

Université de Montréal

Derek Parfit and the Central Problems in  
Personal Identity and the Philosophy of Mind

par

Matthew Furgiuele

Département de Philosophie

Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures  
en vue de l'obtention du grade de  
Maître ès arts (M.A.)

septembre, 1999

© Matthew Furgiuele, 1999



01.8998 11m8

B  
29  
U54  
1999  
V.017

University of Minnesota

Neuroscience and the Philosophy of Mind

Abstracts

Department of Philosophy

Faculty of Arts and Sciences

Abstracts of the Philosophy of Mind

Volume 1

© 1999 University of Minnesota



Université de Montréal  
Faculté des études supérieures

Ce mémoire intitulé:

Derek Parfit and the Central Problems in  
Personal Identity and the Philosophy of Mind

présenté par:

Matthew Furgiuele

a été évalué par un jury composé des personnes suivantes:

M. Daniel Laurier, président-rapporteur  
M. Daniel Weinstock, directeur de recherche  
M. Michel Seymour, membre

Mémoire accepté le: 12 novembre 1999

## SUMMARY

The present study analyses Derek Parfit's writings on personal identity. I explain Parfit's use of teletransportation, thought experiments. While accepting their use I point out how Parfit commits the fallacy of extension of terms by concluding from his arguments against non-Reductionism to his particular version of Reductionism. I next examine two of Parfit's claims about Reductionism: that reality could be described impersonally and that personal identity does not have rational or moral significance. Summarizing the historical debate surrounding identity, I point out how Parfit envisages Relation R meeting the objections to Locke. I look also to Parfit's use of quasi-memory, which he derives from Sydney Shoemaker's paper "Persons and Their Pasts". I analyse both Shoemaker's and Parfit's argument pointing out their similarities. I then discuss the difficulties inherent in the use of distant possible worlds. I suggest that Parfit's version of Reductionism might be the justification for his use of such distant possible worlds, and that correspondingly, Shoemaker might be forced to concede the unimportance of identity. Having established that Parfit needs to defend his particular version of Reductionism in order to be justified in using his thought experiments not just to criticise his opponents, but also to defend his own position, I examine more closely the meaning of Parfit's Reductionism, which runs the spectrum from a simple rejection of Cartesianism to the extreme version which rejects outright the importance of identity. In order to understand this latter position, I propose a thought experiment which compares two different scenarios. In one, a man's life's work is fulfilled after his death by using his meticulous lab notes, in the other a man's replica fulfils his life's work after his death. If Parfit is right, the second man has less reason to fear death than the first. If this is true, then Parfit is asking us to reconsider much more than our belief about identity. Our desires become more important than our fulfilment of them. For instance, one would have to care more about the happiness of his wife than that she be made happy by him. In examining Parfit's justifications for such a change in beliefs, I find none satisfactory. Finally, I assume that somehow Parfit's views have a justification, and I come to the conclusion that if they are successful, then far from rejecting Cartesianism they uphold it!

## RESUMÉ

La présente étude, intitulée Derek Parfit and the Central Problems in Personal Identity and Philosophy of Mind, est une analyse minutieuse des écrits de Derek Parfit, en particulier de *Reasons and Persons* et d'un manuscrit non publié de 1998, *Experiences, Subjects and Conceptual Schemes*. Je commence par expliquer et justifier l'usage que fait Parfit des expériences de pensées, et plus particulièrement, de la télétransportation. Tout en acceptant la validité de cette méthodologie, je souligne qu'en concluant à la vérité de sa version du réductionisme par son argumentation contre le non-réductionisme, Parfit commet le sophisme de l'extension des termes. C'est-à-dire qu'il n'est pas logiquement permis de conclure à la validité d'une version particulière d'un type d'argumentation simplement parce qu'il est démontré que le type d'argumentation opposé est faux. Finalement, parce qu'elles s'avéreront plus tard d'une grande importance, j'examine deux analogies de Parfit: celle des lentilles artificielles et celle des nations et des personnes. Je conclus que la première est circulaire et que la deuxième est fausse.

Dans le deuxième chapitre, j'examine deux des postulats de Parfit concernant le réductionisme: qu'on pourrait décrire la réalité d'une manière impersonnelle et que l'identité personnelle n'a aucune importance rationnelle ou morale. Par conséquent, il pense qu'on doit remplacer cette dernière par la continuité et la connexion psychologiques avec n'importe quelle cause. Parce que ce dernier point possède un lien étroit avec les critères psychologiques de

l'identité, j'examine d'abord le débat historique autour de l'identité personnelle, après quoi je signale la manière dont Parfit prévoit que son critère, la Relation R, répond aux objections contre les positions lockiennes. Pour faciliter cela, j'examine aussi le fameux travail de Sydney Shoemaker "Persons and Their Pasts," duquel Parfit emprunte le concept de quasi-mémoire. Après une analyse de la structure de l'argumentation de Shoemaker, j'examine les arguments de Parfit pour l'impersonnalité, soulignant leur similarité avec ceux de Shoemaker. Je conclus ce chapitre avec la suggestion que la version du réductionisme de Parfit pourrait justifier son recours à des mondes possibles éloignés du nôtre. En conséquence, j'observe que si la version de Parfit justifie cela, Shoemaker devrait accepter la non-importance de l'identité personnelle.

Dans le dernier chapitre, j'établis que Parfit a besoin d'une justification de sa version particulière du réductionisme, pour l'utiliser à la défense de sa propre position, et non seulement comme un argument contre les positions contraires à la sienne. J'examine ensuite le sens que Parfit donne au réductionisme, passant en revue la gamme des sens possibles, de la négation du cartésianisme à l'autre extrême, le rejet de l'importance de l'identité personnelle. Pour mieux comprendre cette dernière position, je propose une expérience de pensée qui compare deux scénarios: dans un cas, le chef-d'oeuvre d'un homme est complété après sa mort grâce à ses notes de laboratoire très détaillées, et dans l'autre, son chef-d'oeuvre est terminé par son double. Je souligne que si Parfit a raison, le deuxième homme a moins de motifs que le premier pour craindre la mort.

Cependant, si cela est vrai, la pensée de Parfit exige d'accepter la non-importance de l'identité personnelle, elle exige d'accepter que nos désirs propres ont plus d'importance que nous-mêmes. Par exemple, je devrais désirer plus le bonheur de ma femme plus que désirer qu'elle soit heureuse avec moi. De plus, en examinant les justifications possibles pour un tel changement de croyance, je n'en trouve aucune. Finalement, en essayant de voir ce que nous devrions accepter d'autre si Parfit avait raison, je suppose qu'il existe pour son changement de position une justification que j'ai manquée, et je conclus que si Parfit réussit dans son argumentation, loin de nier le cartésianisme, il l'embrasse.

## ACKNOWLEDGEMENTS

In any worthwhile endeavour, there are always many people to thank. This thesis is no exception. In the first place, I wish to thank John Rist, emeritus professor at the University of Toronto for having introduced me to the writings of Derek Parfit in my final year at that august institution. It is not often that a classicist of his stature takes the time to acquaint himself with the contemporary philosophical literature. My greatest debt to Professor Rist however lies in his having taught me how to think philosophically, and in the way in which he instilled in me a love for my field. Were it not for him, I might never have continued at the graduate level.

Daniel Weinstock, my mentor and thesis supervisor has been that rarest of breeds: a colleague and a friend. Although he was on leave at the Centre for Human Values at Princeton last year, he never failed to respond promptly to my enquiries. Moreover, whenever my professional future was in doubt, he was always ready with an opportune piece of advice, or encouragement. It is refreshing that one whose primary field is political theory could be so insightful in what amounts to a metaphysical thesis.

Perhaps my greatest debt goes to Derek Parfit himself. I first got in touch with him three years ago while still an undergraduate at U of T, and he responded graciously to my telephone enquiries at what amounted to an ungodly hour, given the time zones. Moreover, when I began my graduate studies two years ago, Parfit made time for me to visit him at Harvard, whereupon he gave me photocopies of his unpublished works. Additionally, he agreed to read all drafts of my thesis and to comment on them. This is a promise which he has kept faithfully. The level of detail and the quality of his comments prove that he is not only a great philosopher, but also a great man. He too has encouraged me in my studies, and has also provided opportune advice as well as necessary references. I look forward to working with him in the future.

At the risk of this sounding like philosophical name-dropping, I must also thank several other philosophers for their advice and encouragement. In



particular, I wish to thank Marya Schechtman at the University of Illinois, Chicago for her long and fruitful e-mail correspondence. Sydney Shoemaker proved helpful for that portion of my work which concerns him. Soren Haqqvist and Roy Sorensen were a wealth of knowledge when I first began and my thesis topic was exclusively concerned with thought experiments. Both men generously provided me with copies of their books on the subject. Tamar Gendler, a student of Parfit's also provided me with extracts of her Doctoral dissertation. It is wonderful to see such a spirit of service alive in the academic community. My good friend Jay Budziszewski also went above and beyond the call of duty in providing me with much needed advice on the better organization of earlier drafts. In this regard, I must also thank Anthony Schratz, David Williams and Fr. David Sands, who also commented on the organizational aspects of previous drafts.

If I leave until the end mention of my parents, it is not because their help has been insignificant. I owe my very being to their love for one another, and I owe my university career to their generosity and selfless determination to raise me well. If my parents had raised me to be a thief, I would certainly not now be finishing off this thesis.

To all who have been named, and to the countless others whose help was not directly related to this thesis, or to whom I have forgotten to name explicitly go my heartfelt best wishes.

## TABLE OF CONTENTS

Summary	iii
Resumé	iv
Acknowledgements	vii
Table of Contents	ix
<b>Introduction</b>	<b>1</b>
<b>Chapter I Derek Parfit's use of Thought Experiments</b>	<b>5</b>
Background and Structure	5
The aim of <i>Reasons and Persons</i>	6
A problem of Extension of Terms?	6
A helpful Analogy	7
The Refutation of Non-Reductionism	9
The Purpose of Parfit's Thought Experiments	9
Parfit's Argument: Hypothetical or Actual?	10
The 'Ordinary View' about Personal Identity	12
A Third Possibility	13
Teletransportation	14
The aim of the Thought Experiment	15
The 2 Criteria of Personal Identity	16
Physical Criteria	18
Are Physical Criterists Committed to the Possibility of Separation?	20
Psychological Criteria of Identity	21
Branch Line Cases	22
Photocopies and Identity	23
Results of the Teletransportation Experiment	24
Concepts and Reality: A Distinction	27
Teletransportation and the Unity of Man	28
Artificial Lenses	29
A Problem with the Analogy	31
A Clarification of Reductionism	32
The Analogy with Nations	33
Problems with the Analogy	35
<b>Chapter II What is Parfit's Reductionism?</b>	<b>39</b>
Parfit's Alternative to Personal Identity	41
Locke's Notion of Personal Identity	42
Reid's Objection to Locke	42
Butler's Objection to Locke	43

Parfit's Use of Shoemaker's Response to Butler	44
Shoemaker on Q-Memory	45
Why Shoemaker Claims that Memory is a Circular Identity Criterion	46
Can Q-Memory Provide us with a non-Circular Identity Criterion	47
The Difference between Q-Memory and Memory	48
How Q-Memory Weakens PAC	48
How Q-Memory Claims to Eliminate IEC	49
A Modification of Q-Memory	50
M-Connectivity	50
A Re-examination of Personal Identity in Light of M-Connectivity	52
A Change in our Notion of Identity Resulting from M-Connectivity	52
Shoemaker's Argument Schematized	54
What our Analysis Demonstrates	55
3 Arguments Against Physical Criteria of Identity	56
A Return to Shoemaker's Argument	59
The Distinction between Logic and Metaphysics	61
The Link Between Parfit and Shoemaker	64
Introduction to Parfit's Argument for Impersonality	65
Parfit's Comparison of Conceptual Schemes	66
Discussion of Parfit's Distinctions	66
Parfit's Argument for Impersonality	70
Comparing Parfit and Shoemaker	74
Schematization of Parfit's Argument	75
A Return to Parfit's Comparison of Conceptual Schemes	76
On the Compatibility of the Various Schemes	77
Why Shoemaker's and Parfit's Schemes are not Mere Notational Variants of Ours	77
Whether the Schemes Differ by Addition from Ours	78
<b>Chapter III Does Parfit's Reductionism Justify Impersonal Conceptual Schemes?</b>	<b>84</b>
The story so far	84
Parfit's Need for a Justification of his Reductionism	85
Parfit's Reductionism	87
Reductionism and Cartesianism	89
Reductionism and Nations	90
Reductionism and Separate Entities	92
Reductionism and Survival	93
Identity as a Matter of Degree	95
Identity and Desire	96
Parfit on 'Ordinary' Survival	98
Impersonality Revisited	100

Parfitian Dualism	101
Parfit and Descartes	103
Conclusion of Chapter III	106
<b>Conclusion</b>	107
<b>Selected Bibliography</b>	111

## INTRODUCTION

The ancients and the medievals had no real problems with personal identity: Plato, Aristotle, Augustine, Aquinas, et al. had very sophisticated and well developed doctrines concerning the nature of the person, and of his persistence through time. It was not until the era of decadent scholasticism with the nominalism of Ockam, and the essentialism of Scotus and Suarez, which later found its way into the modern world via Descartes that personal identity emerged as a problem area. Most of us are familiar with Descartes' famous meditations in which he divides the world into two kinds of substances: thinking and extended, and in which he identifies himself with the thinking type. Perhaps we are less familiar, or less aware, because it is so pervasive a view, of how this move on Descartes' part created the problem. It is true that there were dualists before Descartes, Plato after all identified the person with the immortal soul; however, he had no problem with how the soul came to be in the body, for he held (in *Phaedrus*, for instance) that that soul enters a body as a punishment for being insufficiently attached to the forms in its disembodied state. Aristotle, and through him Aquinas and others held to the notion of the person as a hylomorphic union of body and soul, in which the soul was the form, and the body the matter informed. In this way, there developed the notion whereby the person was precisely the composite entity of matter and form (soul and body). While a detailed analysis of this notion would take us too far afield, suffice it to say that for Aquinas, who also subscribed to Aristotle's physics (with the corresponding metaphysical notions of act and potency), the soul as form was actual enough to

survive the dissolution of the body, and was, moreover capable of re-informing that matter (or potency) at the resurrection of the body, which he believed had been revealed by God. Although this latter belief belongs to faith, Aquinas had a metaphysics which would admit of a rational (i.e. non theological) explanation of this possibility. The important point here is that the soul for Aquinas is not a separate entity, in the way in which the ego is for Descartes. While Descartes would claim that he is his ego, Aquinas would not claim that he is his soul. Rather, he would claim that he is a person who has a body and a soul.

I mention the above points as it is important at least to understand that there is a huge difference between the classical doctrine of the soul, as expounded by Aristotle and Aquinas and the modern Cartesian notion of the ego. The difference is that for Aquinas, the unity of the person and as a result the importance of both the psychology and the physiology (to use modern terminology) to the persistence of that person is easily explained. In contrast, Descartes was reduced to the facile claim that the pineal gland was the point of union between body and soul. Although no one continues to accept such a notion, the fact remains that problems of identity are viewed in light of this Cartesian framework.

Starting with Locke, there has been a tremendous amount of discussion concerning what matters in survival and concerning what it is that survives. There are currently two main lines of thought on the matter: those who maintain that what matters in survival is physiological and those who maintain that what matters is psychological. This means that there are two basic criteria for identity:

physical and psychological. The physical criteria tend to go by the name of animalism, and prominent defenders of these views are P.F. Snowdon, Eric Olson, and Judith Jarvis Thompson. The psychological criteria tend to be specific to various thinkers, and do not have organized names. Defenders of these views include historically John Locke and David Hume, and more recently David Lewis (time-slices), Sydney Shoemaker and Derek Parfit. There is also a third group, who tend to mix the two, and these include Peter Strawson, Quassim Cassam and Christine Korsgaard. These thinkers are, broadly speaking, Kantian.

The essential methodology of Parfit is to devise thought experiments the terms of which are acceptable to all parties in the contemporary debate and to arrive at a point in which the holders of the various identity criteria must concede that although it is no longer possible to ascribe identity in the given situation, all that matters according to their own identity criteria is still present. Parfit refers to questions of identity in such circumstances as “empty” that is, we can no longer claim identity, but we still know all there is to be known. If this is the case, then, or so Parfit argues, identity must not be what matters. Although I never explicitly argue for an Aristotelian position, it is certainly in the background, and I several times assert that because the contemporary landscape does not include all of the possible explanations for identity, Parfit is not justified in concluding from the problems he demonstrates in the current debate to the unimportance of identity.

Along with the two basic kinds of identity criteria, Parfit also provides us with the dichotomy of Reductionist/non-Reductionist as a means of evaluating the

problem. Roughly speaking, for Parfit a view is Reductionist if it holds that personal identity need not be determinate in every conceivable circumstance, and non-Reductionist if it does. I take issue with this claim, pointing out that there is nothing in Reductionism which commits us to the view that identity is not what matters.

I begin with the analysis of the teletransportation thought experiment, arguing that it cannot demonstrate all that Parfit claims it does. I then point out the difficulties with his analogy between artificial lenses and replication, and between nations and people. The first I claim is circular, and the second false. Next I look to Parfit's alternative to personal identity, Relation R. I point out how he is indebted to Sydney Shoemaker's work on quasi-memory, and then I compare Parfit's argument for impersonality with Shoemaker's argument for quasi-memory. I examine the problems associated with the use of distant possible worlds in the defence of a conceptual scheme, and I conclude that both are in need of a justification for such counterfactuals. I allow that Parfit's version of Reductionism might be the justification, and I analyze it to see if it is such a justification. By means of my own thought experiment, I conclude that Parfit's Reductionism cannot justify his use of distant possible worlds, and, by way of conclusion I demonstrate that if Parfit does have a justification for his use of distant possible worlds, then, far from being an anti-Cartesian (as he claims to be) he in fact agrees with Descartes on the most fundamental points of dualism.



## CHAPTER 1

### DEREK PARFIT'S USE OF THOUGHT EXPERIMENTS

#### BACKGROUND AND STRUCTURE

The present study, which is an analysis of the thought of Derek Parfit on personal identity begins with an examination of his methodology, before proceeding to an evaluation of his thought. With this in mind, it will be convenient to divide this study into two parts:

- 1) Why does Derek Parfit write about personal identity
- 2) What does his thought actually demonstrate.

This distinction is particularly useful, if, as I shall argue, what Parfit thinks he has demonstrated and what he has actually demonstrated are two different things. Moreover, I hope to show that much of the literature generated by Parfit can be seen to stem from this tension between what he thinks he has proven and what he has actually proven.

I shall briefly treat of the first question by examining the two main streams of commentary on Parfit which treat of his methodology, pointing out where I think they go wrong. I shall then proceed to put his thoughts on identity into context by critically examining the three central elements of the introductory chapter in Part III of *Reasons and Persons*. Having done this, I shall proceed to an analysis of both Parfit's articles previous to *Reasons and Persons*, and his most recent thoughts on the matter. In my concluding chapter, I shall link everything together, by demonstrating how the tension between Parfit's purported aims and

his actual conclusions is at the root both of the vast literature generated by Parfit and of his own utility. If I am right, Parfit may turn out to be a useful guide to questions of identity in spite of himself.

### **THE AIM OF *REASONS AND PERSONS***

The first question of note is why Derek Parfit wrote *Reasons and Persons*. The answer which Parfit himself gives may come as a surprise to many who have read him with great interest. As Parfit is the foremost proponent of Reductionism in thinking about personal identity, it is often assumed without questioning that arguing for Reductionism is the reason behind his writing. As Parfit himself notes, he was "concerned not to argue for a Reductionist view but to discuss the implications of such a view."<sup>1</sup> Moreover, his "argument is a reductio against those who, while rejecting a non-Reductionist metaphysics, continue to hold beliefs that would be justified only if such a metaphysics were true."<sup>2</sup> These remarks suggest two things: first, that Parfit views the question of identity through the dichotomy of Reductionism/non-Reductionism and second, that his methodology is best understood as an attempt to refute non-Reductionism as opposed to an attempt at a positive expression of Reductionism.

### **A PROBLEM OF EXTENSION OF TERMS?**

The first point is important as I believe that in viewing things in this way, Parfit tends to commit the fallacy of extension of terms. He does this in the following way: in examining any position on personal identity, Parfit first

---

<sup>1</sup>Personal correspondence with Derek Parfit, November 28 1998.

<sup>2</sup>*Ibid.*

considers whether it is Reductionist or not. If he considers it to be non-Reductionist, he proceeds to demonstrate the problems inherent in the position through the application of one of a series of thought experiments. Having thus satisfied himself that no such position is immune to his objections, he is left with only Reductionist positions. His next step is to argue for the inherent similarity of all Reductionist positions. In this respect, he often makes such claims as "if we are Reductionists, we should not try to decide between the different criteria of personal identity."<sup>3</sup> In so doing, he has a tendency to move from the abstract concept of Reductionism to his particular version of Reductionism, with no further argument given.

#### A HELPFUL ANALOGY

Permit me the use of a helpful analogy on this point. If we consider for a moment the distinction animate/inanimate, we may conclude that a given thing, (for instance, a flatworm) is animate, but we cannot, from that fact alone consider what kind of animate thing it is. Traditionally, there have been three kinds of living things: vegetative, sensitive and rational. It is not necessary to go into great details about each of these, or to consider the controversy surrounding this classification. Suffice it to say that vegetative life is the most basic life form<sup>4</sup>, and

---

<sup>3</sup>Derek Parfit, *Reasons and Persons*, Oxford University Press 1984. Reprinted with corrections, 1987. 241 All further references are to this 1987 edition.

<sup>4</sup>The classic doctrine has its origin in Aristotle, especially in his treatise *On the Soul*. For simplicity I shall only refer to his distinction between living (animate) and non-living (inanimate) without referring to his various classifications. "We resume our inquiry from a fresh starting point by calling attention to the fact that what has soul in it differs from what has not in that the former displays life....This [power of self-nutrition] is the originative power the possession of which leads us to speak of things as *living* at all." *On The Soul*, 413a21-3, b1-2. In, *The Complete Works of Aristotle*, Jonathan Barnes, ed. Princeton NJ: Princeton University Press, Volume 1, 1995.

therefore, it has been used as the distinguishing characteristic of living things. That is, there is no living thing which does not possess vegetative powers, but there are living things which possess more and greater powers. In the same way, by distinguishing between Reductionism and non-Reductionism in the broadest possible sense, we may conclude that a particular belief is Reductionist, but we cannot, from that conclusion alone determine what kind of Reductionism it is. Parfit, in dismissing non-Reductionism has a tendency to move directly to his own version of Reductionism, without providing any further proof or explanation.<sup>5</sup> To do so is to commit the same mistake as to conclude from the fact that something is living to the kind of living thing it is, without providing further argumentation. As the fallacy of extension of terms involves the terms in the conclusion of an argument having greater extension than the terms of the premises upon which it is based, I hope to show that at the very least Parfit's methodology suggests this fallacy, and that very likely, he himself inadvertently commits it also.

---

<sup>5</sup>A clear example of this tendency on Parfit's part can be seen in the way in which he dismisses as implausible Bernard Williams' arguments for personal identity: "I believe that they [Williams' criteria] are even more implausible than the only other possible conclusion, which is the Reductionist View. We should therefore now conclude that the Reductionist View is true." Parfit, [1987] 239 Notice that, as I have been claiming throughout, Parfit here makes the jump from the implausibility of non-Reductionism to the truth of Reductionism. Additionally, he makes this claim at the abstract level, and, joined with his other statement that Reductionists need not decide between the various criteria, demonstrates the fallacy of extension of terms above mentioned. If we want to make a choice between being a Reductionist and being a non-Reductionist, it may make sense to accept Reductionism on the basis of the implausibility of non-Reductionism; however, this does not justify our accepting any particular version of Reductionism. If Reductionism and non-Reductionism are at opposite ends of the spectrum, then, by the principle of the excluded middle, if one is true the other must be false. I am claiming that even if we concede that they are so opposed, it does not follow from that fact alone that because one is false a particular version of the other must be true. The most that can be claimed is that some version of Reductionism must be true if non-Reductionism is false.

## THE REFUTATION OF NON-REDUCTIONISM

There are two major lines of criticism leveled against Parfit's methodology: that his thought experiments are riddled with problems and that he fabricates the view of ordinary people concerning their own identity, in order to strengthen his own view by attacking this fabricated view. Amongst those who make the first criticism, the most famous is Parfit's colleague at Oxford, Kathleen Wilkes, who has devoted an entire book to explaining why thought experiments in personal identity are not helpful.<sup>6</sup>

### THE PURPOSE OF PARFIT'S THOUGHT EXPERIMENTS

In response to the first set of criticisms, allow me to reiterate that Parfit's thought experiments are meant to be a *reductio* against non-Reductionist positions; especially against authors like Bernard Williams who continue to maintain that there must always be a definite answer to questions of personal identity.<sup>7</sup> While Parfit might seem to base his beliefs about personal identity upon his thought experiments, it would be better to claim that he bases his examples

---

<sup>6</sup>Kathleen Wilkes, *Real People Philosophy of Mind Without Thought Experiments*. New York: Oxford University Press, 1993. cf. especially Chapter 1.

For an example of articles which also argue against Parfit's thought experiments, see in particular Jerry Goodenough, "On the Methodology of Thought Experiments" found on the internet at <http://astro.ocis.temple.edu/~souder/thought/good.pap> also "Parfit and the Sorites Paradox" *Philosophical Studies*, 83 (2) 113-120 and Geoffrey Madell, "Derek Parfit and Greta Garbo", *Analysis*, 45 105-109

Concerning the second criticism, see the article by Sydney Shoemaker "Parfit on Identity" in *Reading Parfit*, Jonathan Dancy, ed., Malden: Massachusetts, Blackwell Publishers Inc., 1997. 140-144 and the article by Mark Johnston, in the same publication entitled "Human Concerns without Superlative Selves", especially his exposition of 'Minimalism' 149-156 and 175-176

<sup>7</sup>"one example [of those who reject non-reductionism, but continue to hold opinions which require non-reductionism] is Williams, who continues to assume that questions about personal identity must be determinate." Personal Correspondence, November 28 1998.

upon his beliefs. Allow me to clarify this point. It might make sense to dismiss Parfit's conclusions as a result of his thought experiments if the purpose of the thought experiments was to provide us with a positive basis for accepting his own position; if, however, their purpose is to demonstrate the inherent difficulties in positions with which he disagrees, then the most we could claim is that he has insufficiently refuted an opposing position. When we view his thought experiments in this negative light, we can see that questions of their reality or actual occurrence are much less important than would otherwise be the case.

If Parfit is trying to establish something positive with one of his experiments, we might demand that they be actually possible; if he is simply refuting another position, he need only show that his opponent's position commits him to the plausibility of the experiment in question, and then, having established this, the argument would be effectively refuted by the problems thus demonstrated. Unless the opponent can show that his position does not commit him to the plausibility of the experiment in question, even if we were to claim that the experiment were not actually possible, it would be an effective refutation of his position.

#### **PARFIT'S ARGUMENT: HYPOTHETICAL OR ACTUAL?**

The temptation to assume that Parfit does rely upon his experiments to draw positive conclusions is compounded by his tendency to consider his imagined cases as possible and thus to write about them as though they are real. This use of actual as opposed to hypothetical reasoning throws us off guard, and

has a tendency to bring about charges of circularity. For instance, Parfit asks the question 'What happens when I divide?' as opposed to the question 'What would happen if I were to divide?' By arguing in this way, one's first reaction is, 'of course, if you accept that your division is possible that is a good example, but you have already assumed what you set out to prove in using such an example. The conclusion is contained in the premise, circular reasoning establishes nothing; therefore your example and hence your defense of Reductionism is wrong.' If Parfit's examples only worked when expressed as real, this charge would be valid, and would seriously undermine his position; however, as I have argued, we can re-phrase his cases in the conditional (which allows us to dispute their possibility) without weakening their force. I shall recapitulate the aforementioned point that the purpose of the arguments concerning personal identity in *Reasons and Persons*, is to demonstrate the inconsistency of abandoning non-Reductionist metaphysics, while still continuing to hold "beliefs that would *only be justified* if such a metaphysics were true."<sup>8</sup>

We might conclude from the above that there is one significant change in Parfit's arguments if we re-word his examples in hypothetical mode. If Parfit's cases are in fact possible, then they do lend support to his particular beliefs; if however, they merely serve as a *reductio* then they only disprove a non-Reductionist position, without supporting his particular Reductionism. I believe that Parfit's arguments only succeed in refuting non-Reductionism, but this is no

---

<sup>8</sup>*Ibid.*, emphasis my own.

objection, as Parfit himself admits that *Reasons and Persons* is more concerned with disproving non-Reductionism than it is with proving his position. That is, we may accuse Parfit of too hastily concluding from the falsity of non-Reductionism to the truth of his version of Reductionism, but this is not an argument against his refutation of the non-Reductionist position. Moreover, as this is the main aim of his arguments, even if we later conclude that he has not done more than this (worst case scenario) or that he has done little more than this (middle-ground) we must conclude that he has been successful in his endeavour.

#### **THE 'ORDINARY VIEW' ABOUT PERSONAL IDENTITY**

Before proceeding to the second dimension of the question why Parfit wrote *Reasons and Persons*, I shall briefly consider the second objection raised against his methodology: that he fabricates or embellishes what he considers to be the 'ordinary view' about personal identity, in order to strengthen his own position. Parfit claims that "most of us have false beliefs about our own nature, and our identity over time."<sup>9</sup> As noted earlier, Mark Johnston and Sydney Shoemaker have argued that Parfit is mistaken in ascribing such a view to most people; however, when we understand the kind of view that Parfit ascribes to most people, and when we understand his view, we shall see both that most people do indeed hold the view that he ascribes to them, and that most people do not (intuitively) hold his view. To state bluntly the view that Parfit considers the

---

<sup>9</sup>Introduction, [Parfit] 1987



default, most people believe that "to the question 'am I about to die?' there must always be an answer, which must be either and quite simply Yes or No."<sup>10</sup>

According to Parfit, in order for the default view to be true, there must be some kind of 'further fact' that explains something which is 'separately existing' in order for this view to be true.<sup>11</sup> Most likely the critics object to the assertion that most people believe that they are such separately existing entities. This is not really an objection, however; as Parfit does not say that people consciously believe this, only that in order for it to be true that there are definite answers to one's continued existence it would have to be true that we are such entities. In other words, it is by logical extension that they can be taken to believe this.

### A THIRD POSSIBILITY

I believe that we need not accept that we are such entities in order for it to be true that our existence is all or nothing. This distinction between Parfit and myself stems from my claim that he is right about Reductionism, but wrong about the way in which it solves the problem cases. In order to fully defend such a position, it is necessary to revert to a pre-Cartesian metaphysics. As it is reasonable to assume that most people either do not uphold such a metaphysics or are not even aware of it, we should not be surprised if most people do unconsciously think that they are fundamentally such a separately existing entity. Even though it may be possible to separate the two beliefs, I think that Parfit is

---

<sup>10</sup> *Ibid.* p. 445

<sup>11</sup> Parfit is not very explicit about what a further fact might be, except in stating that it could be a version of Reductionism different from his own.

right to assert that most people do not, and since they do not, that he is immune from the objection that he has fabricated a "straw-man" for rhetorical purposes. That is, while a sophisticated metaphysical argument can demonstrate that the question of the all or nothing nature of personal identity does not require that that identity stem from some separately existing entity, it is not unreasonable to concede to Parfit that most people do not distinguish the two, and so, in accepting that their identity is all or nothing they also accept that it is caused by some separately existing entity. It is interesting to note, however, that Parfit himself also thinks that if identity is all or nothing, then there must be some such separately existing entity.<sup>12</sup> In the same way as I claim that Parfit concludes too hastily from the falsity of non-Reductionism to his own version of Reductionism, so too does he claim too hastily that in order for personal identity to be determinate must there be a separately existing entity which underpins that identity.

### PARFIT'S METHODOLOGY

Now that I have explained in general terms the structure of Parfit's methodology, clarifying it in light of the secondary literature, I can proceed to examine his methodology in itself. I shall now begin to look at Parfit's own thought, and determine what is important in it. In order to do this I shall examine three things: a) Parfit's teletransportation thought experiment, b) his analogy with artificial lenses and c) his comparison of persons with nations. With regards to

---

<sup>12</sup>"Only if we are separately existing entities can it be true that our identity must be determinate." *Op. cit.* 216

a), I shall first describe the thought experiment, then show what Parfit is trying to demonstrate after which I shall point out where I think that it is successful. Finally I shall point out where I think that it fails. This last element will lead me to discuss his two famous analogies, that of artificial lenses and that of persons as nations. I shall claim that the first is circular and that the second is invalid.

### **TELETRANSPORTATION**

At the very beginning of Part III of *Reasons and Persons*,<sup>13</sup> Derek Parfit uses a thought experiment involving teletransportation. We are asked to imagine a person stepping into a box on Earth, pressing a button and waking up on Mars. This technology is supposed to enable transportation at light speed between places of great distance. Next, we are told of a technological innovation such that one enters the box and finds oneself on Mars, except that this time, one also remains on Earth. In addition to describing the two procedures, Parfit also asks relevant metaphysical questions such as 'Do I survive?' in the first case, and 'If I am here, how can I also be on Mars' in the second. He further complicates matters by stipulating that, in the second case, due to an unforeseen difficulty in the procedure, the person on Earth will soon die. He then asks what the relevant attitude towards impending death should be.

### **THE AIM OF THE THOUGHT EXPERIMENT**

If we ask ourselves what Parfit hopes to establish by these experiments, then we can see that he wants to bring about a complete rethinking of the

---

<sup>13</sup>*Ibid.* 199-201

importance of personal identity. His strategy is two-fold: in the first case he is careful to ensure that all of the parties of the contemporary debate agree to the inherent plausibility of the experiment. That is, he has chosen an experiment which could tell us something useful about our beliefs concerning personal identity regardless of what criteria we use to establish identity. At this point, he is not yet claiming that identity is not what matters. To achieve this, he needs the second case. Having established with the first case that, according to the various criteria for identity teletransportation is plausible, he goes on to construct a second example which, since the first is plausible is also plausible. He then demonstrates how this second case presents insoluble difficulties for those who consider identity to be what matters. He then goes on to assert that we should no longer consider identity to be what matters, and he proposes a substitute. The consideration of his chosen substitute will form Chapter II, and so it need not concern us here, except insofar as we may note that there is a further possibility, namely that Parfit has overlooked some other criterion of identity which may be able to respond to the objections raised by the experiment in question.

## **THE 2 CRITERIA OF PERSONAL IDENTITY**

In order to establish that the chosen experiment is plausible in the light of current theories of criteria of personal identity we must briefly state the competing criteria. Since Descartes there has been a great deal of thinking devoted to the very nature and structure of man. In light of Descartes' classification of himself as a thinking substance, and in light of his distinction between thinking and extended

substances, there has been much thought concerning what exactly we are. From this background there have emerged two clear streams: those who think that there is something physical which is the basis of our identity, and those who think that there is something mental which guarantees identity. Theories of the former type are called physical; whereas theories of the latter type are called psychological. It is important to note that in the aftermath of Descartes, philosophers seem to have abandoned hope in resolving the problem of the unity of the various elements which constitute a man. If it is true that no one accepts the facile Cartesian solution of the pineal gland being the point of union between body and spirit, it is also true that the dominant trend has been towards monism. I mean by this that in seeking for a basis of personal identity, the essential Cartesian structure has been adopted, but in the light of Descartes' failure to unify the component parts of man, philosophers have chosen either a physical or a psychological basis for identity and are thus unable to account for the unity of man. That is, there have been attempts either to reduce man to a purely material substance, and so to identify everything of importance with something corporeal, typically the brain, or else to focus on the mental processes of man, without being committed to the claim that it is the brain (or some other organ) which is the necessary pre-requisite for this continuity. The consequence for all of this is that, at least in principle, we can conceive of cases in which identity may be transferred; that is, of cases in which the whole man (psychology and physiology, mind and brain, or body and soul) is disunited, but in which his identity still remains. Just as Descartes

identified himself with his ego, and so was at a loss to explain the need of a body for identity, so too, in the contemporary debate do we encounter this difficulty. The various parties to the debate on personal identity have constructed intelligent coherent criteria for identity based upon either the physiology or the psychology of the person, but neither camp has been able to unify the two. The closest the various theorists have come is to avoid the problem by denying the distinction. As will become clear, however, even those who attempt this reduction<sup>14</sup> are unable, in so doing, to provide a unified account of the person.

### **PHYSICAL CRITERIA**

The above is badly in need of clarification. If we look first to physical criteria of identity, we can see that, as their name implies, the identity of man is constituted by something corporeal: usually the brain, but perhaps something else such as the central nervous system or even the whole body. The essential point here, as far as our evaluation of Parfit is concerned is whether or not we can envisage a transfer of identity. To clarify the notion of transfers of identity, we should consider two possible scenarios, one of which involves the transfer of only a part of the person, but the part which is held to be crucial for identity, the other of which involves the transfer (in some way) of the whole person. In the former case, we would have to look at something like a brain transplant or a case of higher-brain death. If one holds that the brain is the basis of identity, then it

---

<sup>14</sup> I am not here using the term reduction in the sense of Reductionism; rather I mean it in the straightforward sense of physicalists who claim that we are our bodies (pure physiology). The psychological theorists, as will emerge throughout the discussion, are essentially heirs to Descartes' problem, in that they can envisage a person's psychology being transferred to another body, and so are at a loss to explain the relevance of a particular body to a particular psychology.

would seem that wherever the brain goes so goes the person. We can now draw two conclusions: in the first place, we could imagine (at least in principle) that science could advance to the point at which the brain can be separated from the body and either transplanted into a different body or else kept functioning in a laboratory. In either case, the person would be said to have survived the procedure and we would then have to say that the new body has now become the person whose brain it houses, or else we would have to say that the brain in the laboratory is the person, even though it lacks a body. In the other case, where the higher-brain functions have irreversibly ceased, such that the body still remains alive in the sense that respiration and digestion still occur, but that thought is no longer possible, we would claim that the person ceases to exist although the body is still alive. Again, the point here is simply that if we wish to reduce the person to a particular element of his physiology, then we must accept both that he exists only insofar as that element is present and that we can envisage a scenario whereby either that element is present elsewhere than in the original body or else whereby the body remains alive in the absence of the element which gives identity. Thus, if the brain is sustained independently of the body, the person is also independent of the body; conversely, if the body is sustained independently of the higher-brain functions, the body is independent of the person.

The other scenario, in which the whole body is transferred would be the case of teletransportation. Teletransportation is different from ordinary transportation as it involves the dissolution and reconstitution of the body. It is

not simply a matter of local motion, as it involves a break in the direct temporal continuity of the body such that there is a moment in time in which the body does not exist and a future moment in which it exists again. The crucial question here becomes whether it is plausible to claim that such temporal continuity is a necessary condition for personal identity. Even if the answer is yes, this would not imply that teletransportation is impossible, only that the person would be destroyed and replaced by an exact replica, such that although everything which constituted the person continues to exist, the person does not. This is certainly what Parfit wants to argue, and the question should now be reformulated in the following manner: if all of the elements of the person continue, but the person does not, is personal identity really what matters? That is, unless we wish to argue that there is some essential unity to the person such that either all of the person is present or none of the person is present, we would have to accept that teletransportation is possible, and that either identity is not what matters, or else that identity is what matters, but that direct temporal continuity is not necessary for that identity.

#### **ARE PHYSICAL CRITERISTS COMMITTED TO THE POSSIBILITY OF SEPARATION?**

Most physical criteria theorists are committed, in principle, to the possible separation of the person from the body. While brain transplantation and teletransportation have not yet occurred, we certainly have examples of higher-brain death, and the overwhelming majority of these theorists claim that when the cerebral cortex irreversibly ceases to function, the person ceases to exist. Thus,



while there is not yet evidence of the brain surviving the death of the body, there are certainly cases of the body surviving the death of the cerebral cortex. If it is plausible to claim that the former is simply the flip-side of the latter, then, if one is willing to accept the former, he should be willing to accept the latter as well. Finally, in the case of teletransportation, the question boils down to whether or not direct temporal continuity is an essential aspect of personal identity. If it is not, then we can continue to hold that personal identity is what matters, but we must also hold that teletransportation is possible. If direct temporal continuity is essential to personal identity, then we should claim, with Parfit, that personal identity is not what matters; unless, of course we have some basis to argue for the unity of the person, such that all of the person is present or none of the person is present. As I have claimed, the attitude of physical criterists towards irreversible cessation of cerebral cortex functioning suggests that they do not have a basis for the unity of the person.

### **PSYCHOLOGICAL CRITERIA OF IDENTITY**

Let us now turn to the psychological criteria and see if they are any more successful in denying teletransportation. The difference between these criteria and physical ones is that instead of looking for a physical basis of personal identity, they look for a mental or psychological one. Typically these theorists look to memory or thought or conscious deliberation as the basis of identity. If the physical criterist looks for the seat of mental activity as being all-important, the psychological criterist looks to the activity itself as being essential. In this

way, we can see that if anything, transfer of identity is more easily accepted by these theorists. The reason for this is straight-forward: if links between memories, or desires or previous actions is what constitutes identity, and if we leave aside the spatial locus of these memories, desires, etc., then we can easily envisage how that link could change its location without violating the identity of the link, precisely because it is the link and not the location which matters!

### BRANCH LINE CASES

We have been considering the case of simple teletransportation, in which a person is said to have been scanned in one location, and reconstituted in another, such that there is no temporal overlap whatsoever. Parfit refers to this as the *Main Line*.<sup>15</sup> In addition to this, however, he asks us to consider a second scenario, much like the first, except that instead of there being no temporal overlap, the person is teletransported while still remaining in his original location. That is, in this scenario, the scanning and replicating of the person does not involve the destruction of the original person. Parfit calls this scenario the *Branch Line*.<sup>16</sup> The important question for us here is whether we can rule out the second scenario if we have accepted that the physical and psychological criteria commit us to the possibility of the first scenario. The crucial admission, it seems to me is that the first case is plausible. If we are committed to a view of personal identity which reduces the identity of the person to some aspect of that person, whether it be physiology or psychology, then we are hard pressed to claim that a transference of

---

<sup>15</sup>*Ibid* 201.

<sup>16</sup>*Ibid*.

that relevant aspect does not result in a transference of identity. This point is most easily seen if we consider the toughest case, or that in which we maintain that the whole body is required for identity. In this case, we must either make the further claim that unbroken temporal continuity is also required, such that were teletransportation to occur, we would cease to have a case of identity, (although we may have a case of exact similarity); or else we must claim that unbroken temporal continuity is not an essential aspect of identity and so, even accepting that there is not direct temporal continuity between the original person and the replica, identity still holds.

### **PHOTOCOPIES AND IDENTITY**

Even in the most all-encompassing physical criterion of identity we would be unable to concede the impossibility of teletransportation; the most we could claim is that what occurred was not a case of identity transference, only a case of replication. This would be akin to photocopying: no one claims that in photocopying the photocopy is the original; however we do claim that it is an exact replica of the original. Thus, for purposes of historical research, etc., we can do as much with the photocopy as we can with the original. For instance, a professor of English literature could as easily work with a copy of a handwritten Shakespeare manuscript as he could with the original. Assuming the quality and authenticity of the copy, he would not reject it simply because it was a copy. While we would not be able to sell a photocopy of a handwritten manuscript of Shakespeare for the same amount as the original, we could still claim that, as far

as the copy is concerned, it serves the same purpose as the original. If we return to the question of personal identity, then we can see that as far as teletransportation is concerned, the issue is not the plausibility of replication, for we have established that it is plausible; the issue rather becomes whether we have an exact copy (as with the manuscript photocopy) or whether we have a case of continuing identity. The *Main Line* case leaves this an open question, even in light of personal identity being what matters. The *Branch Line* case does not leave this an open question, at least assuming that personal identity is what matters. If the *Main Line* permits the claim that what we have is a case of identity, precisely because we need only remove the requirement that direct temporal continuity is required for identity, the *Branch Line*, in order to salvage our criteria of identity requires that we also remove the requirement of distinctness. If we wish to claim that the *Branch Line* is still a case of personal identity, then we have the problem that a person in two different locations at the same time is still the same person. In other words, we would have to reject the principle of non-contradiction to claim that person A on Earth and person B on Mars are the same person at the same time.

### **RESULTS OF THE TELETRANSPORTATION EXPERIMENT**

I have been arguing that Parfit's use of the teletransportation thought experiments is useful and justified as a critique of the idea that personal identity is what matters, unless of course, we are able to find a unified notion of identity. I argued that the first example, the *Main Line* is plausible on both physical and

psychological criteria of identity. I then argued that from that fact, we could not rule out the *Branch Line*, and that precisely because we cannot rule it out, we are faced with a seemingly insoluble dilemma for theories of personal identity. Now I shall further clarify the way in which this combination of the two cases raises these difficulties. Parfit claims that

while believers in the different criteria disagree about imaginary cases, they agree about what is in fact involved in the continued existence of most actual people. They would start to disagree only if, for example, people began to be teletransported.<sup>17</sup>

The above quote raises two questions: why does Parfit claim that believers in the different criteria would agree about ordinary cases, and why does he claim that they would start to disagree in cases like teletransportation? In response to the first question, we should consider my remarks concerning the post-Cartesian failure to account for the unity of personal identity. That is, with the possible exception of higher-brain death (and, even in this case there appears to be remarkable agreement amongst the various theorists) there do not occur instances in which the criteria one uses for establishing identity yield different conclusions about the identity of a person in question. As things currently stand, the physical and the psychological criteria tend to be co-extensional. It is hardly conceivable under current circumstances that there be a case in which the physical or the psychological aspects of a person could be separated in such a way that according to the one identity would be intact, but according to the other it would not be. This fact alone would not be enough to establish that personal identity is not what

---

<sup>17</sup>*Ibid.* 210

matters, for it could either be the case that the two are always so linked as to be inseparable, or it could be the case that one of the two is an inferior criterion on either scientific or philosophical grounds. It might be the case that as medical science or metaphysics improves we are able to refine our criteria such that one or the other or both criteria are wrong. As I maintain Parfit's first scenario demonstrates, neither of the two criteria can establish the unity of the person, so we need to look to the second scenario. If Parfit could show that one of the two criteria is able to account for identity when there are cases in which the two criteria can be separated that would be an advance of the metaphysical type. If however, he shows that neither one of them can plausibly account for identity, that would serve to strengthen his assertion that personal identity is not what matters.

Barring some means of assuring the unity of the person's psychology and physiology, we would seem forced to concede that identity is not what matters. The reason for this is the following: once we arrive at circumstances in which the criteria we use for ascribing identity make a difference to our ascriptions of identity, they prove incapable of providing non-arbitrary answers to questions such as 'Am I about to die?' or 'Is some future person going to be me?' If under normal circumstances there is no disagreement between the various criteria for identity, and under unusual circumstances the criteria break down, then it is difficult to envisage how identity could be an intrinsic feature of reality. When Parfit speaks about questions of identity being 'empty' he means precisely that

whatever criteria of identity we use the answer we give is going to be arbitrary; that is, nothing is added to reality by either ascribing identity or denying it. If this is true, then, as Parfit tells us, we do not disagree about reality, but only about concepts. Whatever concept we use, the reality described remains the same, and so, the concept is much less important.

### **CONCEPTS AND REALITY: A DISTINCTION**

In other words, Parfit wants to argue that when we consider the difference between reality and concepts which describe reality, there is a lot more agreement between theories of identity than the literature generated would lead us to suspect. To sum up the discussion so far, if questions about personal identity are real questions, and not simply matters of conceptual distinction, then we would expect that whenever we can envisage a case in which different criteria yield different results there is a distinction in reality, as opposed to in our use of concepts. However, as neither the physical nor the psychological criteria seem able to account for any real distinction in the hypothetical cases envisaged, it would seem that, at least according to the two criteria, the distinctions are merely conceptual, and being conceptual, are much less important. Note further that if, in the first scenario the conceptual distinctions seem useful (i.e. the distinction between the need for direct temporal continuity for identity or not) in the second scenario, we need also deny the principle of non-contradiction to continue to hold that identity is what matters.

Once again, the crucial point is that if the various criteria of identity are committed to the idea that whatever is responsible for identity can be separated from what is not, and if, as I maintain, we have ample reason to believe that they are so committed, then, as soon as we envisage a case in which what matters for identity is separated from what does not, we are faced with scenarios whereby whatever we thought guaranteed identity loses the characteristic of distinctness. As soon as we give up the need for distinctness, we are no longer dealing with reality, as the principle of non-contradiction is the bedrock of all reality so anything which requires that it be dispensed with cannot be a matter of any really existent thing. In this way, we can see that whatever labels we may choose to attach to our various hypothetical scenarios (even if they were actually to occur) we would of necessity be describing the same phenomenon (as nothing real is not subject to the principle of non-contradiction). Thus, in the *Branch Line* whether we say that both persons are the same, that neither of the persons are the same, or that one of the two is the same, we are describing exactly the same reality, only using different concepts.

### **TELETRANSPORTATION AND THE UNITY OF MAN**

One last point in this regard. I have been assuming throughout that as soon as we accept that we can separate what matters in personal identity from what does not, regardless of the criteria that we are using, then we must concede that it is possible to envisage situations in which what matters is multiply present. As soon as one concedes that a separation of the elements of a man is possible,



such that there is a division along the lines of what a man is (mind and body, soul and body or psychology and physiology) then it is difficult to imagine how, in light of such a separation it must be restricted to unique instances. That is, if we can envisage this separation, then it is difficult to imagine why it should be a separation which retains a one-one relation. In other words, if separation is possible, then something like teletransportation is possible, and if something like teletransportation is possible then *Branch Line* cases are also possible. Unless we can deny that such a separation of elements is possible, or unless we can deny any significance to the separation of the elements of a man, we cannot hold that personal identity is what matters. I have also suggested that in using either a psychological or a physiological criterion of identity we have already agreed both that such a separation is possible and that this separation has significance. However, I also maintain that rather than concluding from all of this that personal identity is not what matters we should look instead for a concept of identity which keeps intact the essential unity of man, and which can therefore rule out such a separation. To examine this possibility now would take us too far afield, so suffice it to say that Parfit's difficulties arise from his refusal to envisage such a solution to his problem cases.

### **ARTIFICIAL LENSES**

I mentioned earlier that Parfit runs into difficulties by too hastily concluding from his critique of non-Reductionism to his own version of Reductionism, which, in addition to being Reductionist also claims that personal

identity is not what matters. Parfit uses two analogies to make his point about impersonality and we would do well to ponder them even if we disagree that he is justified in making this further claim. The two analogies are that of artificial lenses and that of persons and nations. I shall follow the same argumentative structure with these cases as I have with teletransportation.

Parfit has a marked preference for psychological criteria of identity, and he identifies three possibilities<sup>18</sup>:

*Narrow Criterion:* on this version of the criterion, only the normal cause of continuity of identity is permitted.

*Wide Criterion:* on this version of the criterion, any reliable cause of continuity is permitted

*Widest Criterion:* on this version of the criterion, any cause of continuity is permitted.

According to Parfit, if we accept psychological criteria of identity, the particular cause of continuity should not matter. He explains this through the use of an analogy with artificial lenses. We would do well to ponder the analogy, for it seems to me false. Parfit points out that in the case of sight, if we demand that sight have its normal cause, then even though the results of the artificial lens are the same as, or about as good as, if the person had the use of his normal eyes, we would have to say that he cannot see. This conclusion seems valid, but Parfit extends it wrongly to cover identity as well. The syllogism for vision is as follows:

Whatever allows me to have visual experiences in a way causally dependent on external objects, in a reliable way is as good as sight

An artificial eye allows me to have visual experiences causally dependent on external objects in a reliable way.

---

<sup>18</sup>*Ibid.* 207

Therefore, the effects of an artificial eye are as good as sight.

In the case of the psychological criterion, the syllogism is as follows:

Whatever allows for psychological continuity is as good as personal identity

Replication allows for psychological continuity

Therefore, replication is as good as personal identity.

### **A PROBLEM WITH THE ANALOGY**

The reason the second syllogism seems to me invalid is the following: in the case of seeing with an artificial eye, it is indubitable that I am the recipient of the effects of the artificial eye, thus, it is true that *for me* it is as good as sight. However, in the case of replication or teletransportation, it is far from clear that it is I who am the recipient of the continuity. In fact, Parfit himself argues that I do not survive teletransportation, thus, the effects of continuity are felt by a different person than me. While it may turn out that this is all that matters, i.e. that what matters is the continuity, and not that I be the holder of that continuity (this is Parfit's argument), this conclusion cannot be justified by analogy with the artificial eye. We are inclined to accept that an artificial eye is as good as a real one, precisely because it would have the same effects *for us*. If we were told that we could regain our sight by dying and being re-created identically except for now being able to see, we would probably not accept this. While Parfit may be right that we should accept this, we cannot conclude from the existence of artificial eyes that we should accept death and re-creation for the sake of sight.

Again, I am not at this point rejecting Parfit's conclusions, I am simply arguing that they cannot be justified by analogy with artificial eyes and sight. As the wide criterion would seem closer to re-creation than to abnormal transplantation, the analogy fails. In order for the analogy to work, we must already accept that replication is as good as survival, because the effects are the same, even though the effects occur to different agents. However, this diminished importance of personal identity is precisely what the analogy is supposed to convince us of, so it appears circular.

### A CLARIFICATION OF REDUCTIONISM

In attempting to further clarify the Reductionist position, Parfit tells us that Reductionists claim

(1) that the fact of a person's identity over time just consists in the holding of certain more particular facts.

He may also wish to claim

(2) that these facts can be described without either presupposing the identity of this person, or explicitly claiming that the experiences in this person's life are had by this person, or even explicitly claiming that this person exists. These facts can be described in an *impersonal* way.

Moreover, all Reductionists accept

(3) 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.

Beyond this, some claim

(4) A person *just is* a particular brain and body, and such a series of interrelated events.

whereas others claim

(5) A person is an entity that is *distinct* from a brain and body, and such a series of events.<sup>19</sup>

According to these others,

A person is an entity that *has* a brain and body, and *has* particular thoughts, desires and so on. But, though (5) is true, a

---

<sup>19</sup>*Ibid.*

person is not a *separately existing* entity. Though (5) is true, (3) is also true.<sup>20</sup>

### THE ANALOGY WITH NATIONS

It is true that *prima facie* 3 and 5 seem inconsistent, however, as Parfit points out, we tend to think of nations as distinct from the peoples, territory and activities that comprise them, but do not thereby conclude that they are separately existing from all of this. If we accept that a similar explanation of nations is not contradictory, Parfit says that we should accept the same thing about persons. Yet again Parfit has recourse to analogy, and in this he is indebted to Hume. We can now ask ourselves, what is required for persons to be like nations? Then we can ask must a person be like a nation for Reductionism to hold?

To begin with, we should note that Parfit does not quote the whole of Hume's analogy. He quotes only the beginning: "I cannot compare the soul more properly to anything than to a republic, or commonwealth." However, the analogy continues:

in which the several members are united by the reciprocal ties of government and subordination, and give rise to other persons who propagate the same republic in the incessant changes of its parts. And, as the same individual republic may or may not only change its members, but also its laws and constitutions; in like manner the same person may vary his character and disposition, as well as his impressions and ideas, without losing his identity. Whatever changes he endures, his several parts are still connected by the relation of causation.<sup>21</sup>

---

<sup>20</sup>*Ibid.*

<sup>21</sup>David Hume, *A Treatise of Human Nature*, 2nd edition, revised with notes by Peter H. Nidditch, New York: Oxford University Press, 1978. Bk IV part vi

Upon seeing the whole of the quotation, we can ask ourselves if Hume is making the same point as Parfit. Certainly they are related, but it would seem that Hume is not so much saying that a person is ontologically equivalent<sup>22</sup> to a nation as he is that there is a similarity in the persistence conditions. That is, just as a nation continues to exist so long as there is continuity in its laws and traditions, *no matter how radical a change might take place*, so long as there is a causal relation, so too might a person continue to exist through radical change, so long as there is a causal relation. I assume that this means that there must be a legitimate continuity, such that Italy remained the same nation after World War II, despite no longer being a monarchy, precisely because a referendum, supported by the people decided to make the change; whereas in the case of a forced annexation like Tibet, presumably the radical changes that take place are not changes in Tibet, because of the lack of Tibetan consent. In the same way, a person may go from being a communist lesbian vegetarian to a Republican voting cattle rancher, so long as she decides so radically to transform her personality, thus preserving the causal links. The point here is that the analogy applies simply to the manner of persistence through time, and not to the status of the thing persisting.

I cannot see how this analogy to nations could account for the imagined cases that Parfit says it will help us to clarify. For instance, if the US civil war had turned out differently, and the south had won, I do not see how the fact that

---

<sup>22</sup>By 'ontologically equivalent', I mean the idea that both a nation and a person are complex entities which are not distinct from their constituent parts. In other words, the emphasis is on the causal relationships that govern the two as opposed to the ontological status of the entities so governed by causal relations.

the US was now two countries could render the identity question empty. Presumably, in such a case, the North would have continued to be the US and the south would have been something different, such as the Confederate States of America. That is, far from allowing us to circumvent the problem of identity, the Humean analogy presupposes it. We could not tell that a nation had persisted through change unless it obeyed to the persistence criteria. While Hume argues by analogy that the persistence criteria are simply causal links, that does not imply non-identity, it simply argues for a particular understanding of identity. The most that we could say from the analogy is that it remains silent on what the nature of the thing which persists through change is, but not, that the question of its persistence through change is empty. If Hume wanted to demonstrate through his analogy that personal identity is indeterminate in the event of great change, then he would not have said that a person does not lose his identity through change.

### **PROBLEMS WITH THE ANALOGY**

I have tried to show how the analogy with nations might be appropriate when considering persistence conditions for people, and I have argued that even if it were, this would still provide us with (at least the possibility of) definite, non-arbitrary, criteria of identity, even for so-called problem cases. There is also a strong possibility that the analogy does not work. I think that it is non-problematic to assert that much of what constitutes a nation would exist without that nation existing. Broadly speaking a nation involves a territory, and people. Later on, or with a more precise definition, there will of course be things like

judicial, political and financial systems which pertain to the nation as well. However, in order that there be this complex entity called a nation, there must be these constitutive parts.<sup>23</sup> That is, a nation cannot exist without people and territories, but people and territories can exist without constituting (or forming part of) nations. Moreover, it makes no difference to the nature of a person or a territory whether they pertain to a nation or not. I am perfectly aware that from a legal perspective there are important consequences to the question of nationality or statehood; however, I think that we can leave these legal questions aside. While it makes a big difference to international travel what nation one belongs to, as, for instance an Israeli will have great difficulty traveling to Saudi Arabia, while a Syrian will not, I think it a safe assertion that this fact makes no difference to the ontological status of the persons concerned. Again, in terms of international law, the nationality of territory makes a big difference. I cannot drill an oil well in Kuwait and start exporting that oil myself, this would be theft. If, however, I were to discover an island in the South Pacific and find that it was full of diamonds, I would not be stealing if I were to start marketing them. My point is that people and territories are not ontologically enriched by being, in addition to persons or territories, part of a nation. I wish to assert, however, that the constituent parts of a person are ontologically enriched by being part of a person. That is, a person is ontologically fuller than the sum total of his parts. For

---

<sup>23</sup>Here I am following Parfit [1987] "A nation's existence just involves the existence of its citizens, living together in certain ways, on its territory." According to Parfit, this claim is compatible with the following claim: "A nation is an entity that is distinct from its citizens and its territory." p. 212



instance, the body of a man is richer than the body of a corpse from an ontological point of view. This may seem like a strange way of speaking; however, I simply wish to contrast the status of people or territories which are not nations with the status of bodies or minds which are not people. That is, it does not pertain to the nature of a territory or a person to belong to a nation (even though this is in practice always the case); however, it does belong to the nature of a person always to be a mind (or soul) and a body.<sup>24</sup> I must insist that I am not trying to argue that persons are separately existing entities, I think that my remarks are perfectly compatible with Parfit's point (5) [quoted above]; however, the fact that the constituent parts of a person are ontologically diminished when they do not form part of a person and the fact that the constituents of a nation are not diminished in a corresponding way suggests that the Humean analogy between persons and nations fails in one important respect.

I began this chapter with the assertion that it is necessary to examine the context in which Derek Parfit is writing in order to make sense of his methodology. I then looked at some of the errors which various commentators are apt to make as a result of neglecting this. From there I set out to examine the purpose of Parfit's arguments, as well as critically evaluating them in light of his aims. I highlighted the major weakness in his argument, while maintaining that it does not in any way affect his critique of his contemporaries. I then looked

---

<sup>24</sup>I am here accepting Parfit's characterization of a person on page 211 of Parfit [1987] in order to illustrate the breakdown of the analogy as far as nations and persons is concerned, given his characterization of both.

briefly at two other analogies which Parfit uses to try to explain Reductionism. These two examples will prove central in later chapters as he often returns to them. Having finished my examination of the methodology of Parfit, and having put his thought into context, it is to his specific arguments that I now turn, although, as will become clear, it is always necessary to return to the background out of which Parfit's thought springs in order properly to evaluate it.

## CHAPTER II

### WHAT IS PARFIT'S REDUCTIONISM?

In the last chapter, I made three points concerning Derek Parfit's work on personal identity:

1. Parfit relies on *reductio* arguments to make a case against non-Reductionism
2. A careful examination of the analogies that he uses in the setting up of his argument reveal that he commits the fallacy of extension of terms, by concluding from the *reductio* against non-Reductionism to the validity of his particular version of Reductionism
3. As a result of 2, he weakens the strong case presented in 1.

There are many aspects of Reductionism. In a recent manuscript, Parfit outlines 7 theses:

- (A) Even if we are not aware of this, many of us are inclined to believe that, in all conceivable cases, our identity must be determinate.
- (B) For this assumption to be true, our existence would have to involve the existence of some ultimate and simple substance, such as a Cartesian ego.
- (C) There are no such entities.
- (D) Our existence consists in the existence of a body, and the occurrence of various interrelated mental processes and events. Our identity over time consists in physical and/or psychological continuity.
- (E) We can imagine cases in which questions about our identity would be indeterminate: having no answers. These questions would be in this sense empty: they would not be about different possibilities, but only about different descriptions of the same course of events. Even without answering such questions, we would know what would happen.
- (F) Reality could be fully described in impersonal terms: without the claim that people exist.

(G) Personal Identity does not have, as is widely assumed, rational or moral importance. But some of this importance can be had by psychological continuity and connectedness, with any cause.<sup>1</sup>

The view expressed by (D) is what Parfit calls Reductionism; nevertheless, he tells us that the most important claims are (A) to (C) and (G). This is because even if we accept (D), "many of us don't fully accept the implications of this view. We think about ourselves and our futures in ways that would be justified only if something like a Cartesian view were true."<sup>2</sup>

I argued in the last chapter that Parfit tends to conclude from his anti-non-Reductionism, [the view expressed in (C)], to the validity of his version of Reductionism. That is, looking at his 7 theses, I argued that from (C) Parfit concluded not simply (D) but also (E) through (G). I referred to this as the fallacy of the extension of terms. I claimed that, even if (C) and (D) are true, barring further arguments (E) - (G) could not be claimed as true, and that Parfit already spoke as if they were. As evidence of this, I appealed to Parfit's use of two analogies: that of artificial lenses, and that of comparing people with nations. I argued that in order for the first analogy to be true, we had to already accept that personal identity was unimportant; in this way, I concluded that the argument was circular. In the second case, I argued that in Parfit's intended use, the analogy was false. I claimed, contra Parfit, that persons are not like nations. Perhaps my criticisms were not entirely justified; perhaps according to Parfit, all of (D) - (G)

---

<sup>1</sup>Derek Parfit, *Experiences, Subjects, and Conceptual Schemes*, rough draft of a manuscript, September 4, 1998. pp. 1-2

<sup>2</sup>*Ibid.*, pp. 2-3

follow from (C). If that is true, then indeed, as Parfit claims, most of us do not fully accept the implications of (D). In that case, (A) - (C) and (G) would be the most important of his claims.

### PARFIT'S ALTERNATIVE TO PERSONAL IDENTITY

In this chapter, I wish to look at Parfit's (F) and (G). I shall examine the alternative that Parfit proposes to personal identity. There are two things which need to be examined, the first is Parfit's chosen psychological criterion, and the second is Parfit's argument that reality can be understood in impersonal terms. To do the first, we must briefly examine the history of psychological criteria, in order to understand the problems involved in holding such positions. We can then examine the particular criterion Parfit proposes, to see whether he is able to meet the traditional objections. After this, we can examine the second aspect, which is that reality can be understood in impersonal terms.

Parfit has a preference for psychological criteria, in fact, he proposes for us as his version of the psychological criterion *Relation R*. This is defined as "psychological connectedness and/or continuity with the right kind of cause."<sup>3</sup> As it is more controversial, he adds as a separate claim, that the right kind of cause "could be any cause."<sup>4</sup> Recall both Parfit's definition of connectedness and continuity and his motivations for appealing to them. Parfit defines psychological connectedness as follows: "the holding of particular direct psychological connections." Psychological continuity on the other hand is defined as "the

---

<sup>3</sup>Parfit, [1987] 215

<sup>4</sup>*Ibid.*

holding of overlapping chains of *strong* connectedness."<sup>5</sup> Parfit gives us one further explanation:

Since connectedness is a matter of degree, we cannot plausibly define precisely what counts as enough. But we can claim that there is enough connectedness if the number of direct connections, over any day is *at least half* the number that hold, over every day, in the lives of nearly every actual person. Where there are enough direct connections, there is what I call *strong* connectedness.<sup>6</sup>

### LOCKE'S NOTION OF PERSONAL IDENTITY

Parfit uses these notions as a corrective on the simple Lockean criterion of identity. In a justly famous passage of the *Essay*, John Locke defines personal identity in the following way:

a thinking intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing in different times and places...as far as this consciousness can be extended backwards to any past Action or Thought, so far reaches the Identity of that *Person*.<sup>7</sup>

### REID'S OBJECTION TO LOCKE

A contemporary of Locke, Thomas Reid pointed out the difficulty of identity thus defined, with his example of the general and the school boy. The critique runs as follows: suppose a boy to have been flogged at school, and later in life to have taken the standard from the enemy in his first campaign; suppose that the officer remembers being flogged as a boy. Further suppose that this same officer is, later in life, named a general. The general remembers taking the

---

<sup>5</sup>*Ibid.*, 206

<sup>6</sup>*Ibid.*

<sup>7</sup>John Locke, *An Essay concerning Human Understanding*, Peter H. Nidditch, ed., Oxford: Clarendon Press, 1975, reprinted with corrections 1979. Book II Chap. XXVII, § 9.

standard, and the officer who took the standard remembers having been flogged, but the general does not remember having been flogged. As far as Locke's definition goes, the boy and the officer are identical, and the general and the officer are identical, but the general and the boy are not identical.<sup>8</sup> While Parfit does not say that he offers his revised Lockean account as a result of Reid's criticism, it is easy to see how it avoids the criticism, while remaining true to Locke's intentions. If, as Reid says, the general cannot remember the boy, this is not problematic for Parfit's account, since there are enough chains of strong connectedness throughout his life, to enable him to be identical, even in the absence of direct memories.

#### **BUTLER'S OBJECTION TO LOCKE**

Although this revised version of psychological criterion for identity is able to avoid one of the problems that Locke's project of explaining identity fell prey to, there is another and more fundamental objection which it does not meet. Joseph Butler argued that any criterion of identity which relied on memory (as evinced by consciousness of past actions) in order to establish sameness of person was viciously circular, for surely, Butler argued,

one should really think it self-evident, that consciousness of personal identity presupposes, and therefore cannot constitute, personal identity, any more than knowledge, in any other case, can constitute truth, which it presupposes.<sup>9</sup>

---

<sup>8</sup>Thomas Reid, "Of Mr. Locke's Account of Our Personal Identity", in *Personal Identity*, John Perry, ed., Los Angeles: University of California Press, 1975. p. 114

<sup>9</sup>Joseph Butler, "Of Personal Identity", in Perry [1975] p. 100

### PARFIT'S USE OF SYDNEY SHOEMAKER'S RESPONSE TO BUTLER

With little reflection, we can see that, if Parfit's Relation R avoids the first objection, it does not, of itself avoid the second. Unless Parfit can give us independent grounds for a psychological criterion of identity being able to explain identity without presupposing it, he has not given us a valid criterion. Following Sydney Shoemaker, Parfit appeals to the concept of quasi-memory, or Q-memory, in an effort to provide just such an explanation. I said earlier that this chapter was going to examine Parfit's defense of Reductionism. As I argued, we can accept (C), so (G) proves to be the crucial premise. (A) is merely an observation of the status of people's beliefs, which I argued in the last chapter was accurate. (B) on the other hand I consider to be wrong; as with the second element of (G), however, I shall discuss this in the next chapter. This postponement is unproblematic, given that I do not think that (B) is required of Reductionists. This is because, as I see it, (C) and (D) are the essential points of Reductionism. As I see it, there are two dimensions to (G); the first argues that some psychological criterion could have the importance that we mistakenly apply to personal identity; the second is that personal identity cannot have the importance that we assign to it. I shall leave my evaluation of this latter point until chapter III; for the time being, let us examine the former: that some psychological criterion can have the rational and moral significance that we give to personal identity.

Modern theories of personal identity originate with John Locke, especially with the passage cited earlier. As we have seen, even Locke's contemporaries saw



problems with his solution to the question of personal identity. The various objections to Locke's argument have taken two forms: those, like Reid's, which seem to accept in principle Locke's approach, but which fault certain aspects of it; and those, like Butler's, which seem to argue that Locke's criterion is fatally flawed. In other words, Parfit must, if he is going to argue for some form of Lockean identity, respond to both kinds of arguments. We have seen how Relation R meets Reid's objections; it remains to be seen whether he can respond to Butler's. With this in mind, we can proceed first to examine Shoemaker's argument for Q-memory, and second, starting from Parfit's 1971 paper Personal Identity, through *Reasons and Persons*, and into his later, unpublished manuscripts, to examine Parfit's attempts to work out the full implications of Shoemaker's ideas.

### SHOEMAKER ON Q-MEMORY

The idea of quasi-memory was first expressed by Sydney Shoemaker, in 1970 in a paper entitled "Persons and Their Pasts", in which he asked

whether there could be a kind of knowledge of past events such that someone's having this sort of knowledge of an event does involve there being a correspondence between his present cognitive state and a past cognitive and sensory state that was of the event, but such that this correspondence, although just like that which exists in memory, does not necessarily involve that past state's having been a state of the very same person who has that knowledge. Let us speak of such knowledge, supposing for the moment that it is possible, as "quasi-memory knowledge," and let us say that a person who has this sort of knowledge of a past event "quasi-remembers" that past event.<sup>10</sup>

---

<sup>10</sup>Sydney Shoemaker, "Persons and Their Pasts", in *Identity, Cause, and Mind Philosophical essays*, Cambridge: Cambridge University Press, 1984. All further references are to this text, and will be cited as Shoemaker [1984]. pp. 23-4

Shoemaker is concerned in the above passage with something that is truly a memory, and not a mere illusion; that is, he is concerned with an event or action which actually occurred at some past time, as opposed to some illusory or hallucinogenic perception. What he is concerned with, then, is whether something which is a memory could be had by someone without it being a memory of something which happened to that same person.

### **WHY SHOEMAKER CLAIMS THAT MEMORY IS A CIRCULAR IDENTITY CRITERION**

Given that this enquiry is pursued in order to determine whether using memory as a criterion for identity is circular or not, we should examine what it is about memory which implies that a remembered action or event is a remembered action or event of the same person who remembers. According to Shoemaker, there are two dimensions of memory, the combination of which guarantee the identity of the person. He refers to these as the previous awareness condition (PAC) and the immunity from error condition (IEC). The previous awareness condition is defined as follows:

it is a necessary condition of its being true that a person remembers a given past event that he, that same person, should have observed or experienced the event, or known of it in some other direct way, at the time of its awareness.<sup>11</sup>

Shoemaker illustrates the second aspect of memory by considering a statement of the type "I shouted that Johnson should be impeached" claiming that such a statement would be immune to "error through misidentification."<sup>12</sup> Shoemaker concedes that "one could misremember such incidents, but it could not be the case

---

<sup>11</sup>*Ibid.* p 19

<sup>12</sup>*Ibid.* p 20

that I have a full and accurate memory of the past incident but am mistaken in thinking that the person I remember shouting was myself."<sup>13</sup> The general point is that such memory judgments are "immune to error through misidentification with respect to the first-person pronouns, or other "self-referring" expressions contained in them."<sup>14</sup> The combination of these two notions make Locke's use of memory as a criterion of identity circular. If it is both true that if I remember something, I was previously aware of it (PAC), and that if I ascribe a past action to myself, I cannot be in error about that self-ascription (IEC), then memory contains the notion of identity, and so as a criterion of identity feeds upon itself. If we cannot find some way around these two conditions, then memory based criteria are circular.

#### **CAN Q-MEMORY PROVIDE US WITH A NON-CIRCULAR IDENTITY CRITERION?**

According to Shoemaker, however, Q-memory does avoid charges of circularity. This fact requires an explanation, given that he concedes that, "In our world all quasi-remembering is remembering." What we must consider then "is whether the world could be such that most quasi-remembering is not remembering."<sup>15</sup> The first question we should ask is how the two could be separated, given that in our own world they are always identical. Shoemaker does

---

<sup>13</sup>*Ibid.*

<sup>14</sup>*Ibid.*

<sup>15</sup>*Ibid.*, p 24

this by stipulating that "quasi-remembering, as I shall use the term, includes remembering as a special case."<sup>16</sup> This stipulation is significant, as it implies that Q-memory should be our primary and more general reference. All remembering is Q-remembering, but not all Q-remembering is remembering. Even with this conceptual distinction, we would still need to imagine some way in which a Q-memory could *in fact* fail to be a memory; otherwise, the conceptual distinction is of little use in discussing actual cases. Shoemaker has to show that it would be possible to Q-remember without remembering in order for Q-memory to be a non-circular identity criterion.

#### **THE DIFFERENCE BETWEEN Q-MEMORY AND MEMORY**

The main difference between memory and Q-memory is that Q-memory "is subject to a weaker previous awareness condition than the latter. Whereas someone's claim to remember a past event implies that he himself was aware of the event at the time of its occurrence, the claim to quasi-remember a past event implies only that someone or other was aware of it."<sup>17</sup>

#### **HOW Q-MEMORY WEAKENS PAC**

The above would seem to weaken the PAC for it no longer implies that the Q-rememberer is identical with the person involved in the thing remembered, but only that someone was involved in the thing remembered. However, this weakening of the PAC alone is not sufficient to avoid charges of circularity; we need to see a corresponding weakening of the IEC. To do this, Shoemaker first

---

<sup>16</sup>*Ibid.*

<sup>17</sup>*Ibid.*

explains to us the distinction between Q-memories from the outside and from the inside. A Q-memory from the outside would be one "whose corresponding past cognitive and sensory state belonged to someone who was watching someone else do the action" whereas a Q-memory from the inside would be one in which the "corresponding past cognitive and sensory state belonged to the very person who did the action." That is, these correspond to the difference between "remembering an action of someone else's...and, on the other hand, remembering *doing* an action, which can be equated with remembering *oneself* doing the action."<sup>18</sup> According to IEC, all memories "from the inside" are necessarily immune to error.

#### HOW Q-MEMORY CLAIMS TO ELIMINATES IEC

The question then is whether Q-memories are also subject to IEC. As a first approximation, we are told that

a world in which there is quasi-remembering that is not remembering will be one in which it is not true that any action one quasi-remembers from the inside is thereby an action he himself did.<sup>19</sup>

Furthermore, IEC "exists only because remembering requires the satisfaction of the previous awareness condition." As soon as we eliminate the PAC, "this feature disappears."<sup>20</sup> While it would seem that we have just concluded that Q-memory avoids circularity as, according to Shoemaker, by weakening PAC we eliminate IEC, and by eliminating IEC we render intelligible questions about the

---

<sup>18</sup> *Ibid.* p 27 (all three quotations)

<sup>19</sup> *Ibid.*

<sup>20</sup> *Ibid.*

identity of subjects of memory, thus enabling memory to be a plausible identity criterion.

### A MODIFICATION OF Q-MEMORY

Shoemaker goes on to discuss examples which suggest that PAC is not quite what we thought it was.<sup>21</sup> In order to avoid some of the problems which follow from the PAC for Q-memory as our definition of it now stands, Shoemaker argues that we must strengthen our definition by including

the requirement that a veridical quasi-memory must not only correspond to, but must also stand in an appropriate causal relationship to, a past cognitive and sensory state of someone or other.<sup>22</sup>

The new condition is still one of Q-memory, for we are not yet claiming that the causal relationship exists between the same person at two different times, only that it is a relation between persons (maybe the same, maybe not) at different times. However, to distinguish this new type of Q-memory, Shoemaker suggests that we now speak of quasi<sub>c</sub>-memory.

### M-CONNECTIVITY

In addition, we should use the term

"M-type causal chain" to refer to the sort of causal chain that must link a quasi<sub>c</sub>-memory with a corresponding past cognitive and sensory state if they are to be "of" the same event, or if the former is to be "of" the latter.<sup>23</sup>

---

<sup>21</sup>As this is not a thesis about Shoemaker, but about Parfit, I need not concern myself with all of Shoemaker's argument. Suffice it to say that as a result of these problem examples, Shoemaker concludes that we must introduce a causal element to discussions of memory. This necessity forces us to rework our understanding of Q-memory in a way which requires further explanation to be non-circular when used a criteria of memory.

<sup>22</sup>*Supra.* 34

<sup>23</sup>*Ibid.* p 35

Moreover,

Since quasi<sub>c</sub>-remembering is to be as much like remembering as is compatible with the failure of the strong previous awareness condition, M-type causal chains should resemble as much as possible the causal chains that are responsible for actual remembering, i.e., should resemble them as much as is compatible with their sometimes linking mental states belonging to different persons.<sup>24</sup>

That is, we must ensure that these Q<sub>c</sub>-memories are as similar to memories as possible without being subject to the strong PAC. They should be as similar as possible without it always being the case that the person having the memory is identical to the person who had the experience, whether from the "outside" or the "inside". Let us next consider that

At any given time a person can be said to have a total mental state which includes his memories or quasi<sub>c</sub>-memories and whatever other mental states the person has at that time.<sup>25</sup>

In order to understand M-connectivity, we should consider two mental states existing at different times, they are

directly M connected if the later of them contains a quasi<sub>c</sub>-memory which is linked by an M-type causal chain to a corresponding cognitive and sensory state contained in the earlier.<sup>26</sup>

From this it follows

by way of giving a recursive definition, that two total mental states are M-connected if either (1) they are directly M-connected, or (2) there is some third total mental state to which each of them is M-connected.<sup>27</sup>

---

<sup>24</sup>*Ibid.* pp. 35-6

<sup>25</sup>*Ibid.* 36

<sup>26</sup>*Ibid.*

<sup>27</sup>*Ibid.*

### A RE-EXAMINATION OF PERSONAL IDENTITY IN LIGHT OF M-CONNECTIVITY

We are now faced with an alternative: either the world is such that there will always be at most one total mental state which is M-connected to another, or it is not. Either there is still going to be a one-to-one M-connection or there is not going to be a one-to-one M-connection. If there is only the one-to-one connection, and Shoemaker allows that "this is presumably the situation which exists in the actual world,"<sup>28</sup> then, we can say that these M-connected total mental states will be "copersonal," that is, states of one and the same person. This does not mean, however, that where branching occurs there will not be co-personality. Even in a world where branching occurs, the presumption of copersonality would hold "as long as there was no evidence that the M-type causal chain linking the past action or experience with the subsequent quasi<sub>c</sub>-memory had branched during the interval between them."<sup>29</sup>

### A CHANGE IN OUR NOTION OF IDENTITY RESULTING FROM M-CONNECTIVITY

Shoemaker then examines a series of counterfactual cases in which branching does occur, as, for instance, a world where people split like amoebae,<sup>30</sup> and he concludes that even though such branching does not occur in our world, this need not prevent us from saying that the offshoots both remember the actions of the pre-fission person. He further notes that in his discussion "it was assumed

---

<sup>28</sup> *Ibid.* p37

<sup>29</sup> *Ibid.* pp. 37-8

<sup>30</sup> This discussion occurs on page 38 of the text.



that remembering...involves the satisfaction of the strong previous awareness condition."<sup>31</sup> If this is the case, that even with the strong PAC we could still claim that the post fission pair remembers the activities of the pre-fission person, then

quasi<sub>c</sub>-remembering turns out to be just remembering, and the previous awareness condition for remembering turns out to be the causal requirement [M-connectivity] rather than the stronger condition I have been assuming it to be."<sup>32</sup>

That is, if, given certain counterfactual assumptions, such as the notion that we could undergo fission, we could still describe the memories possessed by the two post-fission persons as being memories of the pre-fission person, then, it would seem that our original notion of PAC is flawed, and that what we in fact have in mind is M-connectivity. Finally, if this is the case, then "the logical connection between remembering and personal identity is looser than [we] have been supposing it to be."<sup>33</sup> We can now say, even though these counterfactual assumptions turn out not to be the case, that we can avoid the objections raised about memory as a criterion for identity because our PAC would in fact be that of quasi<sub>c</sub>-remembering. Shoemaker even suggests that we speak of "'remember" in a "weak" sense," which he calls "remember<sub>w</sub>."<sup>34</sup> The reason we are unaccustomed to speaking in this way is that, as Shoemaker remarked earlier, there is no branching of M-type causal chains in our world, and so, we have had no need of distinguishing between remembering<sub>s</sub>, which is the strong sense of remembering,

---

<sup>31</sup>*Ibid.*, p 40

<sup>32</sup>*Ibid.*, p 41

<sup>33</sup>*Ibid.*

<sup>34</sup>*Ibid.*, p 42

and which obeys to the original PAC, and remembering<sub>w</sub>, which is identical to quasi<sub>c</sub>-remembering, and which does not imply identity.

### SHOEMAKER'S ARGUMENT SCHEMATIZED

The foregoing description of Shoemaker's position has been long and detailed, and of necessity included many quotations from his work, in order to establish his position. I shall now attempt to summarize this position in a way which will facilitate discussion. As I understand Shoemaker's argument, it is made up of the following 15 points:

1. We need to ask ourselves whether memory implies identity. If it does, then all memory based identity criteria are circular. If it does not, then they are not.
2. In order to imply identity, memory must satisfy both the PAC and IEC.
3. There is a kind of memory which does not imply identity, which we call Q-memory.
4. All memory is Q-memory, but of a special kind; namely Q-memory + PAC & IEC
5. All memory (whether Q- or not) must obey certain causal connections, let us call these M-type causal chains. Let us call this type of memory Quasi<sub>c</sub>-memory.
6. Either Q<sub>c</sub>-memory is identical to memory, or it is not.
7. In order for Q<sub>c</sub>-memory not to be identical to memory, there must be a possibility of branching.
8. Either there is a possibility of branching, or there is not.
9. If there is no possibility of branching, then Q<sub>c</sub>-memory = memory
10. In our world, there is no branching.
11. Even though there is no branching, we can still ask ourselves questions concerning memory which would only require answers if branching did occur.
12. We find our discussion of these questions, and our answers to them intuitively plausible.
13. Our discussion is intuitively plausible, and it indicates that Q<sub>c</sub>-memory is a useful concept even if, in our world, Q<sub>c</sub>-memory = memory,
14. Therefore, we can then claim that Memory<sub>s</sub> (which obeys to the strong PAC) = Memory<sub>w</sub> (which obeys to the weak PAC of Q<sub>c</sub>-memory)
15. Therefore, memory does not imply identity, even though the counterfactuals used to arrive at this conclusion can be admitted not to occur.

## WHAT OUR ANALYSIS DEMONSTRATES

It now seems as though the argument about the non-circularity of memory based identity criteria lacks the force which many partisans have attributed to it. When we schematize Shoemaker's argument, we see more clearly (even though he does not deny this himself) that  $Q_c$ -memory is a construct. That is, even though Shoemaker is careful to point out that his counterfactuals are precisely counterfactuals, and that they do not occur in our world, he makes the move from what might be the case, if these counterfactuals did occur to what is the case even though they do not. It would seem that the crucial steps in the argument are 11-15. While Shoemaker is willing to conclude that we can move from hypothetical cases to actual cases, precisely because he claims that the hypothetical cases are argued on the basis of our actual definition of memory (which is what happens from 11-15) I do not find the argument (as it stands, without any further reasons) persuasive. Rather than abandon the Lockean notion because we need to envisage hypothetical constructs in order to salvage it, Shoemaker wants us to reconsider our common sense notion of memory, in order to salvage the Lockean concept.

When we examine Shoemaker's argument in a schematized manner it becomes clear that at the crucial steps in the argument he asks us to change our beliefs about actually occurring states of affairs in light of states of affairs which are not merely hypothetical, but quite probably impossible as well. That is, although Shoemaker concedes numerous times throughout his article that he is making a move from a distant possible world to our actual world, he finds the

tremendous gaps from the imagined world to our world to be unproblematic. An awareness of this move should give us reason to inquire first what motivation we can have for allowing such a move and second whether such a move is justified. The reason for posing the question in this two-fold manner is that it might be the case that we have no choice but to make such moves in order to solve our problem. If this is the case then for the sake of understanding ourselves as we actually are, we might be forced to have recourse to such distant possible worlds.

Given that Shoemaker undertakes his discussion in order to respond to objections raised against Lockean identity criteria, we would do well to examine the alternatives. We can ask ourselves both why we think that Locke's notion is worth salvaging and whether the need is sufficient to justify such uses of distant possible worlds.

### **3 ARGUMENTS AGAINST PHYSICAL CRITERIA OF IDENTITY**

In response to the first question, it would seem that the arguments against physical criteria of identity are very persuasive indeed. As I see it, there are effectively three strong arguments against bodily criteria, two of which are experiential and one of which is theoretical. The two experiential criteria can be referred to as the argument from death, and the argument from radical personality change. The argument from death basically asserts that if bodily criteria satisfy identity requirements, then there seems to be an implausible diminishing of the difference between a living being and a corpse. When a person dies (barring something like an explosion which destroys the body), the body is for all intents

and purposes the same body as that of the living person, except for the cessation of physiological functions. While there are undoubtedly many biochemical changes involved in death, this does not negate the fact that the body remains the same. Even though decomposition sets in, we can take measures to prevent this, such as embalming or cryopreservation. If bodily criteria suffice, then we should not speak of Lenin's body as being in Lenin's mausoleum; rather, we should speak of Lenin's being in Lenin's mausoleum. Such a change in description strikes us as unsatisfactory.<sup>35</sup>

The argument from radical personality change appeals to our observation of radical changes in personality taking place within the same body. Consider the example of Benito Mussolini, who went from being the editor of a Communist newspaper to fascist dictator of Italy. Again, the body of Mussolini was the same body before and after, but his fundamental beliefs about politics were totally different. It would seem implausible then, to use sameness of body as evidence of sameness of person. All of us know many examples of people who undergo such radical changes (perhaps we have even experienced them ourselves). I am not here trying to argue that one does not survive such a radical change, but only to assert that it is implausible to claim that our reaction to such occurrences is to say "Well, X's body is the same, so X must be the same, despite changing from belief A to belief Z." Clearly, there is a difference between personhood and personality.

---

<sup>35</sup>In correspondence Derek Parfit has pointed out that we do sometimes speak in this way. My point, however, is that even if we do make such a reference, we intend it as a reference to the corpse and not the man.

I do not wish either to confuse or to blur this distinction. I wish to claim that it is implausible to assert that personhood survives radical personality changes simply because the same body is present before and after the change. This modest claim is all that I need to assert, in order for the argument from radical personality change to be effective.

The theoretical argument has to do with so called sorites reasoning. Recall that sorites reasoning has to do with such things as heaps of sand or the famous ship of Theseus. According to this argument, if we remove one grain of sand from a heap, it is still a heap, and if we remove one more, it is still a heap and so on until there is one grain left. The point of such reasoning is that there can be no clearly defined point at which the removal of one further grain of sand results in the destruction of the heap. In the case of Theseus's ship the argument has to do with the gradual removal and replacement of planks, until eventually not a single one of the original planks remains. The question is 'is it still Theseus's ship?' According to most thinkers, it is arbitrary to try to come up with a point beyond which the ship ceases to be identical to the original. Something trivial like the removal of a plank cannot be the difference between identity and difference.

One could of course attempt to salvage the first two criteria by attempting to locate one specific part of the body as being crucial; however, as Derek Parfit has shown, this would then become a sorites question. That is, assume (for instance) that the brain is the important element of the body for questions of

identity. The problem then becomes 'how many brain cells must I loose in order to cease to be me?' The answer surely, is 'it would be arbitrary to decide.'<sup>36</sup>

The objections to bodily criteria of identity explain why we are interested in salvaging Locke's psychological criterion; however, as I suggested in the first chapter, when discussing the teletransportation example, we can further ask whether we need to choose between physical and psychological criteria. Might it not be the case that there is some alternative criterion of identity which does not require a further fact or separately existing entity, but which nevertheless accounts for the unity of a person's physical and psychological aspects? For the moment, however, if we assume that we must choose the one or the other, we can see that there is a very strong case against physical or bodily criteria. In fact, the case is so strong, that one might argue that if the cost involved in upholding a psychological identity criterion is simply that of having recourse to distant possible worlds, as opposed to being saddled with the empirically and intuitively unsatisfying physical criteria then it is a price we should be willing to pay. We may need no further justification for adopting a psychological criterion than the pragmatic one: it works and it's more satisfying than the alternative!

### **A RETURN TO SHOEMAKER'S ARGUMENT**

Perhaps my last remarks were unjustified; perhaps we need not defend a shift from one conceptual scheme to another. I have been assuming that the need

---

<sup>36</sup>Parfit [1987] 234-36. See also the discussion of the two requirements of Bernard Williams 266-70, where Parfit maintains that no physical criteria can meet the second requirement, that against trivial facts making a difference to questions of identity. "It is a trivial fact whether some future person has half my brain or slightly more than half." 270

to have recourse to distant possible worlds should give us reason to suspect that something has gone wrong in our discussion. One reason for this is that I fear it blurs the distinction between logic and metaphysics. Allow me to specify this distinction. As I understand things, logic is the branch of philosophy that deals with the relation between ideas or concepts, whereas metaphysics deals with the relation between ideas and reality. That is, logic deals with the internal consistency of ideas and patterns of thought, whereas metaphysics seeks to determine whether those ideas, concepts or patterns of thought conform to reality, what may be loosely referred to as 'the external world.' I have been assuming throughout the discussion that we cannot make positive judgments about actually occurring states of affairs based upon radically different counterfactuals.

This is perfectly in keeping with my acceptance of the validity of the teletransportation example in Chapter I. Recall that I took great pains to separate the usefulness of teletransportation as a refutation of non-Reductionism from its usefulness as a demonstration of a particular version of Reductionism. I argued that if our aim is to refute a contrary position, or to demonstrate that there are problems with it, then we are not constrained by the proximity of a possible world. If our aim is refutation, then we can do this by pointing out how our opponent is committed to the plausibility of a given possible world, even if we ourselves find it implausible. However, if our aim is to be constructive, and to say something positive about our world, as opposed to clarifying a misunderstanding, then we are not entitled to have recourse to distant possible worlds. That is, if we



are criticising a given position, then we are entitled to speak about what would be the case if that position were true; conversely, if we are trying to establish what in fact does occur, then we have much less right to generate hypothetical cases.

### **THE DISTINCTION BETWEEN LOGIC AND METAPHYSICS**

It seems to me that if we wish to claim that we have the same right to generate far-flung hypotheses about what is the case as we do when trying to point out the problems of a given argument if things were as the given argument claims then we are substituting logic for reality. In other words, we may have recourse to distant possible worlds to determine whether a given position is coherent or not but we may not, based upon the coherence thus demonstrated, make the transition to the reality of that given position. I am claiming that if we wish to uphold a distinction between concepts and reality, or between logic and metaphysics then we should also accept that the coherence of a given conceptual scheme does not guarantee its reality. An incoherent or contradictory scheme must not be true, but a coherent or non-contradictory conceptual scheme need not be true. The contrary view, which claims that once we have given an affirmative answer to the question 'is this coherent' we have answered the only question which matters I shall refer to as the view that logic is sufficient to explain reality (LER).

It may seem like I am merely splitting hairs with the above distinction, but upon further reflection, we can see that this is not the case. If there is one thing which the history of philosophy should make clear to all it's that two mutually exclusive doctrines may be perfectly internally coherent. For instance, idealism

and realism are mutually exclusive, but they are both internally coherent. If we cannot make reference to something beyond our constructed system in order to verify our system then we have blurred the distinction between logic and reality and we are forced to concede that which of our two systems we choose is simply a matter of choice, with nothing 'objective' to justify it. We would be reduced to merely pragmatic justifications for the conceptual scheme with which we describe reality.

The above clarification might still be too harsh, as perhaps we need not be constrained by the plausibility of a possible world in the construction of a conceptual scheme. Perhaps we can claim that a conceptual scheme which makes use of a distant possible world in its defense is acceptable so long as it does not require the plausibility of that world in order to describe the actual world. I have been concerned that the use of such distant possible worlds might diminish the distinction between a concept and the reality of which it is a concept. We can now ask ourselves whether in Shoemaker's argument he needs counterfactuals to describe the actual world, or whether his conception of the actual world qua actual world does not rely on such counterfactuals. That is, even in employing such counterfactuals Shoemaker might still be able to claim that there is a distinction between his conceptual scheme and the world so described.

In order to determine whether Shoemaker's conceptual scheme is merely pragmatically justified, or whether it has a theoretical justification, we can ask ourselves first 'is anything missing in Shoemaker's redescription of memory<sub>s</sub> as

memory<sub>w</sub>?' And second, 'does Shoemaker's conceptual scheme make claims about reality which are only justified in light of his counterfactuals?' The answer to the first question is, I think, no. Shoemaker's redescription is both coherent and complete. The problem seems to arise when we ask the second question. It would seem that if Shoemaker does not utilize the counterfactual hypotheses, then he cannot avoid the charge that memory based criterion of identity are circular, and so he is in need of them in order for his argument to work. However, if this is the case, then we are back to needing a justification for the use of counterfactuals which need not actually occur. I have claimed, however, that if LER is true, then we are reduced to pragmatic reasons for choosing between rival conceptual schemes. If this is true then we should agree with Parfit's 1971 claim that it "may be a logical truth that we can only remember our own experiences. But we can frame a new concept for which this is not a logical truth."<sup>37</sup>

I am now in a position to claim that either we should reject Shoemaker's arguments as being merely pragmatically justified or else we should look for another means of justifying the use of counterfactual reasoning. Shoemaker himself seems to recognize that he is walking a fine line with his argument when he points out that if his redefinition were in addition to being logical also metaphysical; i.e. to in fact obtain in reality, then we would have reason to think

(1) that some of our concepts, perhaps including the concept of a person, would necessarily undergo significant modification in their application to such worlds, and (2) that in such worlds personal

---

<sup>37</sup>Derek Parfit, "Personal Identity", in Perry [1975] p 209

identity would not *matter* to people in quite the way it does in the actual world.<sup>38</sup>

In his conclusion, Shoemaker once again focuses on the actual world by concluding that

In the actual world it is both true that (1) remembering<sub>w</sub> is always remembering<sub>s</sub>...and that (2) the primary focus of a person's "self-interested" attitudes and emotions is his own past and future history. It is surely no accident that (1) and (2) go together.<sup>39</sup>

### THE LINK BETWEEN PARFIT AND SHOEMAKER

If Shoemaker shies away from the claim that we should revise our common sense notions of the importance of personal identity based upon his argument for Q-memory, Parfit sees no need to maintain such reservations, telling us that both Shoemaker's and his discussion "suggests a bolder claim. It might be possible to think of experiences in a wholly "impersonal" way."<sup>40</sup> One of the major differences between Parfit and Shoemaker is that Parfit considers that a consequence of a successful defense of psychological criteria of identity results in a diminished importance of the concept of identity. I think Parfit is right about this. If we can claim that Shoemaker's argument is acceptable, and we can avoid the circularity objection, then it is also the case that we can describe reality using our psychological criterion without making the claim that persons exist. I intend to show that Parfit's argument for impersonality mirrors the strategy of Shoemaker and so the two stand or fall together. That is, if, as I am going to claim, Parfit

---

<sup>38</sup>Shoemaker [1984] 47

<sup>39</sup>*Ibid.* 48

<sup>40</sup>Parfit in Perry [1975] 211

uses the same technique as Shoemaker, then either what works for Shoemaker works for Parfit, or it does not work at all. This is a bold claim, but I believe that I can demonstrate that the argumentative structures are the same. It will also become clear that either Parfit succeeds in arguing for impersonality, or else we have very weak reasons for accepting Shoemaker's Q-memory.

### **INTRODUCTION TO PARFIT'S ARGUMENT FOR IMPERSONALITY**

It is not an easy task to reconstruct Parfit's argument for an impersonal scheme, because the various parts of the argument are interspersed with responses to possible objections rather than being given all at once. I believe that if we assume that the argument for impersonality is close to that for Q-memory, then we can leave aside the various other objections, and consequently the aspects of the argument designed to meet those objections. If Parfit's strategy does not work as a whole, then the fact that certain of his remarks are designed to answer objections which are not against the whole of the argument but only against parts of it is superfluous. If the argument as a whole is sound, then an analysis of Parfit's fine-tuning is beyond the scope of this present work. An advantage of my proposal is that we can agree with Parfit that his "view about persons...is closer to Shoemaker's than he believes, and so, more likely to be true."<sup>41</sup> I shall start with a discussion of Parfit's comparison of conceptual schemes, as it will help us to see

---

<sup>41</sup>Parfit [1998] 1

how his strategy is close to that of Shoemaker, as well as providing continuity with my earlier discussion of the distinction between logic and metaphysics.

### PARFIT'S COMPARISON OF CONCEPTUAL SCHEMES

According to Parfit, there are four ways to compare conceptual schemes, which he describes as follows:

1. Mere notational variants: That would be true if, for every thought that one scheme makes possible, there could be a close equivalent in the other scheme. Such conceptual schemes, or the languages in which they were expressed, would be close to being mutually translatable. As that remark implies, this relation is a matter of degree.
2. Differ only by addition: One of two conceptual schemes may, next, either include or be included in the other. This relation holds, for example between some ordinary pre-scientific scheme and an enriched version of that scheme which includes the concepts of modern science. These two schemes are not notational variants, since there are many facts, claims and questions which cannot be recognized or expressed in the pre-scientific scheme. But these schemes may not conflict.
3. Different but compatible: This relation holds in miniature, between our concept of a river and the concept of a continuous flowing of water.
4. Different but incompatible: This relation holds, for example between a Newtonian spatio-temporal scheme, and the space-time scheme of modern physics. These schemes cannot both truly apply to reality.<sup>42</sup>

### DISCUSSION OF PARFIT'S DISTINCTIONS

I mention these distinctions now, even though Parfit brings them up towards the end of his discussion, as they are a useful orientation towards his discussion. Clearly, Parfit needs to avoid 4, as this would render the impersonal conceptual scheme not merely impoverished, but wrong. Notice that 1-3 could all

---

<sup>42</sup>*Ibid.* 38-9

be acceptable ways of describing reality impersonally; however, it seems to me that 2 and 3 require a greater justification than does 1. I assume that if two schemes are mere notational variants then they are wholly compatible and so, presumably describe the same reality. This would be the case for instance with schemes that use the terms 'Morning Star', 'Evening Star' and 'Venus' presumably there is no significant difference between the three terms, such that what can be said of one can be said of the others as well. That is, there is nothing about any of the three terms which precludes a complete description of the reality referred to.

In the case of 2 and 3, however, we are accepting that there is enough difference between the two schemes that we cannot reduce the one to the other. In the case of 2, this reduction cannot occur as one of the two schemes is more enriched than the other. The question then revolves around the significance of the added or subtracted element in the two schemes. Parfit notes that in 2, it need not be the case that the two schemes conflict. We can consider in this light the difference between traditional herbal medicine and modern pharmacology. It is often the case that modern science develops drugs as a result of traditional remedies. For instance, native Americans might notice that chewing a certain bark leads to a reduction of fever, and this observation, through laboratory refinement might then lead to the isolation of a specific compound that produces a pain killer. The point is that while the modern refinement leads to more knowledge, it is not the case that there is anything wrong with the primitive scheme. In fact, one might as effectively combat fever by chewing the bark as by

taking a pill; however, the scientific understanding is greater than the primitive one and so cannot be reduced to it.

In the case of identity, we would have to determine what the difference is between a personal and an impersonal scheme and then to determine whether the difference was so significant as to render one of the two descriptions less useful, or whether, by contrast the enriched description was not significant. In this context, we should ask ourselves what we mean by 'significant difference.' In keeping with my desire to uphold the distinction between logic and metaphysics, the question would be whether there was something real in the enriched description that was missing in the impoverished one. That is, we would want to know whether there is something real behind the concept of personal identity that is not captured by a scheme that lacks the concept such that, even though the two schemes were not conflicting there is something in the personal scheme which is lacking in the impersonal one. To return to the bark versus the pill example, we can say that the scientific scheme knows more about the properties of the bark than does the primitive scheme, but this difference is not significant as both schemes agree about the reality they are describing: that the bark reduces fever. It might be the case that there are other differences between the two schemes, as for instance, the scientific knowledge of the properties of the bark might enable us to make further discoveries about the bark that the primitive scheme does not allow. We might discover that another property of the bark is to cure skin cancer, and, if that were the case, it would seem to imply that there was something missing in the



primitive scheme; however, this need not be the case. For instance, the natives might be able to cure skin cancer by successive applications of the bark to any tumours which develop. If this were the case, then it would not imply that there was anything significant missing, only that the natives had not yet discovered the additional properties of the bark in question. The primitive scheme would only prove to be metaphysically inferior if we could show that there was something about the bark which their conceptual scheme could neither discover nor explain, as opposed to there being something about the bark which their scheme had not yet discovered or explained.

The distinction between 1 and 3 is subtle, as 1 implies that anything which can be said in one scheme can be said in another, only differently. In the case of 3, it would seem that there are things which cannot be said in the one scheme which can be said in the other, but that these gaps in knowledge can be accounted for nonetheless. In this respect, we can consider the example of glaciers versus continuous movements of ice in a certain pattern. As Parfit says

If we used that other concept, we could not think of two such movements as colliding. That would be a category mistake. When dancers collide, it is the dancers and not their movements, which collide. But, though two movements of ice could not collide, they could interact and affect each other. By thinking of the interactions of these processes, we could know as much about what happens when glaciers collide.<sup>43</sup>

---

<sup>43</sup>*Ibid.* 43

Notice that we cannot claim that the two concepts are mutually translatable; however, this does not prevent us from saying that they both account for the same facts.

### PARFIT'S ARGUMENT FOR IMPERSONALITY

We are now in a position to examine Parfit's argument for impersonality. To begin with, he reasserts that an Impersonal Scheme would be no worse (INW) than a personal one. Responding to an objection of John McDowell concerning the possibility of INW, Parfit says that "it would be enough to show that we can coherently imagine thinkers who could understand the facts to which a Reductionist account appeals, even though they did not have the concept of a person, or the wider concept of a subject of experiences."<sup>44</sup> He further specifies that if there were beings with such an impersonal scheme, then, if "we met these imagined beings, we could teach them the concept of a person in the way that McDowell doubts is possible: as a construction out of impersonal elements which they already understood."<sup>45</sup> In the same way as Shoemaker tries to lessen the difference between his hypothetical beings and us, so as to render Q-memory plausible, Parfit argues that if his imagined beings

could think about their experiences without even having the concept of a person, or the wider concept of a subject of experiences, the conceptual dependence of experiences on subjects may not be ontologically significant.<sup>46</sup>

---

<sup>44</sup>*Ibid.* 4

<sup>45</sup>*Ibid.* 6

<sup>46</sup>*Ibid.* 11

Moreover, in the same way as Shoemaker accepts that the concept of Q-memory would differ from our concept of memory (at least for the beings who have this concept), Parfit admits that

since our concept of an experience is the concept of an event that involves a subject, these imagined beings may not have our concept of an experience. But they might have a variant of this concept, and one that is similar enough to count as applying to the same part of reality.<sup>47</sup>

Parfit first describes his imagined beings, and then accounts for their thoughts.

Apart from lacking the concept of a person, and whatever else that implies, my imagined beings think like us. In place of our concept of a person, they have concepts of two closely related entities: living bodies, and unified sequences of interrelated mental processes and events, such as thoughts, experiences and acts. The unity of each sequence they take to consist in various psychological connections between these events, and in their direct relations to the same body.<sup>48</sup>

Parfit next describes the difference between 'us' and 'them'

In describing how these beings think about their lives and about 'the flow of experience', we might describe them as thinking, for example, of what is involved in first seeing something, then thinking something, then feeling something. But that description may not be impersonal, since it may imply that there is some entity which first sees, then thinks, then feels. These beings might think instead of what is involved in something's being seen, followed in the same sequence by something's being thought. Or they might think of what is involved in a seeing of something, followed in the same sequence by a thinking of something.

Next, we are told that these beings would have names for themselves (as we would refer to them), or their sequences (as they would refer to themselves).

---

<sup>47</sup> *Ibid.* 12

<sup>48</sup> *Ibid.* 11-2

We can next suppose that, just as we give people names, these beings give names to particular sequences. Where we might claim, for example, that Tenzing climbed Everest, they would claim that in Tenzing--that is, in the sequence with that name--there was a climbing of Everest. This sequence does not climb Everest; nor does its associated body. Rather, this sequence includes a climbing, achieved with this body.<sup>49</sup>

These beings are even capable of a form of self-reference, and of making a form of distinction between themselves.

In place of the pronoun 'I', these beings might have a special use of 'this' which referred to the sequence in which this use of 'this' occurred. Where one of us would say, 'I saw the Great Fire', one of them would say, 'This included a seeing of the fire.' In place of 'you', they might have a corresponding use of 'that', which referred to the sequence to which it was addressed. Where we would say, 'Did you see the fire?', they would say, 'Did that include a seeing of the fire?' They might also have a special use of 'here', so that, instead of 'I am angry', they would say, 'Anger has arisen here'.<sup>50</sup>

Parfit even claims that his imagined beings are aware of their decisions and of what they do. However, "they do not think of their decisions as made by them, or of their acts as done by them."<sup>51</sup> In the same way as Shoemaker tried to account for the possibility of the interiority of Q-memories, Parfit replies to another objection of McDowell's, which claims that an impersonal scheme cannot make sense of interiority. The objection, as interpreted by Parfit, is that there are two conditions for the 'inner' character of consciousness. "We must think of experiences that are both had by a subject, and are thought of by this subject as

---

<sup>49</sup>*Ibid.* 13

<sup>50</sup>*Ibid.*

<sup>51</sup>*Ibid.*

had by it."<sup>52</sup> Clearly the imagined beings do not have the concept of a subject; however, Parfit says

While they do not think of experiences as being theirs, they could think of them as being these--these present experiences, of which, in the conscious state that includes this thinking of a thought, there is a direct awareness. And they could think of other experiences as either being, or not being, in this sequence: the one that contains this experience.<sup>53</sup>

McDowell's first condition is also met, for

Even though my imagined beings would not think of themselves as subjects, that would be what they were. And what they call 'sequences' would be continuing lives. So even if they could understand the 'interiority' of experiences in abstraction from the idea of a subject, we have not, in imagining these beings, made sense of one of these ideas without the other.<sup>54</sup>

We can now see that just as for Shoemaker, even though Q-memories (in his imagined conceptual scheme) turn out to be memories (in our world) we are still able to use the concept, so too for Parfit, our imagined beings are subjects; nevertheless, our thought experiment is useful for giving an impersonal description. That is, just as for Shoemaker the important point was not whether Q-memory is a concept describing an actually occurring phenomenon, but whether it could be an intelligible concept, so too for Parfit the important point is not whether beings who perceived themselves impersonally would actually not be subjects or persons, but whether they could make sufficient sense of themselves,

---

<sup>52</sup> *Ibid.* 14

<sup>53</sup> *Ibid.*

<sup>54</sup> *Ibid.* 15

or sufficiently understand the facts in which their existence consists without the belief that they were persons or subjects.

If my imagined conceptual scheme is coherent, and metaphysically no worse than ours, there could be beings who understood both what experiences are like, and how experiences at different times can form unified sequences, without even having the concept of a subject. Such beings would have what McDowell doubts is possible, an impersonal understanding of psychological continuity 'which might subsequently enter into the construction of a derivative notion of a persisting subject'. It is irrelevant that these beings would themselves be subjects.<sup>55</sup>

### COMPARING PARFIT AND SHOEMAKER

I am suggesting that this strategy is both grounded in, and mirrors that of Shoemaker. Recall that in the analysis of Shoemaker's 11-15, I argued that he claims that even if Q-memory should reduce to memory (in the real world), then, provided that Q-memory is an intelligible concept in and of itself, and provided that it can account for all that memory can account for, it is a useful criterion to employ. For Parfit, the argument runs as follows: If we can imagine an impersonal conceptual scheme such that it is able to account for all that a personal conceptual scheme accounts for, then, even if the beings of that conceptual scheme should turn out to be subjects (as, in an analogous way, Q-memories turn out to be memories in Shoemaker's scheme) then this fact is irrelevant. So long as nothing significant is missing, our impersonal scheme is both successful, and usefully employable in our world. Perhaps I have been too ready to link Parfit's and Shoemaker's form of argumentation. I shall now attempt a schematization of

---

<sup>55</sup>*ibid.*

Parfit's argument, which will show, in detail, the points of similarity. Of course, not every step in the process is identical, but this need not concern us, given that I need only show that 1) Parfit's reasoning is based upon Shoemaker's and 2) Parfit's impersonality is a consequence of that reasoning, which is to say that both conclusions (that of the permissibility of the redefinition of memory, and that of the unimportance of identity) follow from the same argument.

### SCHEMATIZATION OF PARFIT'S ARGUMENT

1. We need to ask ourselves whether experiences require subjects. If they do, we should ask ourselves whether those subjects need an awareness of their being subjects or not.
  2. In order for a personal description of reality to be necessary, it must be the case that all conceivable rational subjects be aware of their being subjects.
  3. We can envisage rational beings who are subjects, but who are not aware of this fact.
  4. Such beings would either be capable of understanding reality fully, or they would not.
  5. If they were capable of understanding reality fully, then their being unaware that they were subjects would not be significant.
  6. Either we could make sense of the thought processes of such beings or we could not.
  7. We can make sense of the thought processes of these beings.
  8. Since we can make sense of their thought processes, then they have a way of understanding reality which is compatible with ours.
  9. If there were something fundamental missing from their way of understanding reality, then it would not be compatible with ours.
  10. Nevertheless, they do not have the concept of subjects of experiences.
  11. Therefore we may conclude that an impersonal conceptual scheme is possible for us, unless there is a mistake in their conceptual scheme
  12. There is no mistake in their conceptual scheme
  13. Therefore, we may conclude that an impersonal conceptual scheme is possible for us.
- In order to make the link with Shoemaker more explicit, I add the following two premises:
14. At no point in this argument did we claim that these imagined beings actually exist.
  15. Therefore, these beings need not exist for our conclusion to hold.

## A RETURN TO PARFIT'S COMPARISON OF CONCEPTUAL SCHEMES

There are still a couple of points which need clarification in order to make the link between Shoemaker and Parfit complete. I shall make that link complete using Parfit's discussion of the comparisons of conceptual schemes. Such a move is permissible given that both Shoemaker and Parfit are talking about rival conceptual schemes and the role they play in the redefinition or better understanding of our own schemes. In the case of Shoemaker, the conceptual schemes to be compared are those of memory and Q-Memory. In the case of Parfit the conceptual schemes to be compared are those of personality and impersonality. We must recall that, at a certain point, Shoemaker argues that since we have really been using our criteria to discuss his imagined cases<sup>56</sup> then it is permissible to redefine our criteria for what actually occurs on the basis of the hypothetical cases considered. Such a move does not explicitly refer to Parfit's methods of conceptual comparison; however, we are certainly entitled to ask if and why such a move is permissible.

---

<sup>56</sup>In the previous sections it was assumed that remembering, as opposed to (mere) quasi<sub>c</sub>-remembering, necessarily involves the satisfaction of the strong previous awareness condition; that is, it was assumed that in any genuine case of event memory the memory must correspond to a past cognitive and sensory state of the rememberer himself...If this is a correct way of describing [cases of fission] then perhaps my second sort of quasi-remembering, i.e., quasi<sub>c</sub>-remembering, turns out to be just remembering, and the previous awareness condition for remembering turns out to be the causal requirement discussed in the previous sections [M-type causal chains] rather than the stronger condition I have been assuming it to be. If the suggestion just made about the conditions for remembering is correct, the logical connection between remembering and personal identity is looser than I have been supposing it to be.



### ON THE COMPATIBILITY OF THE VARIOUS SCHEMES

If we re-examine the four kinds of schemes of conceptual analysis expressed by Parfit, then we can see that, in the first place two conceptual schemes are either compatible or they are not. If they are not compatible, then one could replace the other only if it were correct about something that the other was in error about. Shoemaker does not claim that 'memory' is the wrong way to describe memory, so his claim is not that the two schemes are incompatible. Likewise, Parfit does not claim that we are not persons or subjects of experience, so it cannot be the case that a personal and an impersonal scheme are incompatible. Among compatible schemes, then, there are three options: they can be mere notational variants, they can differ by addition or they can be different but compatible. Let us examine these three variations with respect to Q-memory and impersonality.

### WHY SHOEMAKER'S AND PARFIT'S SCHEMES ARE NOT MERE NOTATIONAL VARIANTS OF OURS

If two conceptual schemes are mere notational variants, then, anything which can be expressed or inquired about in one can be equally expressed or inquired about in the other. This does not seem to be the case for Q-memory and memory, for, it makes sense to ask of a Q-memory whether it is a memory of the person who has it, but it does not make sense of a memory to ask that question. Clearly then, there is greater difference and variation than would be permitted were the two schemes mere notational variants. In the case of personality and

impersonality, the opposite holds, but for the same reason. That is, in a personal scheme it makes sense to ask whether a future person will be identical to a current one, but in an impersonal scheme this makes no sense. Parfit acknowledges this point when he says that the "difference [between an impersonal and a personal conceptual scheme] is particularly clear when we consider the imaginary 'problem cases' on which discussions of personal identity have so often turned. For my imagined beings, such cases would not raise similar problems."<sup>57</sup>

#### **WHETHER THE SCHEMES DIFFER BY ADDITION FROM OURS**

If two conceptual schemes differ by addition there are notions or ideas in one scheme which are lacking in the other. Depending on which of the two schemes we consider, there will be more or less contained within; if we consider the primitive scheme, there is less; if we consider the newer scheme, there will be more. I argued earlier that the crucial question here is whether what is missing is logical or real. If the additional notions in the enriched scheme are simply concepts or ideas, but do not require that there is something missing in the reality described by the primitive scheme, then there is nothing problematic about holding either scheme. It might be convenient to hold to one rather than the other; i.e. it might be easier to describe a person than to describe a sequence of events, but that would not imply that there was something missing in the primitive description. If however, a person is more than a sequence of events, then we have a scheme which is not just inferior, but impoverished. Such an omission, I argued

---

<sup>57</sup>Parfit [1998] 48

would be grounds for the rejection of the impoverished scheme. If, however, we claim that it is only at the level of concepts that there is something lacking, then we can choose either scheme. Moreover, if we find that one scheme is simpler than the other (as presumably the primitive scheme would be) and if the omission were merely conceptual; if in all important respects the scheme which lacked the particular concept were ontologically as rich and conceptually simpler, this could be grounds for favouring that simpler scheme.

When we examined Shoemaker's scheme we saw that there was something missing: namely, the notion of the strong PAC and IEC which memory has but Q-memory does not. However, we were also able to claim that this scheme was not lacking anything significant, for Q-memory was seen to be able to apply to all things which memory applied to. We said that all memory is Q-memory, but that not all Q-memory is memory. The problem here is that we argued earlier that Shoemaker had only pragmatic grounds to defend his non-circular identity criterion. I referred to this as pragmatic for two reasons: first because the argument is motivated in part by the desire to avoid physical criteria for identity and second because Shoemaker does not give us a theoretical justification for his use of distant possible worlds. Even though there are good reasons to reject physical criteria, it does not follow from this either that we should adopt psychological criteria, or that the best or only reason for adopting psychological criteria is pragmatic. If we find some other grounds for accepting psychological criteria, then this is better. Given the similarity between Parfit's argument for

impersonality and Shoemaker's argument for Q-memory, we should see if Parfit is able to give us some defense of his reasoning.

For his part, Parfit tells us that his impersonal scheme

differs by subtraction from our ordinary scheme, since it lacks the concepts of a person, subject, thinker, and agent. When one scheme differs from another by lacking certain concepts, these two schemes might be related in any of the other ways just sketched. If the missing concepts are unimportant, these schemes may be close to being mere notational variants. If the missing concepts are important, these schemes may instead be like a pre-scientific scheme and its scientifically enriched version. The scheme with fewer concepts may be coherent and compatible with the enriched scheme, but provide less knowledge of the world<sup>58</sup>.

As Parfit does not specify in which of the other two ways we should view his impersonal scheme, we must try to understand this for ourselves. This is not terribly significant; however, as we know that it is compatible with our scheme, so we have already answered the crucial question. It now seems as though the question of whether Parfit's impersonal scheme differs by addition or whether it is different but compatible depends upon the degree to which we can translate the impersonal scheme into a personal one. If there are things in the impersonal scheme which impede those who use it from ever comprehending a personal conceptual scheme, but which remain at the level of concepts, then we would say that an impersonal scheme is different but compatible. The difference comes from the impossibility of describing all that a personal scheme can describe in impersonal terms, the compatibility would come from the missing or 'untranslatable' elements being mere concepts or ideas without any reality behind

---

<sup>58</sup>*Ibid.* 39-40

them. If however, we could in principle express all that a personal scheme describes in an impersonal one, even if that is not now possible, then it would turn out that the two schemes differ by addition.

As to which of these two relations hold between the personal and the impersonal schemes we should bear two things in mind. The first is that Parfit tells us that although his imagined scheme "is in one sense impersonal, this scheme does not deny that experiences are had by subjects, and thoughts by thinkers."<sup>59</sup> Although his imagined beings do not consider themselves subjects and thinkers, they do not deny that this is what they are. We still need to see how translatable their impersonal ideas are with personal ones; however and to do this we should return to Parfit's discussion of glaciers and continuous movements of ice in a certain pattern.

We can ask if the relation between Q-memory and memory and an impersonal and a personal conceptual scheme is like that of glaciers and continuous movements of ice in a certain pattern. If this is the case, then, unlike mere notational variants, we cannot claim that the two schemes are mutually translatable; however, unlike schemes which differ by addition, we would not need to claim that our alternative schemes are missing anything (in any significant sense). That is, although our ideas would not be directly translatable because there would be a real difference,<sup>60</sup> we could still describe everything in the one

---

<sup>59</sup>*Ibid.* 45

<sup>60</sup>I am not here using real difference in the metaphysical sense; rather I mean that the differences are such that we would not be able to translate directly from one scheme into the other. For instance, the English word cat is directly translatable into the French word chat; conversely, glacier

scheme which we could describe in the other. Recall that Parfit admits that there are things which can be said of glaciers that cannot be said of continuous movements of ice in a certain pattern, but that nevertheless the same events can be fully described using either conceptual scheme. In an analogous way, we would have to see whether or not all that is described personally can be described impersonally, even though the descriptions would not be identical.

It now looks as though Parfit is as much in need of a justification of his use of distant possible worlds as is Shoemaker. As with Shoemaker, my reason for claiming this is that Parfit needs to make use of these counterfactuals in order to establish that an impersonal conceptual scheme is not lacking anything significant from a personal one. I noted that Shoemaker is unable to provide such a justification; however, perhaps Parfit is able. Earlier on I suggested that Parfit and Shoemaker stand or fall together. Given that Parfit's arguments in addition to being a defense of a psychological criterion of identity are also a defense of impersonality, if it turns out that Parfit provides us with a defense of impersonality, then it will also turn out that Shoemaker needs to accept impersonality in order for his arguments to work. That is, it may turn out that a defense of psychological criteria of identity is only successful if personal identity is not what matters! This sounds paradoxical, but in the next chapter, I shall

---

and continuous movement of ice in a pattern are not directly translatable. That is, anything which is understood by cat can be understood in the same terms by chat; however, not everything which is understood by glacier can be understood by continuous movement of ice in a pattern. Nevertheless, there would seem to be nothing about glacier that cannot be said about continuous movement of ice in a pattern; we would just have radically different ways of describing them.

examine whether or not Parfit's Reductionism provides us with the justification for both Parfit's and Shoemaker's arguments which we have not yet found. However, a central element of Parfit's Reductionism is his premise (G) with which we began this chapter, and so, paradoxical or not, we appear to have come full circle.

## CHAPTER III

### DOES PARFIT'S REDUCTIONISM JUSTIFY IMPERSONAL CONCEPTUAL SCHEMES?

#### THE STORY SO FAR

In my first chapter I argued that Derek Parfit concludes too much from the arguments that he gives concerning personal identity. I did this by examining his teletransportation thought experiment and explaining what I believe that it demonstrates. I then examined his analogy of artificial lenses and his analogy of persons with nations. I concluded that the first analogy was circular, and that the second analogy was false.

In my second chapter, I suggested that we needed to examine the plausibility of psychological criteria of identity and that in order to do this, we needed to look to the pioneering work of Sydney Shoemaker. I also claimed that an examination of Shoemaker's work is essential to an understanding of Parfit's. I then claimed that Parfit's defense of an impersonal conceptual scheme mirrored that of Shoemaker's defense of a non-circular psychological identity criterion. This discussion naturally led to an examination of Parfit's guidelines for the comparison of conceptual schemes. I argued that since the arguments of Parfit and Shoemaker are similar we could expect they were governed by the same conceptual comparisons. Next I showed that Parfit claims that his impersonal scheme is not a mere notational variant of a personal one. As I took for granted that it was not incompatible with a personal conceptual scheme (since that would require the claim that we are not persons, a claim which Parfit denies, in order for



it to be a valid conceptual scheme) I said that we were left with two options: either a personal scheme differs by addition from an impersonal one, or else it is different but compatible. I next claimed that either of these options required further justification. I said that either there was some theoretical basis for the validity of the comparison, or there was merely a pragmatic one.

I maintained that if our reasons were merely pragmatic, then we could accept Shoemaker's arguments without accepting Parfit's; however, I said that if there were a theoretical defense of the comparisons it would be better. Moreover, if there is a theoretical defense of the comparisons, we cannot accept Shoemaker's arguments without also accepting Parfit's. I ended with a promise to examine the role of Parfit's Reductionism in defending psychological criteria of identity.

### **PARFIT'S NEED FOR A JUSTIFICATION OF HIS REDUCTIONISM**

This chapter will ultimately be an attempt to examine whether Parfit's Reductionism can justify psychological criteria in the context of discussions about personal identity. I claimed in Chapter 2 that we should not be satisfied with merely pragmatic reasons for accepting Shoemaker's and Parfit's arguments, and that this left us with two choices of schemes for conceptual comparison. I now wish to assert that these two options, since they both require Parfit's Reductionism, do not differ fundamentally. Since we are claiming that we are not content with mere pragmatic justifications, then we must claim that on the first of our two options, that a personal conceptual scheme differs by addition from an impersonal scheme, that either nothing important is missing from the impersonal

one, or else something important is missing, but that it is only conceptual. If we take the first option, that nothing important is missing from our impersonal scheme, then that sounds a lot like what we have is a mere notational variant of a personal conceptual scheme. However, Parfit himself denies that this is the case. While a mere notational variant would be perfectly acceptable, that is not, according to Parfit, what his scheme is. If, on the other hand, what is missing is merely the concept of a person, then we are back to needing Parfit's Reductionism to justify the non-necessity of that concept. This is not problematic, but it does demonstrate that, without a justification for Parfit's Reductionism, we are no further ahead.

If we take the second option, that the two schemes are different but compatible schemes; i.e. if we accept that rather than there being something missing from an impersonal scheme, what we have is simply two different descriptions of the same reality, then we are again asserting Parfit's Reductionism. The difference with our first scheme is that with this first scheme, we cannot explain everything in an impersonal way, but we can explain why we do not need to. With the second scheme, we can explain everything that we need to explain impersonally, but the different way of explaining it is sufficiently different to deny that the two schemes are mere notational variants. In other words, with this second option, we have different concepts to explain the same reality, but these concepts are sufficiently different as to not be directly mutually translatable. Again, we need to defend Parfit's Reductionism to render this option viable.

Since this discussion demonstrates that in order to have more than a pragmatic defense of psychological criteria of identity, we need to defend an impersonal conceptual scheme, we have to ask ourselves what can justify the use of an impersonal conceptual scheme? One thing is clear: we cannot simply assert that the giving of an impersonal description is self-justifying, for this justification is precisely what is at stake! Unless we want to claim that an impersonal scheme is a mere notational variant of a personal scheme, then there must be some significant difference between them, and, in order to accept our new scheme, we must either explain away that difference, or account for it. As I argued at the end of my second chapter, we need Parfit's Reductionism to account for the difference, because we cannot explain the difference away, without claiming that we have either a mere notational variant of a personal conceptual scheme or an incompatible but ultimately better scheme, neither of which Parfit claims.

### PARFIT'S REDUCTIONISM

We must now ask ourselves whether Parfit's Reductionism could explain the difference. In order to do this, we must ask ourselves what exactly is Parfit's Reductionism? Parfit gives us a couple of different explanations. In *Reasons and Persons*, Parfit said that Reductionists claim

(1) that the fact of a person's identity over time just consists in the holding of certain more particular facts.

They may also claim

(2) that these facts can be described without either presupposing the identity of this person, or explicitly claiming that the experiences in this person's life are had by this person, or even explicitly claiming that this person exists. These facts can be expressed in an *impersonal* way.

(3) 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.

Some Reductionists claim

(4) A person *just is* a particular brain and body, and such a series of interrelated events.

Other Reductionists claim

(5) A person is an entity that is *distinct* from a brain and body, and such a series of events.<sup>1</sup>

In a recent unpublished manuscript, this is stated as

(D) Our existence consists in the existence of a body, and the occurrence of various interrelated mental processes and events. Our identity over time consists in physical and/or psychological continuity.<sup>2</sup>

If we return to the previous statement, we see that Parfit tells us that

On this version of the Reductionist View, a person is not merely a composite object, with these various components. A person is an entity that *has* a brain and body, and *has* particular thoughts, desires, and so on. But, though (5) is true, a person is not a *separately existing* entity. Though (5) is true, (3) is also true.<sup>3</sup>

We can ask ourselves if the view in *Reasons and Persons* is compatible with the more recent view, and the answer is yes. The reason for this is that (3) and (D) are compatible, even though (D) no longer mentions the brain, and Parfit maintains that (3) is compatible with (5). In any event, at least this much is true: both notions are Reductionist, and Parfit counsels us that "if we are Reductionists, we should not try to decide between the different criteria of personal identity."<sup>4</sup>

---

<sup>1</sup>Parfit [1987] 210-1

<sup>2</sup>Parfit [1998] 1

<sup>3</sup>*op. cit. Ibid.* 211

<sup>4</sup>Parfit [1987] 241

## REDUCTIONISM AND CARTESIANISM

We might be able to claim then that Reductionism is the means of explaining the difference between the two conceptual schemes; however, the above statements are not clear in themselves. We still need to ask what the definition of Reductionism means. It would seem (and this is confirmed by Parfit's arguments against the validity of the Cartesian position) that Reductionism means the denial of the existence of Cartesian egos. If this is what Reductionism means, then I fully support it, and I suspect that most people do also. However, if Reductionism means only this, we can still ask ourselves if the rest of Parfit's arguments follow. I deny this. If Reductionism simply means that we are not Cartesian egos, which is (C) of Parfit's 7 Reductionist theses, then I do not think that this is sufficient to justify the impersonal conceptual scheme. My reason for claiming this is that there might be some way to account for the unity of the various elements of the person such that an impersonal description would be missing that unity and so would be an incorrect description of persons. This suggestion leads to another of Parfit's seven theses, namely, (B) which claims that identity can only be determinate in any conceivable circumstances if there exists some ultimate and simple substance like a Cartesian ego. I deny that there need be such a Cartesian ego, in order for identity always to be determinate. It will therefore be good to return to the question of the meaning of Reductionism, and to see if there might not be some other meaning to it.

If Reductionism means more than the denial of Cartesian egos, or other such simple substances, then I cannot find any satisfactory explanation of this in Parfit's writings. I looked, in the first chapter at two ways in which Parfit tries to make sense of Reductionism. The first was an analogy with artificial lenses, where he argued that from the effects of a given thing we should not worry about the cause. That is, in drawing the analogy between artificial lenses, which result in sight, but without the normal cause, and replication which results in the continuity of a person's psychology, but without the normal cause, he claimed that, if we accept that what the artificial lens give us is as good as sight, then we should say that what replication gives us is as good as survival. I claimed that this reasoning was circular. This was because in the case of an artificial lens, it is clear that the same person is the benefactor of sight; however, in the case of replication, it is not the same person. While my claim was not in itself an argument against personal identity not being what matters, neither was Parfit's claim a valid argument for personal identity not being what matters. That is, we would only accept the argument if we already accepted that personal identity is not what matters. Again, this claim, in order to be valid needs a justification, and Reductionism, as so far expressed, does not give us one.

### **REDUCTIONISM AND NATIONS**

The other description given, the one which Parfit used to explain how (3) was compatible with (5) I argued was wrong. Parfit claimed that just as a nation was not a separately existing entity, distinct from, or over and above the facts

about the people, territories and actions of that nation, that the same was true of persons. I argued that this could not be true, for, it would imply that a corpse was not ontologically enriched while it was the body of a living person, and not just a corpse. Correspondingly, the distinction between being dead and being alive would be rendered insignificant (even if one were a materialist). I further argued that if one's psychology could be separated from oneself, through teletransportation or brain transplantation, that this psychology would also be ontologically diminished. This is not to be construed as a denial of the importance of psychological continuity; rather, it should be construed as a denial that one's psychological continuity in and of itself is what matters. While Parfit might claim that this is precisely his point, and that this psychological continuity is all that should matter; i.e. that what is important is that Relation R continue to hold, regardless of the way in which it holds, I maintain that we cannot simply assert this fact, without giving it a justification. While a person may very well be simply the sum total of his parts, this assertion (which seems to be a Reductionist assertion) cannot further justify that these parts, were they able to exist in a disunited form, are not, in this disunited form of existence, ontologically diminished by the fact of their disunity. While I might say that by being a Canadian citizen, I am not ontologically enriched, and that correspondingly, should I cease to be one, I am not less of a human being, I could not say that were my psychology to be separated from my body, even were it to survive this, it was not ontologically diminished by its separate existence. Likewise, I could not

claim that, were my body embalmed or cryogenically preserved, that it was not ontologically diminished by no longer being a part of me.

### **REDUCTIONISM AND SEPARATE ENTITIES**

If it turns out that Reductionism is really just the doctrine which rejects that a person is a separately existing entity, the next question we must ask ourselves is "what is meant by 'separately existing entity'?" I maintain that by separately existing entity, we cannot mean anything which is part of a person, but which can survive the death, dissolution or destruction of the composite that is that person. To hold this would be to deny the very examples which are taken to defend Reductionism. If 'separately existing entities' refers to any part of a person which survives the dissolution of the person, then either we must affirm what Reductionism denies, which is that in fission, teletransportation or brain transplantation both of the resulting persons are the original person, which is to say that we now have one person inhabiting two bodies, or else we must deny what Reductionism affirms: namely that psychological continuity can hold between some past person and some future person, without the two being identical.

If "being destroyed and Replicated is about as good as ordinary survival"<sup>5</sup> then it would seem that there is some kind of link between the person destroyed and the person replicated. We can now ask what is meant by "about as good as" in the above sentence. Presumably Parfit means by this that the link

---

<sup>5</sup>*Ibid.* 201



between the person destroyed and the person replicated, while not constituting identity, is something stronger than the bond between two persons who do not share Relation R, and that, while this bond could not be called identity, it contains all that matters.

### **REDUCTIONISM AND SURVIVAL**

I need to explain that last point. Consider the following two scenarios:

(1) Mr. X is a brilliant research chemist who is on the verge of discovering the cure for cancer, but he has just had a massive stroke, and knows that he is going to die in 48 hours. He also knows that because he was such a meticulous researcher, who scrupulously documented all of his research and because some other researcher is almost as brilliant as he is, the discovery will almost certainly still be made. Perhaps his death will set the project back a few months, but it will go ahead.

(2) Mr. Y is a 40 year old brilliant research chemist who is on the verge of discovering the cure for cancer, but he has been diagnosed with a rare degenerative disease that will kill him within six months. However, there is a new radical cure possible: in the manner of teletransportation, we can destroy and replicate him in a new body which will be free from the disease. While Mr. Y will not survive the operation, Mr. Y\* who will be the resulting person will be identical to Mr. Y in every way, except for being cured of his disease. He will now be able to live out the rest of his life, (according to current life expectancy another 30-40 years) and he will be able to make the discovery.

In scenario (1) Mr. X is only responsible for the discovery in a very tangential way: his research has made possible a discovery that, had he lived longer, would have been made by him. While he can be comforted by the thought that his life's work is not in vain, that someone else will make 'his' discovery, his only link to that discovery is his lab notes. While we can say that this is important, and is a better prospect for him than his not being able to envisage the discovery being made, we cannot say ( even if we accept that Relation R is what matters) that the connection between the eventual discovery of a cure for cancer and his research "is about as good as ordinary survival". That is, Relation R cannot hold between Mr. X's work and a later discovery made based on that work.

At first glance, in scenario (2) Mr. Y seems to be more intimately related to Mr. Y\* and his discovery of the cure for cancer than Mr. X is to the researcher who discovers the cure in (1). Even though Mr. Y cannot claim to survive as Mr. Y\*, that is, although Mr. Y is not Mr. Y\*, scenario (2) seems to involve Mr. Y much more than scenario (1) involves Mr. X. This would be true even though the result is the same in both, that the cure for cancer is made by someone who is not identical to the original researcher. If we accept that Relation R is what matters, and we further accept that any cause of Relation R's holding is acceptable, which we would have to do to make the claim that Mr. Y is R Related to Mr. Y\*, then we are really claiming that what matters in identity is also present in Relation R. This last point is crucial: unless there is some way to explain how Mr. Y and Mr. Y\* are more closely related than Mr. X and the researcher who finds the cure for

cancer, then we have no reason to suppose that Mr. Y should be any less concerned about his impending death than Mr. X. Even leaving aside the emotional dimension (which Parfit himself concedes is very tough to overcome) I fail to see how we can simply decree by fiat that Mr. Y is better off than Mr. X.

### **IDENTITY AS A MATTER OF DEGREE**

I find it puzzling that simply by asserting that Mr. Y\*'s relation to Mr. Y is "about as good as" Mr. Y's relation to himself, we have come up with a criterion that would not also hold between Mr. X and the researcher who discovers the cure. Bear in mind that if this distinction is going to hold, then we have to accept that Mr. Y's relation to himself is not significantly different than Mr. Y's relation to Mr. Y\*. If we want to say that identity is not what matters, then Mr. X and Mr. Y are in the same situation; this would be akin to saying that if we desire something that is good *simpliciter*, then we should desire it in itself, as opposed to desiring it for ourselves. It may very well be the case that both Mr. X and Mr. Y should be equally pleased at the thought of the impending discovery of cancer, regardless of whether he discovers it, but this is not an argument for the proposition that Mr. Y bears a closer relation to Mr. Y\* than Mr. X bears to the researcher who discovers it. If we hold that Mr. Y is not Mr. Y\*, then why should we not also hold that Mr. Y\* is simply someone with the same desires and capacities as Mr. Y? And if Mr. Y\* is simply someone with the same desires and capacities as Mr. Y, is it not also true that the discoverer of the cure following the death of Mr. X is also such a person? Certainly there is less similarity between

Mr. X and the other researcher than there is between Mr. X and the later discoverer of a cure for cancer, but it does not follow from this that Mr. X should be any less consoled than Mr. Y.

### **IDENTITY AND DESIRE**

If identity is really just a matter of degree, then why, if for both Mr. X and Mr. Y their greatest desire is going to be fulfilled should the one be any less consoled than the other? We could of course begin to bring up all sorts of other considerations such as Mr. Y\*'s taking care of Mr. Y's wife and children, his finishing off the painting that Mr. Y would have finished, etc.; however, if identity is not what matters, then all that should matter is that Mr. Y's desires all be fulfilled. The question of who fulfills these desires now becomes irrelevant. If however the question of who fulfills his desires becomes irrelevant, then we are conceding a lot more than the unimportance of identity. We could say that since Mr. Y\* is exactly like Mr. Y that he could best carry out Mr. Y's desires, that the trauma his family faces by his death is lessened by their having an exact replica rather than someone else looking after them, etc.; however, once again, this does not seem to follow. If his family is really consoled by this fact, then it means that they too wish for his desires to be fulfilled, as opposed to being fulfilled by him. They cannot claim that he is fulfilling them, so, for instance, they cannot appeal to anything like a shared history, i.e. Mr. Y\* would not be Mrs. Y's husband, or the father of their children, etc., so she could not reminisce about their wedding day, the first words of their son, etc. If this fact (that what they now have is someone

exactly like Mr. Y, but who is nonetheless not Mr. Y) really consoles them, then it means that what they love is not Mr. Y, but Mr. Y's characteristics. This would explain how someone exactly like Mr. Y, but who was not Mr. Y, would be more of a consolation to them than someone totally different taking care of them. Moreover, it would also seem to imply that if Mr. Y knows of someone who can better fulfill his desires than he can, that he should desire that that other person do so, instead of himself.

For instance, perhaps Mr. Y desires both that his wife and children be well cared for and that a cure for cancer be discovered. Suppose also that he cannot devote himself wholly both to the happiness of his family and to the finding of a cure for cancer. Quite apart from considerations of replication in place of death, it would seem that in order to fulfill both desires, he should decide whether his desires are better fulfilled by being a chemist or by being a father and husband, and which of his two desires could most easily be fulfilled by someone else. That is, either, his family will be happier with him as father and husband, and someone else can find the cure for cancer, so he should give up his research or else, if he sees that his family would be happier with someone else looking after them, and that no one else is better suited to find a cure for cancer, he should leave his family and devote himself wholly to research. I am going to leave aside the possibility of a dilemma, where, no matter which option he chooses some desire will suffer, for my only point is that if we take impersonality seriously, then it turns out that we are subordinated to our desires. A consequence of taking

impersonality seriously would be that wherever someone else could better fulfill a good desire that we had, we should leave that up to the other person. In other words, Mr. Y would have no justification (or at least no strong justification) to desire that he look after his wife and kids if he knew that someone else could better fulfill this desire, unless identity mattered. I mean this in the sense of mattering really, i.e. being metaphysically significant, as opposed to simply mattering from the point of view of emotional attachment. At the very least, we should say that Mr. Y should have a stronger desire that his wishes be fulfilled in the best possible way than that he be the one to fulfill them, if, by his fulfilling them himself, they are imperfectly fulfilled.

#### PARFIT ON 'ORDINARY' SURVIVAL

Parfit does, in fact, seem to endorse a view remarkably like the one just described, for he says, in reference to his teletransportation example, and why he may fear that his replica on Mars might not be him because it would be missing some "further fact" that "What I fear will be missing is *always* missing." Moreover, "*Ordinary survival is about as bad as being destroyed and Replicated.*"<sup>6</sup> Parfit also says that

My death will break the more direct relations between my present experiences and future experiences, but it will not break various other relations. This is all there is to the fact that there will be no one living who will be me. Now that I have seen this, my death seems to me less bad. Instead of saying, 'I shall be dead', I should say, 'There will be no future experiences that will be related, in certain ways, to these present experiences'.<sup>7</sup>

---

<sup>6</sup>*Ibid.* 280

<sup>7</sup>*Ibid.* 281

Parfit sees the radicality of his view as positive:

Is the truth depressing? Some may find it so. But I find it liberating, and consoling. When I believed that my existence was such a further fact, I seemed imprisoned in myself. My life seemed like a glass tunnel, through which I was moving faster every year, and at the end of which there was darkness....There is still a difference between my life and the lives of other people. But the difference is less. Other people are closer. I am less concerned about the rest of my own life and more concerned about the lives of others.<sup>8</sup>

Additionally, we are told that

As I have said, I care less about my death. This is merely the fact that, after a certain time, none of the experiences that will occur will be related, in certain ways, to my present experiences. Can this matter all that much?<sup>9</sup>

Notice, however, that Parfit has not given us a single justification for his point of view. If we review the case as presented by Parfit, we see that at every stage where we stop and ask for a justification, we either find none, or find a problematic one. That is, we cannot appeal to the possibility of a non-circular psychological criterion, for, when we ask why we should accept that hypothetical reasoning make a difference to actual cases, we find none. Neither can we appeal to the analogy with artificial lenses, for we have shown that it is circular. The comparison of persons with nations does not work, because the parts of a person, unlike the parts of a nation are not complete apart from being a person. Neither can we accept that since we can imagine hypothetical beings who have an impersonal conceptual scheme that we too could have an impersonal conceptual

---

<sup>8</sup>*Ibid.*

<sup>9</sup>*Ibid.* 282

scheme. This is because, in examining the possible justifications of comparing the conceptual schemes, we have found something lacking each time. Finally, we have seen that Reductionism cannot be the justification either, because its validity presupposes some or all of the previously offered justifications. If none of the justifications works separately, it cannot be the case that all of them work when combined.

### IMPERSONALITY REVISITED

At this point, it will be helpful to return to the comparison of the justification of Q-memory schemes and impersonal conceptual schemes, for perhaps they can extricate us from the above difficulties. In the same way as Shoemaker proposes that Q-memory should be our primary referent, and that memory should be considered a special case of Q-memory, namely Q-memory that guarantees identity, perhaps we can say that for Parfit, Relation R should be our primary referent and that identity is but a special case of Relation R, namely Relation R with its normal cause, or non-branching Relation R; that is, Relation R that holds for the same person at different times. Moreover, just as Shoemaker claims that since in our world branching does not occur, Q-memory equals memory (in our world), we might also say that since in our world branching does not occur, Relation R equals identity. In this context, we can finally make sense of Parfit's remark that "personal identity nearly always coincides with psychological continuity, and roughly coincides with psychological



connectedness."<sup>10</sup> We can understand the 'nearly always' as implying that without either branching or an abnormal cause (such as replication with destruction) some psychological criterion (whether Relation R, or some other) is what constitutes identity. This is certainly in keeping with our intuitions that personality, and other psychological traits are more important than simple physical identity or continuity; however, Parfit says that "personal identity is not what matters. It is merely true that, in most cases, personal identity coincides with what matters."<sup>11</sup>

### PARFITIAN DUALISM

To the question "What does matter *in the way in which* personal identity is mistakenly, thought to matter?"<sup>12</sup> Parfit replies that there are four options:

- (1) Physical continuity
- (2) Relation R with its normal cause
- (3) R with any reliable cause
- (4) R with any cause<sup>13</sup>

We already know that (1) is not a viable option, according to Parfit (2) is what normally occurs, however "physical continuity is the least important element in a person's continued existence."<sup>14</sup> This leaves (3) and (4). (3) is not really a viable option, because, if R is what matters, then, the cause is insignificant. Returning to the teletransportation example, Parfit says

- (a) that my replica will not be me  
would just consist in the fact
- (b) that there will be no physical continuity,  
and

---

<sup>10</sup> *Ibid.* 215

<sup>11</sup> *Ibid.* 282

<sup>12</sup> *Ibid.*

<sup>13</sup> *Ibid.* 283

<sup>14</sup> *Ibid.* 284

(c) that, because this is so, R will not have its normal cause. Since (a) would just consist in (b) and (c), I should ignore (a). My attitude should depend on the importance of facts (b) and (c). These facts are all there is to my Replica's not being me....It cannot matter much that the cause is abnormal. It is the *effect* which matters. And this effect, the holding of Relation R, is in itself the same.<sup>15</sup>

Notice however that Parfit has now radically separated our psychology from our bodies, in the way in which Descartes did. It would seem that we cannot account for the unity of our psychologies with our bodies. I am suggesting that even if we accept some kind of psychological criterion of identity, we need to find some way to link that psychology to the whole person otherwise we shall end up holding a Cartesian position. Parfit's Reductionism, in its attempt to offer a justification of the unimportance of identity, leads to dualism of a Cartesian type. While Parfit no longer places an emphasis on the importance of identity, such that he will not claim that a replica is identical to the person it is a replica of, he does say that the replica "contains all that matters." I submit that if the replica contains all that matters, then even if we do not claim identity; indeed, even if we go so far as to maintain that there is nothing of the original person in the replica, then we are still dualists of a Cartesian type. In fact, we are ultra-Cartesian, for, Descartes at least demanded that something (the ego) persist; here we have neither the body, nor the ego, but we still have all that matters. And what is it that matters: Relation R. What is Relation R? It is psychological connectedness and/or continuity with any cause. Of what we may ask is it continuity? The answer must

---

<sup>15</sup>*Ibid.* 286

be (since Parfit does not accept that there is an ego) *of itself*. This is not to say that there are experiences without subjects, for there must be a subject; however, it is to assert that the subject does not matter. That is, while whatever is required for conscious experience (be it just a brain and body, or be it something not totally physical), must be present for the continuity or connectedness to occur, it need not be the same subject for that continuity to occur. Where the subject is the same, we have identity in the normal sense (Relation R with the normal cause). Where we have different subjects of experience, we have Relation R without identity. Ultimately, then, if Parfit's arguments are to be successful; that, is, if we assume that at least one of the objections raised in this thesis can be responded to, then what we are left with looks like Descartes position, only stripped of identity. If we accept that experiences must occur to someone, that experiences do require a subject, but that the identity of that subject is unimportant, then we, like Descartes hold to a radical dualist position. I am suggesting that, in a fundamental way, the price of Parfit's success is to endorse the view he rejects. Cartesianism without identity is still Cartesianism!

### PARFIT AND DESCARTES

As things currently stand, Parfit is caught in the horns of a dilemma: if Relation R is separable from the person of whom it is Relation R, we are lead to the kind of Cartesianism which he wishes to deny. On the other hand, it would seem that if there is something which persists, then we are also lead to the kind of

Cartesianism which he denies. This claim needs an explanation, so let us now compare Parfit and Descartes. In the sixth meditation, Descartes says

my essence consists solely in the fact that I am a thinking thing. It is true that I may have (or, to anticipate, that I certainly have) a body that is very closely joined to me. But nevertheless, on the one hand I have a clear and distinct idea of myself, in so far as I am simply a thinking, non-extended thing; and on the other hand I have a distinct idea of body, in so far as this is simply an extended, non-thinking thing. And accordingly, it is certain that I am really distinct from my body, and can exist without it.<sup>16</sup>

Notice what Descartes has done: he has affirmed that there exist two distinct things: a thinking thing and an extended thing. The thinking thing is the ego, and the extended thing is the body. Descartes identifies himself with the thinking thing, for his idea of a thinking thing does not include that of an extended thing. He therefore concludes that he is distinct from the extended thing and that he can exist without it. This certainly involves talk of identity, for he does claim to be the thinking thing. Notice, however, that Parfit makes the identical point, only without talk of identity. That is, while Descartes says both that the ego matters and that he is the ego, whereas Parfit says only that Relation R is what matters, but does not make the further claim that Relation R is what he is, this fact is not significant. For Parfit, no less than Descartes, envisages a conceptual scheme in which there is no essential link between the thinking thing and the extended thing.

I do not mean that Parfit's arguments rule out materialism (for matter alone might be the locus of Relation R); however, Parfit claims that Relation R, is, at

---

<sup>16</sup>Rene Descartes, *Meditations on First Philosophy*, in *The Philosophical Writings of Descartes*, Volume II, John Cottingham, Robert Stoothoff, Dugald Murdoch, trans. Cambridge: Cambridge University Press, 1984. p 54

least conceptually, separable from whatever its locus is. Given Parfit's argument, whether R is located in matter or in something apart from matter, it is still separable from its normal locus. Otherwise, his talk of teletransportation, fusion, fission, etc., is meaningless, as it would imply that R always and everywhere had only its normal cause, and we would then be forced to concede that Shoemaker is right, and that we could establish a non-circular identity criterion, but that it would still be an identity criterion, so that identity would still be what matters. Identity would still be what matters because we could not claim that Relation R is actually separable from identity. The most we could claim is that identity reduces to Relation R. In place of Shoemaker's Q-memory, we would then be speaking of Relation R as our criterion of identity. Recall, however, that Shoemaker favours  $Q_C$ -memory and that an important component of it is "M-type causal chains". Recall further that "M-type causal chains" and Relation R are the same thing. If Parfit in fact has a different argument from Shoemaker, then a consequence of this would seem to be that he is a Cartesian, but a special kind of Cartesian, one who does not think that identity is what matters. So far we have only been able to distinguish Parfit from Shoemaker at the cost of making him Cartesian. The cost of his distinction has been to accept what he denies, for, in order to be different from Shoemaker, and to accept that identity is not what matters, he must accept the central notion of Descartes, that what matters could be separated from a particular body. The fact that this separation would not entail identity does not imply victory for Parfit, it simply implies a rejection of the substance based

metaphysics of Descartes. The cost, however, is to concede Descartes most fundamental point; namely that what matters and the unity of the person (ego and body, or Relation R and body) are separate, clear and distinct. Moreover, this is not simply true conceptually, but actually. Again, the reason that this must be metaphysically true as opposed to conceptually true is that if this separability is only a conceptual truth, then what we have is an identity criterion instead of a "what matters" (without constituting identity) criterion.

### CONCLUSION OF CHAPTER III

By way of conclusion I submit that the aforementioned point is the most serious objection to Parfit's work. Even if we understand 'separately existing entity' in the widest possible sense, such that it need not refer to something which is a substance, and even if we assume that Parfit's argument is both consistent and valid, then either identity does matter, or we are forced into a fundamental agreement with Descartes in order to deny the importance of identity. That is, even if I am wrong in claiming that Shoemaker has been unable to justify his use of distant possible worlds to construct a non-circular psychological identity criterion (and it was precisely this claim that led me to examine the role of Parfit's Reductionism in this regard) then either Parfit's arguments reduce to Shoemaker's or else Parfit is as much a dualist as Descartes.

## CONCLUSION

I began this study with the intuition that there was something missing in the current debate surrounding personal identity. The more research I did, the clearer it became that Descartes is responsible for the contours of the current debate. Given Derek Parfit's argument that there is no empirical evidence for Descartes' claims, and given his argument against quasi-Cartesian views such as Bernard Williams', I had hoped in Parfit to find the resolution to the problem. However, as I proceeded to analyse the published materials of Parfit, and as he continued to provide me with his more recent and as yet unpublished material, I became increasingly aware that he was as much tied to the essential Cartesian paradigm as are the other parties to the debate. Moreover, as I began to imagine what the world would be like if Parfit's arguments were true, I realised that it would be a world in which Descartes was at home. This is not to claim that Parfit and Descartes have identical positions, for Descartes accepted the existence of the ego, which Parfit denies. However, if we compare the person for Descartes, (considered as a Cartesian ego which has a body), and the person for Parfit (where Relation R is what matters) and if we add to this Parfit's claim that the cause of Relation R does not matter, then notwithstanding the absence of an entity like an ego, the difference is insignificant. Even as Descartes could not find an adequate explanation for the unity of the ego with the body, neither can Parfit find such an explanation. While we might be tempted to think that Parfit's person is more acceptable to our scientifically and technologically enlightened age than would be

Descartes' person (with his mysterious ego), we cannot claim that Parfit has been any more successful than Descartes in uniting the various components which make up the person.

In presenting my argument, I have had to examine the relationship between metaphysics and logic, in order to point out the dangers involved in the use of distant possible worlds as a means of better explaining our own. To be sure I am less weary of such techniques than either Wittgenstein or Quine; nevertheless, I maintain that we must have clearly in mind the difference between our concepts and the reality they describe if we are going to benefit from such reasoning without getting carried away by it. In referring to the eternal recurrence, Nietzsche once remarked that "I drew this conclusion, but now it draws me." Such, I believe, is the danger of mixing hypothetical worlds with our own actual world, if we are not prudent beforehand.

I believe that Parfit's writings are illuminating, and although I do not accept his solution, I certainly share his diagnosis. In this study I have not had the time to fully develop an Aristotelian or a Thomistic theory of personal identity which takes into account both current scientific discoveries, and the contemporary analytic philosophical scene; however, I am convinced that this is our only hope to arrive at the definitive solution to the problem. It is perhaps ironic, but not for that reason untrue, that we should look to a metaphysics which was developed long before modern science to solve problems that we could not even have been aware of before its development.



I maintain that it is precisely because philosophers, and especially metaphysicians are too willing to jump on the latest scientific research about psychology or physiology that we have been unable to get to the bottom of our human nature. Science can tell us many things about how we function, and develop but it cannot tell us what we are. If you want to understand the physiological processes which are undergone in human development you should certainly look to science; however, if you are interested in the nature of the person who undergoes these changes, and in whom these processes occur, you need a metaphysician. Science certainly gives us a more profound description of reality, but to understand and interpret it, you need a different level of analysis. Such analysis belongs to metaphysics.

I fear that metaphysics is currently too willing to accept the physical science paradigm in order to attempt to solve the deepest questions of reality. I need only point to the increasing use of counterfactuals as a means of clarifying our concepts. Kant maintained that the world of things in themselves was inaccessible to our empirical self, and so he argued that the task of metaphysics (like that of physics) was to construct hypotheses which were capable of describing reality as we know it. I do not wish to argue that metaphysics should contradict experience; however I do maintain that if we are ever going to understand reality fully then we must learn to abstract from the sensible to the intellectual. If Kant, and most philosophers since are sceptical about this possibility, I remain more optimistic. I believe that not only can we grasp the

intelligible through the sensible, I believe that even to understand the sensible we must.

The question of personal identity is particularly well suited to this larger project of understanding reality, as we cannot but understand reality as persons. A well known doctrine of Thomistic metaphysics is that “whatever is received is received according to the condition of the receiver.” If this is true, then it is imperative that we understand well exactly what kind of beings we are, in order to understand the world of which we are a part. Elizabeth Anscombe is famous for her remark that we should stop doing ethics until we have an adequate philosophical psychology. This is also my belief, and I hope that this study of the limitations as well as the usefulness of Derek Parfit’s thought serve as an opening onto this larger question.

### SELECTED BIBLIOGRAPHY

- Aristotle *On the Soul*, In *The Complete Works of Aristotle*, Jonathan Barnes, ed. Princeton NJ: Princeton University Press, Volume 1, 1995.
- Butler, Joseph "Of Personal Identity," in *Personal Identity*, John Perry, ed., Los Angeles: University of California Press, 1975. 99-105
- Descartes, Rene *Meditations on First Philosophy*, in *The Philosophical Writings of Descartes*, John Cottingham, Robert Stoothoff, Dugald Murdoch, trans. Cambridge: Cambridge University Press, 1984. Volume II
- Goodenough, Jerry "On the Methodology of Thought Experiments," found on the internet at <http://astro.ocis.temple.edu/~souder/thought/good.pap>
- \_\_\_\_\_. "Parfit and the Sorites Paradox" *Philosophical Studies*, 83 (2) 113-120
- Hume, David *A Treatise of Human Nature*, 2nd edition, revised with notes by P.H. Nidditch, New York: Oxford University Press, 1978.
- Johnston, Mark "Human Concerns without Superlative Selves," in *Reading Parfit*, Jonathan Dancy, ed., Malden: Massachusetts, Blackwell Publishers Inc. 1997. 149-179
- Locke, John *An Essay Concerning Human Understanding*, Peter H. Nidditch, ed., Oxford: Clarendon Press, 1975, reprinted with corrections 1979.
- Maddell, Geoffrey "Derek Parfit and Greta Garbo" *Analysis*, 45 105-109
- Parfit, Derek *Experiences, Subjects and Conceptual Schemes*, rough draft of a manuscript, September 4, 1998. 54 pages
- \_\_\_\_\_. "Personal Correspondence" November 28, 1998.
- \_\_\_\_\_. "Personal Identity," in *Personal Identity*, Perry, [1975] 199-223
- \_\_\_\_\_. *Reasons and Persons*, New York: Oxford University Press, 1984. Reprinted with corrections, 1987.
- Reid, Thomas "Of Mr. Locke's Account of Our Personal Identity," in Perry, [1975] 113-118
- Shoemaker, Sydney "Parfit on Identity," in Dancy [1997] 135-148

---

\_\_\_\_\_. "Persons and Their Pasts," in *Identity, Cause and Mind Philosophical Essays*, Cambridge: Cambridge University Press, 1984.  
19-48

Wilkes, Kathleen *Real People Philosophy of Mind Without Thought Experiments*, New York: Oxford University Press, 1993.