductionist would be responsible for providing a detailed account of why and in what way the concern is a misguided emotion instead of merely labeling it as an illusion.

However it should be noted that in giving this account of why the concern is mistaken, the reductionist is not in a position to assume that there is no self and from that conclude that self-related concerns are mistaken. Such an account would be uninformative if not question-begging. So, if the extreme claim is correct – again in the sense that the concern and reductionism are incompatible – and the concern is mistaken, then the reductionist can appeal to non-reductionist habits in thinking that there is an enduring entity and reasons for thinking in this way in order to explain why there is a

concern. While I do not have the space here to discuss the incentives for non-reductionist habits in thinking, let me just note that methodologically speaking, this appeal to non-reductionist habits would resemble Hume's (1739-40/1978, p.253-6) appeal to principles of resemblance and causation in explaining why we have a propensity to ascribe an identity to what are in fact various, separately existing different perceptions. Hume notes that even though we occasionally correct ourselves, more frequently we yield to the action of the imagination (p.254). Likewise there could be psychological reasons for yielding to non-reductionist ways of thinking about our lives.

Still, even if one believed in the existence of this mysterious, persisting entity that is over and above psychological states as a consequence of non-reductionist habits of thinking, as mentioned before, I don't see what it is about this thing that I would be concerned. Either way, this is only a suggestion with regards to the third conception of the extreme claim, that is, granted that the extreme claim is true. I on the other hand have argued that a reductionist account of the concern is possible and that reductionism and the special concern are compatible.

CONCLUDING REMARKS

Inspired by Hume, and more naïvely inspired by the constant failure of persons to become anything different from what they have been shaped into as a result of their past, my purpose in this work was to see whether or not we have any good reasons to think that persons are anything over and above their bodies and a series of interrelated experiences. To think of persons as merely a body and a family of mental states is highly counter-intuitive, but it is not clear at all what more they are supposed to be. Philosophers do not like the idea of persons as Cartesian egos. They do not like the idea of persons as brains either. And rightly so. Positing the further fact in terms of Cartesian egos or brains brings more problems than solutions.

Yet we cling onto the idea of persons as something more than bodies and related experiences, even if we do not know how to articulate what precisely the nature of this *something more* is. Perhaps understandably so, because whatever this further fact is, it seems to be expected to do a lot of explanatory work and it might just be too difficult to find one common essential feature that can be used to define it. Strawson (1999, p.100) lists about twenty five different conceptions of the self. So I argued that if reductionism is to be defended, one needs to render the further fact explanatorily redundant and since there is no one way to do that, I argued for different reductionist accounts: a reductionist account of personal identity, of conscious mental states, and of one's concern for one's future.

However I do not thereby claim that these accounts exhaustively cover all areas where subjects seem to do some explanatory work. For instance, if one endorses the

reductionist view, a person's inability to change can no longer be conceived as a failure. By the same token, their occasional ability to change can not be conceived as a success either, because in the reductionist framework, there is no longer an entity that can change or fail to change, and therefore there is no basis for praise or blame. The list of things there is no longer a basis for hardly stops with praise and blame. After all, there is no longer an entity that thinks, or loves, or hopes, or lies. And our lives might all of a sudden seem too bare. Maybe it is too difficult, if not impossible, to render subjects explanatorily redundant, and even though reductionism may seem to be the metaphysically more coherent thesis, one may argue that it simply is not worth it.

Unfortunately, too much is invested in non-reductionist ways of thinking about human lives, and this way of thinking is further supported by linguistic habits on a regular basis. Even I did not have the patience and the determination to remove all the pronouns from this essay. Or rather, the desire to remove all the pronouns and the patience required to complete the job was lacking. But the more unfortunate fact is the conception of reductionism as a too sterile a thesis and consequently a lack of appreciation for what a more coherent metaphysical thesis might have to offer. According to reductionism, the existence of persons consists in the existence of bodies and interrelated psychological states, and since guilt, pride, shame, hope etc. are among these psychological states too, the reductionist accepts their existence and takes them seriously as things that persons partly reduce to. As I mentioned in chapter 4, the reductionist's aim is not to simply disregard such states as illusions, but to understand and explain them. The ontology of reductionism is rich with perceptions, emotions, thoughts, beliefs,

desires and so on. And it is inconceivable to me how our lives would be impoverished in the absence of selves, but in the presence of endless number of mental states.

It is true that reductionism implies a different way of thinking about many of our concepts, e.g of blameworthiness or praiseworthiness. However this is not necessarily a weakness. When accepted, reductionism is usually treated as a highly abstract metaphysical theory that does not need to be compatible with human experiences. I believe that further work can not only undermine this conception but also offer a more comprehensive way of thinking about human experiences. The reductionist account of one's concern for one's future, insofar as it is not asked of the reductionist to justify the concern, was an initial attempt at undermining such conceptions.

But how would reductionist accounts offer more comprehensive and possibly richer ways of thinking about mental lives without giving in to selves? Consider the example of praise and blame again. Insofar as there are psychological states of guilt, shame, regret, etc., I think it should be possible to accommodate some sense of blameworthiness or praiseworthiness in the reductionist's framework perhaps with the help of some bridge laws. The relational matrix plays an important role in the kind of reductionism that I endorse and its non-atomistic nature is what makes reductionism different from a Humean bundle theory. Given the role of the relational matrix, the reductionist would be interested in why psychological states like regret or shame or pride emerge in one's mental life in the first place. Such questions encourage an understanding of the *gradual* development of the mind during the course of a person's life, not to mention the *gradual* development of the human mind during the course of evolution.

When thinking about mental states in philosophy, we tend to start with already a fully functioning mind which may have its disadvantages because that is simply not the case. Humans are not born with fully functioning minds. Nor did the homo sapiens emerge all of a sudden with their uniquely human characteristics. Experiences taken in this sense are processes. Whatever is happening in mental lives, it is taking place on a continuum and squeezing in subjects or an understanding of oneself as a subject somewhere along that continuum seems not only unnatural but also unjustified. It would be absurd to say that homo sapiens started to diverge from their lineage after their respective subjects emerged. Likewise, it would be absurd to claim that children start to have more complicated experiences around the age of three or four because they can think of themselves as subjects around that age. It is in this sense that reductionism may offer a more comprehensive way of thinking about mental lives, because the notion of a relational matrix calls for an understanding of both how the mind of the human race has evolved and how the mind of an individual develops. I suspect that as there are more answers to such questions, there will be less room and tolerance for subjects, whether they are taken to be distinct entities, or a formal condition of certain types of mental states.

REFERENCES

- Armstrong, D. (1968). *A materialist theory of the mind*. London: Routledge.
- Bercic, B. (2004). Carnap's Loop: Epistemological reductionism about the self and the argument from the unowned experience. *Synthesis Philosophica*, 19(1), 297-305.
- Behrendt, K. (2003). The new neo-Kantian and the reductionist debate. *Pacific Philosophical Quarterly*, 84(4), 331–350.
- Block N., Flanagan O. and Güzeldere G. (1997). *The nature of consciousness: philosophical debates*. Cambridge: MIT Press.
- Butler, J. (1736/2008). Of personal identity. In J. Perry (Ed.), *Personal identity*, 2nd ed. (pp.99-105). Berkeley, CA: University of California Press.
- Byrne, A. (2004). What phenomenal consciousness is like. In R. Gennaro (Ed.), *Higher order theories of consciousness: An anthology* (pp. 203-226). Philadelphia: John Benjamins Publishing Co.
- Campbell, J. (1994). *Past, space, and self*. Cambridge: MIT Press.
- Cassam, Q. (1989). Kant and reductionism. *The Review of Metaphysics*, 43, 72-106.
- Cassam, Q. (1992). Reductionism and first-person thinking. In D. Charles, & K. Lennon (Eds.), *Reduction, explanation, and realism* (pp. 361-381). New York: Oxford University Press.
- Caston, V. (2002). Aristotle on consciousness, *Mind*, 111, 751–815.
- Carruthers, P. (2004). Suffering without subjectivity. *Philosophical Studies*, 120(2), 99-125.
- Carruthers, P. (2005). Why the question of animal consciousness might not matter very

- much. *Philosophical Psychology*, 18(1), 83-102.
- Crane, T. (2003). Mental substances. In A. O'Hear (Ed.), *Minds and persons* (pp. 229-250). Cambridge, UK: Cambridge University Press.
- Damasio, A. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. New York: Harvest Books.
- Descartes, R. (1985). *The philosophical writings Descartes*. Translated by J. Cottingham, R. Stoothoff, and D. Murdoch. New York: Cambridge University Press.
- Frankfurt, H.G. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy*, 68, 5-20.
- Frankfurt, H. (1988). The importance of what we care about. In H. Frankfurt, *The importance of What We Care About* (pp. 80-94). New York: Cambridge University Press.
- Gallagher, S. & Zahavi, D. (2008). *The Phenomenological Mind*. New York: Routledge.
- Garrett, D. (1997). *Cognition and commitment in Hume's philosophy*. New York: Oxford University Press.
- Goldman, A. (1970). *A theory of human action*. Englewood Cliffs, NJ: Prentice-Hall.
- Hume, D. (1739-40/1978). *A treatise of human nature*. Ed.s L.A. Selby-Bigge and P. H. Nidditch, 2nd ed. Oxford, UK: Oxford University Press.
- Jackson, F. (1986). What Mary didn't know. *Journal of Philosophy*, 83(5), 291-295.
- James, W. (1890/1950) *The principles of psychology*, vol. 1, New York: Dover.
- Johansson, J. (2007). Non-reductionism and special concern. *Australasian Journal of Philosophy*, 85(4), 641-657.

- Kant, I. (1781-7/1965). *Critique of pure reason*. Translated by N. K. Smith. New York: St. Martin's Press.
- Kind, A. (2004). The metaphysics of personal identity and our special concern for the future. *Metaphilosophy*, 35(4), 536-553.
- Kriegel, U. (2004). Consciousness and self-consciousness. *The Monist*, 87(2), 182-205.
- Korsgaard, C. (1989). Personal identity and the unity of agency. *Philosophy and Public Affairs*, 18, 101-132.
- Langsam, H. (2001). Pain, personal identity, and the deep further fact. *Erkenntnis*, 54, 247-271.
- Lichtenberg, G.C. (1959). *Lichtenberg: A doctrine of scattered occasions*. Ed. J. P. Stern. Bloomington, IN: Indiana University Press.
- Locke, J. (1694/ 1975). *An essay concerning human understanding*. Oxford: Clarendon Press.
- Locke, J. (1694/2008). Of identity and diversity. In J. Perry (Ed.), *Personal identity*, 2nd ed. (pp. 33-52). Berkeley, CA: University of California Press.
- Lowe, E.J. (1996). *Subjects of experience*. Cambridge: Cambridge University Press.
- Lycan, W.G. (1996). *Consciousness and experience*. Cambridge, MA: MIT Press.
- Markosian, N. (2008). Restricted composition. In T. Sider, J. Hawthorne, and D.W. Zimmerman (Ed.s), *Contemporary debates in metaphysics* (pp. 341-364). Oxford, UK: Blackwell Publishing.
- McDowell, J. (1997). Reductionism and the first person. In J. Dancy (Ed.), *Reading Parfit*. Malden, MA: Blackwell.

- Nagel, T. (1979). What is it like to be a bat? In T. Nagel. *Mortal Questions* (pp. 165-180). New York: Cambridge University Press.
- Nagel, T. (2008). Brain bisection and the unity of consciousness. In J. Perry (Ed.), *Personal identity*, 2nd ed. (pp. 227-245). Berkeley, CA: University of California Press.
- Noonan, H.W. (1989). *Personal identity*. London: Routledge .
- Olson, E.T. (1997). *The human animal, personal identity without psychology*. Oxford, UK: Oxford University Press.
- Parfit, D. (1984) *Reasons and Persons*. New York: Oxford University Press.
- Parfit, D. (1998). Experiences, subjects, and conceptual schemes. *Philosophical Topics*, 26, 217-270. .
- Perry J. (2008). *Personal Identity*, 2nd ed. Berkeley, CA: University of California Press.
- Ramachandran, V.S. and Blakeslee S. (1999). *Phantoms in the Brain: Probing the Mysteries of the Human Mind*. New York: HarperCollins.
- Reid, T. (1785/2008). Of identity. In J. Perry (Ed.), *Personal identity*, 2nd ed. (pp.107-112). Berkeley, CA: University of California Press.
- Reid, T. (1785/2008). Of Mr. Locke's account of our personal identity. In J. Perry (Ed.), *Personal identity*, 2nd ed. (pp.113-118). Berkeley, CA: University of California Press.
- Rosenberg, J. (2005). *Accessing Kant: A relaxed introduction to the Critique of Pure Reason*. New York: Oxford University Press.
- Rosenthal, D. (2002). Persons, minds and consciousness. In R. E. Auxier, B. Smith, L.E.

- Hahn, & C. Korsmeyer (Eds.), *Philosophy of Marjorie Green, Vol. 29* (pp. 199-120). La Salle, IL: Open Court Publishing Co.
- Rosenthal, D. (2005). *Consciousness and mind*. New York: Oxford University Press.
- Schechtman, M. (1990). Personhood and Personal Identity. *Journal of Philosophy*, 87(2), 71-92.
- Schechtman, M. (1996). *The constitution of selves*. Ithaca, NY: Cornell University Press.
- Shoemaker, S. (1970). Persons and their pasts. *American Philosophical Quarterly*, 7, 269-285.
- Shoemaker, S. (1985). Critical notice of *Reasons and Persons*, by D. Parfit. *Mind* 94, 443–453.
- Shoemaker, S. (1986). Introspection and the self. *Midwest Studies in Philosophy X*, 101-120.
- Simons, P. (2000). *Parts: A study in ontology*. New York: Oxford University Press.
- Sorabji, R. (2006). *Self: Ancient and modern insights about individuality, life, and death*. Chicago, IL: The University of Chicago Press.
- Strawson, G. (1997). The self. *Journal of Consciousness Studies*, 4(5-6), 405-428.
- Strawson, G. (1999). The self and the sesmet. *Journal of Consciousness Studies*, 6(4), 99-135.
- Strawson, G. (2003). What's the relation between an experience, the subject of an experience and the content of an experience? *Philosophical Issues*, 13, *Philosophy of Mind*, 279-315.
- Strawson, G. (2004). Against narrativity. *Ratio*, 17(4), 428-452.

- Trefil, J. (1997). *Are We Unique? A Scientist Explores the Unparalleled Intelligence of the Human Mind*. New York: Wiley Publishing.
- Tye, M. (2007). The problem of common sensibles. *Erkenntnis*, 66, 287-303.
- Van Cleve, J. (2008). The moon and sixpence: A defense of mereological universalism. In T. Sider, J. Hawthorne, and D.W. Zimmerman (Ed.s), *Contemporary debates in metaphysics* (pp. 321-340). Oxford, UK: Blackwell Publishing.
- Van Inwagen P. (1990). *Material beings*. Ithaca, NY: Cornell University Press.
- Whiting, J. (1986). Friends and future selves. *The Philosophical Review*, 95(4), 547-580.
- Williams, B. (2008). The self and the future. In J. Perry (Ed.), *Personal identity*, 2nd ed. (pp.179-199). Berkeley, CA: University of California Press.
- Wolf, S. (1986). Self-interest and interest in selves. *Ethics*, 96(4), 704-720.
- Wollheim, R. (1984). *The Thread of Life*. London, UK: Yale University Press.