Derek Parfit and Personal Identity

Is Parfit's Relation R All That Matters?

Anne Newburg
Department of Philosophy
McGill University, Montreal
March 1991

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Arts.

ABSTRACT

This thesis examines Derek Parfit's theory of personal identity. Parfit argues that what matters in the continued existence of persons through time is psychological connectedness and continuity (relation R), and that the identity relation does not matter. He makes this claim through a series of arguments which, he says, inevitably lead to the conclusion that relation R is the only relation that matters, in all cases. I argue that Parfit does not convincingly demonstrate that relation R is in fact all that matters. In examining each of Parfit's arguments, I show that it is possible to draw conclusions that are inconsistent with those drawn by him. I argue that this shows Parfit's position to be an arbitrary one. If Parfit's arguments do not necessarily lead to the conclusion that relation R is all that matters in questions of survival, then his theory is not an adequate solution to the problem of personal identity.

RESUME DE THESE

Ma thèse examine la théorie d'identité personelle de Derek Parfit. Parfit prétend que ce qui importe dans l'existence continue des personnes dans le temps sont la connexité et la continuité psychologiques (relation R), et que la relation d'identité n'a pas d'importance. Parfit fonde cette théorie à travers une série d'arguments qui, d'après lui, mène inévitablement à la conclusion que la relation R est en fait tout ce qui est important dans tous les cas. Je soutiens que Parfit ne démontre pas de façon convaincante que la relation R est la seule critère qui importe. En examinant chacun des arguments de Parfit, je démontre qu'il est possible de tirer des conclusions qui ne sont pas consistantes avec celles tirées par Parfit. J'affirme que ces conclusions démontrent que sa position est arbitraire. Si les arguments de Parfit ne mènent pas nécessairement à la conclusion que la relation R est primordiale quant à la question de survivance, on peut constater que sa théorie n'est pas une solution adéquate au problème d'identité personnelle.

ACKNOWLEDGEMENTS

I wish to thank Jim McGilvray who advised me in the writing of this thesis. His discussions with me were thoughtful and insightful, and his comments on various drafts of the text were extremely helpful. I also wish to thank Mane Hajdin who first introduced me to Parfit's theory, and encouraged me towards this topic through many challenging discussions about personal identity. And I am greatly indebted to Ian Gold for his perceptive comments and suggestions, and for his support. His help was important to the completion of this paper.

TABLE OF CONTENTS

CHAPTER ONE

1. Introduction

In his book Reasons and Persons Derek Parfit defends the Reductionist view of personal identity. He constructs a number of arguments in support of Reductionism and, through these arguments, he claims that personal identity is a series of experiences connected by a particular relation, which he calls relation R.

In defending the Reductionist view of personal identity, Parfit is offering a theory that is very much based on the argument offered by Hume in A Treatise of Human Nature.[1] While Hume is the antecedent to Parfit's position on personal identity, the methods and arguments that Parfit uses in defending his views are sufficiently original and challenging to warrant individual treatment. For this reason I will restrict myself to the original work of Parfit and will not explicitly trace the historical context of his claims.

I will examine Parfit's theory and will show that the arguments he puts forth do not sufficiently support his view of Reductionism. I will *not* defend a specific position on

the Reductionist/non-Reductionist spectrum; rather, I will consider Parfit's arguments and, through counter-arguments, I will challenge the claim that his view is, as he asserts, the "only plausible view", or the "true view", of what matters in questions of personal identity.

The problem of personal identity is broad and may be approached from a variety of angles, each emphasizing different aspects of the problem. For instance, the question of personal identity may evoke arguments concerning the nature of persons, individual identification, individual reidentification, individual differentiation, or, alternatively, class differentiation. As Parfit's analysis of personal identity is concerned with defining what is necessarily involved in the identity of persons over time, this discussion will limit itself to this particular aspect of the problem.

2. Parfit on Personal Identity

Parfit argues that there is nothing further to the existence of persons than a series of experiences unified by way of relation R.[2] Relation R is physical and psychological connectedness and continuity, with any cause.

This means that when I consider whether or not I will survive, or whether I will have a given experience, I am considering nothing further than whether a certain kind of relation will hold between my present self and those particular, future experiences. I am not concerned with my soul, or with any other kind of separate substance that identifies me as being who I am; I am not wondering whether my ego will be harmed or affected; nor am I interested in protecting some "furcher fact" that distinguishes me from others, making my experiences peculiarly mine. I am not concerned with any of these things because my existence involves no such "further fact". All that it involves is relation R: physical and psychological connectedness and continuity, with any cause.

This, Parfit says, represents the true view of personal identity. If I believe that there is in fact something further to my identity than relation R, it is because I am misled. Once I recognize that I am nothing further than a series of physical and mental states interrelated in a certain way, then I will care less about my self, and about my experiences, both present and future. Since I will care less about my own experiences, I will, according to Parfit, care more about the experiences of others. This is because the true view of identity serves to reduce the barriers that we maintain between ourselves and

others. The effect of this will be a breakdown of my reasons in favor of self-interest.

Parfit claims that relation R accounts for all that matters in questions of identity, and that certain significant implications follow from this fact. If, however, relation R does not take care of all that it claims to take care of, then it is difficult to accept Parfit's view, at least as a sufficient explanation of what matters in questions of identity.

My view is that Parfit does not succeed in meeting his goal. To show this, I will begin by reviewing Parfit's position.

Parfit is a Reductionist. He does not believe that there is a further fact, or separate entity, that distinguishes us from one another. Our existence does not entail the existence of a soul, or of an ego, such as that discussed by Descartes, without which we would be incapable of experience. Such beliefs represent the non-Reductionist view of personal identity which is, for Parfit, the false view. The true view of identity claims that our existence just involves the existence of a brain and body, the doing of certain deeds, the thinking of certain thoughts, and the occurrence of certain experiences.[3]

Parfit's version of Reductionism claims:

1) that "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"; and

2) that "a person is an entity that is distinct from a brain and body, and such a series of events."[4]

This view takes the position that a person is an entity that has particular thoughts, does particular deeds, and so on. While a person's existence just involves a series of interrelated physical and mental events, it is true that we can refer to persons as being separate from these events. But this does not mean that "person" involves anything further than physical and mental events. As Parfit explains, this way of talking is similar to the way we refer to other objects that have two names. Venus, for example, is also called the Evening Star. It is possible to refer to Venus without using the words "Evening Star", just as we can refer to the Evening Star without referring to Venus. Both accounts are complete; we are not leaving anything out by saying "I've been to Venus" when the Evening Star is not mentioned.

In the same way, in saying:

1) There exists a particular brain and body, and a particular series of interrelated physical and mental states,

I am saying the same thing as:

2) A particular person exists.[5]

The point is, although we do speak of persons as existing, this does not mean that the existence of persons involves some deep or further fact. Rather, our concept of a person may be accounted for without ever referring to anything other than a series of interrelated physical and mental states. It follows from this that it is possible to fully describe what is involved in the existence of a person without reference to a person. This is the view that a complete description could be *impersonal*.[6] In describing a series of interrelated physical and mental states it is not necessary to claim that a person exists, for that fact is included in the very description of those interrelated states.

This view of persons is important to Parfit because he believes it tells us what matters in questions of identity.

In determining what is involved in the identity of person A through time, what is relevant is the relation that holds between that person's physical and mental states, and past physical and mental states. This is relation R: physical and psychological connectedness and continuity.

Psychological connectedness exists between a present mental state and a past mental state when there are a sufficient number of memory links between these two states. This means that I am connected with the self who read the newspaper yesterday if I remember the experience of reading a newspaper yesterday.

Our concept of memory is that the memories we have refer only to our own experiences. On this view, the continuity of memory and relation R would both presuppose personal identity. To address this objection, Parfit offers an expanded definition of memory, which he calls "quasi-memory".

On this defintion, my memory of reading the newspaper yesterday would be accurate if and only if:

1) I seem to remember having the experience,
2) *someone* did have this experience, and
3) my apparent memory is causally dependent, in the right kind of way, on this past experience.[7]

Quasi-memory is a broader concept than ordinary memory in that the experience remembered may be anyone's experience, not just the experience of the person having the memory. In broadening the definition of memory in this way, Parfit claims that he avoids the problem of circularity that Locke has been accused of with respect to his memory criterion.[8]

It is probable that many psychological connections hold between my present state and my reading the newspaper yesterday. Connectedness between these states can hold to any degree. In order to be the same self as I was yesterday, Parfit maintains, there must be strong connectedness between my self then and my self now. What constitutes strong connectedness cannot be precisely defined, but it should generally be understood as the degree to which we can say there are enough direct connections. Parfit arbitrarily decides that we can say there are enough direct connections if:

> ...the number of connections, over any day, is at least half the number of direct connections that hold, over every day, in the lives of nearly every actual person.[9]

Connectedness is not a transitive relation. I may now be connected with the self who was reading yesterday, but

such connectedness will not necessarily hold in five weeks. Rather, I may then be connected with the self of the week before, who may in turn be connected with yesterday's self. It is through a chain of such connections that we view a person as being the same person over time.

Although it is unlikely that enough direct psychological connections will hold between myself now and the experiences I had ten years ago, it is probable that there exists a continuity of strong connections, and such continuity will determine that I am the same person now as I was then. Psychological continuity is the holding of overlapping chains of strong connectedness. It is this relation which links connected experiences, and distinguishes a particular set of experiences as being those had by the same person.

I may not remember a sufficient number of experiences from ten years ago, but strong connectedness probably holds between myself now and my experiences of three years ago. My self of three years ago is, in turn, probably connected with my self three years before that, and so on. In this way, my experiences are linked through a continuity of my mental states, such that each subsequent period of connectedness is continuous with the previous period.

Continuity, then, <u>is</u> a transitive relation. My mental states today are related to those of yesterday by way of

connectedness and continuity, and yesterday's mental states are related to those of the previous year by way of connectedness and continuity. None of these states are connected to one another, but the chain of the various states provides continuity. This means that my experience of, say, eating ice cream three years ago is continuous with my experience of reading the newspaper yesterday.

In summary, Parfit argues for the following:

1. We are not separately existing entities, apart from our brains and bodies, and various interrelated physical and mental events. Questions of identity are actually concerned with relation R, and identity is not in fact what matters when we ask whether someone today is the same person we knew sixty years ago. Such questions do not really pertain to numerical or qualitative identity, but rather to psychological and physical connectedness and continuity, with any cause.

2. Identity need not be determinate. There are cases in which it is not possible to determine whether I am the same person as I was yesterday. Such cases do not matter, since we can still know all there is to know about the situation without having answers to these questions.

3. It is not necessary to ascribe experiences to a particular person. It is possible to explain the existence of experiences without saying that these experiences belong to anyone. We can describe out lives in an impersonal way.[10]

CHAPTER TWO

1. Parfit's Methodology

In dealing with the question of what is involved in the continued existence of the same person over time, Derek Parfit defends the Reductionist view of personal identity. Parfit offers a series of arguments that are designed to tell us several critical things about personal identity. First, that an impersonal description of persons is possible; secondly, that personal identity is not in fact what matters when we consider questions of survival; and, in conclusion, that the relation that we ought to be concerned with is that of psychological connectedness and continuity, for it is this relation that matters most.

Essentially, Parfit is replacing one relation with another. His claim is that theories concerning personal identity have been committing a fatal error: they have been seeking to preserve and explain the wrong relation in their exploration of what we have in mind when we use the language of personal identity. We may think that identity is what matters, but, Parfit tells us, if we look at what is

actually at stake we will see that it is not the identity relation after all.

This claim is supported by way of arguments based on imaginary cases that force us to decide on what is crucial in questions of survival. The advantage to this method is that it brings to our attention distinctions that ought to be drawn with respect to what is most important in the continued existence of persons. Ordinary discourse muddles these distinctions because we do not normally need to employ them.

It is significant that Parfit believes that there is one critical relation that matters in questions of survival. He attempts to dispel views in defense of what has traditionally been regarded as the relevant relation, and in turn defends an alternative critical relation. We are told by Parfit that his alternative accomplishes two things:

1) It avoids all the errors inherent in the established view that the identity relation is what matters; and

2) it solves problems that cannot be explained by this claim.

However, I argue that this approach leaves Parfit open to many of the same criticisms that he launches against the

identity relation. Parfit must convince us that his criterion is what matters, and he must also convince us that his criterion can withstand his own arguments against the established view.

I essentially agree with Parfit that the identity relation can not be defended as the critical relation in questions of survival. But I do not accept his claim that psychological connectedness and continuity is the relation that matters, and that it is all that matters. I will argue that this claim is inadequate in that:

1) It is based on arguments that depend on a number of inconsistencies and that draw arbitrary conclusions; and

2) it is itself subject to some of the very arguments used by Parfit against the identity relation.

2. Parfit-Mars

Parfit begins his defense of the "true view" of personal identity with an imagined trip to Mars. This trip involves entering a Teletransporter and pressing a button that will scan Parfit's cells, producing a perfect replica of him on Mars. The new, more sophisticated Teletransporter does not destroy Parfit on Earth, as did the obsolete

machine. Rather, if all goes well, Parfit-Earth will be able to speak to Parfit-Mars within several minutes of pressing the button. The replica will be psychologically continuous with Parfit-Earth, meaning that he will have direct psychological connectedness with him: a sufficient number of psychological connections -- of memories, intentions, and beliefs -- will be present such that psychological continuity will hold. This fact makes Parfit-Mars qualitatively, though not numerically, identical with Parfit-Earth.

In this particular case the Teletransporter is deficient in one way: its scanner harms Parfit's cardiac system and Parfit can now expect to die of heart failure within the next few days -- Parfit-Earth, that is.

Given the existence of his replica on Mars, Parfit argues that it is most rational to view this situation as being about as good as ordinary survival. It is different from ordinary survival to the extent that the identity relation does not hold in the normal way. It is also different in that Parfit-Mars, though psychologically continuous with Parfit-Earth, will be branching off from his original, thereby reducing connectedness quite rapidly. However, these differences do not affect what is most crucial: direct psychological connectedness between the two.

As his replica will have all the same intentions and

memories that Parfit-Earth has formed over his lifetime, Parfit's fears about death should be greatly reduced. In some sense he will not be dying, at least not in the ordinary sense.

The case of Parfit-Earth and Parfit-Mars is paradigmatic of Parfit's entire argument and it summarizes most of what he argues for in his defense of Reductionism. Parfit offers his general argument in various parts, with each part presented as a new argument that is meant to show that traditional views of identity fail to explain crucial questions. These arguments build on one another, and are meant to point to a conclusion that Parfit views as being the natural and the only rational outcome of each step taken.

The essential point made by the case of the Teletransporter is that what matters most in the question of Parfit's survival still holds. This claim is based on the hypothesis that experiences may be described in an impersonal way and that it is possible to fully describe a person's life without ever referring to a particular person. In other words, experiences need not be explained as being those of a particular subject.

This is a claim directly against the non-Reductionist view that the existence of persons necessarily involves a separately existing entity, such as a Cartesian ego, or a

soul. If we can describe an experience without referring to a particular subject of that experience, then it follows that there need not be a particular subject. This refutes the claim that a separate substance or entity is essential in the having of an experience.

Parfit's defense of the claim that experiences may be described in an impersonal way involves three arguments. Each argument is important in its own right in that each builds on the others and their respective assumptions and conclusions. These arguments are:

The case of Jane and Paul

Divided Minds

Parfit's Division

In examining these cases I will show that Parfit's arguments in defense of the impersonal view are inadequate, and that they are subject to the very criticisms that Parfit himself launches against the identity relation.

3. Jane and Paul

In the case of Jane and Paul, copies of Paul's memory-traces are implanted in Jane's brain. The result of this is that Jane has certain memories that are actually memories of experiences had by Paul. They are Paul's memories, but Jane is experiencing them as her own.

Jane has memories of Paul's experiences and she is not able to distinguish between those experiences and her own, except where she knows that it is impossible for certain memories to refer to her own experiences. Thus, Jane is able to distinguish between her own actual memories and those of Paul's only by way of rational thought: the memories themselves are just like her own; the difference is that in some cases Jane remembers things that she knows she did not experience. She knows that some of Paul's memory-traces are in her brain, and so she is able to infer that certain memories that conflict with others must be those of Paul's. For instance, Jane has memories of being in Venice. She knows that she has never been to Venice, and so she is able to conclude that these memories cannot be her own.

But there are some memories that are interchangeable, and in the case of these memories there is no way that Jane can tell which are memories of her own actual experiences and which are memories of Paul's experiences.

One example of this might be, according to Parfit, the hearing of a tune:

> She might have to say, 'I do vividly seem to remember hearing that tune. But I do not know whether it was I or Paul who heard that tune.'[11]

This case is important to Parfit's overall argument for two reasons. First, it introduces the concept of quasi-memory (Q-memory). Quasi-memory is a wider concept of memory that is meant to take care of Butler's criticism that memory presupposes personal identity and so cannot qualify as a criterion of personal identity. Q-memory differs from ordinary memory in that the person who originally had the experience remembered and the person having the memory of that experience need not necessarily be one and the same.

Jane's case offers an example of quasi-memory: Jane has a memory of having an experience, and it is true that someone did have that experience; it is also true that Jane, the subject of the apparent memory, and Paul, the subject of the experience, are not one and the same.

Jane and Paul's case is also important in a more general way. In arguing for Q-memory, Parfit is implicitly offering a critical premise concerning the nature of experience: namely, the premise that experiences are by nature interchangeable -- that two people having a particular experience, such as looking at a particular object, can have exactly the same experience. Parfit's example of the tune with unknown origins necessarily commits him to this specific epistemological position.

Jane has a memory of hearing a tune, yet is not sure who actually heard it -- "Was it I or Paul?" This

description of what happens when Paul's memories are transferred to Jane's brain can be accurate if and only if it is exactly the same for Jane to have experience X (e.g., to hear a tune) as it is for Paul to have experience X. In saying "I heard that tune", the most that Jane could claim (according to Parfit) is that she knows she heard it because she knows she was present when that tune was played. There is nothing about the experience *itself* that would reveal to Jane that it was her experience, and not Paul's.

Jane's memory of Venice must be a memory of Paul's experience because Jane has never been to Venice. If, however, Jane had been to Venice and had memories of being in Venice, she would not be able to distinguish between the two sets of memories. According to Parfit, the only difference between her own experiences and those of Paul's is one of circumstance -- the fact that Jane happens to be the one having the experience, rather than Paul.

Essentially, this example reveals that Parfit holds the view that, given experience X, the experience is identical whether it is had by person A or by person B. Parfit is claiming that it is qualitatively the same in all respects for person A to have experience X as it is for person B to have experience X. It is important to recognize this position as an implicit part of Parfit's definition of quasi-memory.

If, however, we were to adjust the case of Jane and Paul according to an alternate view of experience, we would be faced with a conclusion opposite to the one drawn by Parfit. Suppose, for example, that we assume specific experiences are not identical for different individuals having those experiences, but rather that specific experiences are unique for each individual.

This new view might make the following claim:

The having of an experience, such as looking at a painting in the Modern Wing of the Metropolitan Museum, is the result of a certain process that is unique to each individual. On seeing an object, the senses are stimulated, and this in turn activates a mental process that interprets what has been seen. This process of digesting information follows a certain pattern that reflects the particular mental structure and perspective of the individual. Each person's perspective is slightly different from that of others, according to the combination of their past experiences, and so results in an experience unique to that individual.

Under this new premise, when Paul's memory-traces are transplanted to Jane's brain, Jane would face no confusion in distinguishing between her own memories and those of Paul's. In fact, it is unlikely that Jane would regard Paul's memories as being memories at all. Paul's memories

would be so fundamentally different from anything Jane has ever experienced that Jane would immediately identify Paul's memories as being foreign to her own memories. Were Jane to experience an entirely new sensation that is fundamentally different from any of her usual five senses, there would be no confusion in her mind that this sensation was unlike any she had ever experienced. Experiencing one of Paul's memories could be understood as this remarkable kind of experience -- one so unlike her own that she would regard it as a unique experience and would classify it as such.

Consider, for example, the case of Jane and Paul under slightly different circumstances. Suppose that the case is the same, with the qualification that Jane is a concert cellist, and Paul is an accountant who listens occasionally to popular music and knows nothing about classical music. Suppose further that Jane and Paul are in the same room, listening to Beethoven. Although the conditions of listening to the music are the same, it is not hard to imagine that the experience of listening to Beethoven will be entirely different for Jane than it will be for Paul. The music may trigger sensations in Jane that are unlike the ones Paul will experience, and it may elicit images for Paul that Jane could never experience. Jane's knowledge of music may cause her to respond to subtleties in the music that Paul would be unable to detect. And Paul's vivid imagination, which he

nurtures during his hours spent in an office tower, may bring about images that Jane could not fathom.

If we accept this description then it follows that, if Jane were to experience Paul's memory of hearing a tune, Jane would know what the experience was like for Paul -- there would be no doubt that the memory did not come from her own experience. A memory transfer of Paul's experiences would give Jane knowledge of what it is like for Paul to have given experiences. Jane would know about Paul's experience of listening to Beethoven, but she would never confuse that experience with her own.

If it is agreed that, if Jane is a concert cellist, her experience of listening to Beethoven will be fundamentally different from what that experience will be like for Paul, then we are adopting the view that one's experiences are dependent on the conditions under which those experiences are had. To listen to a tune is not simply a matter of being within proximity of a record being played. The nature of the experience is determined by certain external factors, such as the acoustics of the room, and is also determined by the particular individual listening to the record. It the person is very familiar with the music, the experience will be of one sort. If, on the other hand, the individual listening has never heard Beethoven's music and expects to hate it, the experience will be of another.

And this point need not be limited to such obvious differences. We may extend it to the more general claim that all experiences are dependent on the particular individual having the experience -- that all sensations are interpreted by the individual's point of view -- and that the sensations experienced by one individual may be radically different from those of another, or all other individuals.

Parfit would probably argue that this point of view can be explained in terms of relation R. However, I would not consider this a reasonable response since quasi-memory does not hold in the case of Jane and Paul. Strictly speaking, when Jane experiences Paul's memories, Jane <u>is</u> experiencing memories. Since these memories are memories of experiences had by someone, this new case does, in one sense, satisfy Parfit's criterion for Q-memory. But his criterion is satisfied in only a limited and trivial way, and certainly not to the extent intended by Parfit. As far as Jane is concerned, her experiences of Paul's memories may be regarded not as memories but rather, as I have tried to show, as new kinds of experiences that cannot be defined as memories of either Jane's or Paul's experiences. It is only if Jane is told that these strange sensations are to be known as "Paul's memories" that she will know to call them that.

If Paul's memories are not experienced by Jane as memories then *there is no actual link between Jane's memories and those of Paul.* Under this interpretation of memory transfer, continuity does not hold between Jane's mental states and Paul's mental states.

Parfit's definition of quasi-memory is that there is a memory of an experience, that someone did have that experience, and that the memory is causally dependent on that past experience.[12] In Parfit's example of Jane and Paul, the important assumption is that a link will exist between Paul's memories and Jane's memories. It is this link that will satisfy the criterion of psychological connectedness and continuity to Jane's mental states, and that will provide her with "all that matters" regardless of whether the experiences remembered are her own or Paul's. Parfit's new definition of memory is meant to show that the continuity of memory does not necessarily presuppose personal identity.

However, my description of what does actually occur with the transfer of Paul's memory traces to Jane's brain tells us that Parfit's position is not conclusive. Rather, it is possible to take the same case, alter its circumstances slightly, and draw quite opposite conclusions.

The case of Jane and Paul, then, is relevant only if we accept a certain "hidden" view of what is involved in the

having of an experience. If we hold an alternative view of experience, the conclusions Parfit draws from the case no longer follow.

This tells us something important about Parfit's argument:

Parfit's claim concerning quasi-memory is based on a significant premise which he neither addresses nor defends. Given the lack of defense of this premise concerning the nature of experience, the conclusion Parfit draws from the case of Jane and Paul appears to be a dubious one. And, of course, if Parfit's argument in favor of quasi-memory is not sound, then Butler's criticism must still be addressed.

4. Divided Minds

In case we are not convinced by the case of Jane and Paul, Parfit offers another argument that is meant to show, in speaking of experiences, that it is not necessary to refer to a subject of those experiences. This is the argument from divided minds.

The case of divided minds involves one person whose mind is divided in half. While this is an imaginary case, Parfit regards it as not being "deeply impossible" since it is an extension of the established scientific view that two

hemispheres of the brain are capable, under certain circumstances, of operating independently from one another.[13] This fact is meant to support Parfit's use of this case.

To begin with, Parfit imagines that he is able to divide his left hemisphere from his right and that he applies this talent to his performance on a physics exam. In the physics exam Parfit is faced with a problem that seems to have two possible solutions and only fifteen minutes in which to solve it. He decides to divide his mind so that each half is able to work on one possible solution, and after ten minutes reunite the two halves so that he has time to write down the best result.

The point of this case is to consider what it would be like to experience such a division. Parfit expects that his experience of the physics exam would be as follows:

After deciding to divide, he is unaware of the division. Both his right and left hands are working on the problem at the same time and neither half of his brain is aware of the work being done by the other. Parfit's right hemisphere is using his left hand in the calculation, and it can see that his right hand is also working on the problem. But this, Parfit says, is akin to the awareness he might have of his neighbor's hand. He has no special knowledge of the thoughts or experiences taking place in his left

hemisphere. Thus, what is normally experienced as one stream of consciousness is now experienced as two simultaneous streams, each stream operating independently from the other.

As long as the two streams remain divided, neither stream is aware of the experiences of the other. However, once Parfit reunites his two hemispheres he becomes equally aware of both sets of experiences. He can remember working out the problem from two angles -- one attacked by his right side, the other by his left -- and yet he knows that at the time of division neither side was conscious of the other.

In this case relation R continues to hold between Parfit's mental states prior to division and each of his two sets of mental states after division. Psychological connectedness and continuity does not hold between the two sets themselves as each stream has branched from the main; but when Parfit unifies his two streams each stream is connected to the single stream and so relation R remains uninterrupted. In thinking back on the division Parfit will clearly remember each stream as something he experienced.

According to Parfit this imagined case refutes the ownership view of experience. The claim that experiences must refer to a single subject -- that one's mental states necessarily *belong* to one's self -- assumes that there is a single consciousness that subsists in the having of all experiences. During division, however, there are two streams

of consciousness, each of which is equally conscious, and each of which displays a separate unity of consciousness.

Parfit claims that since this case involves two unities of consciousness it is not coherent to argue for a single unity. The ownership view suggests that we can always ascribe experiences to a single subject. But how, Parfit asks, is ownership to be explained in the case of division? On his view it is not plausible to claim that both sets of experiences belong to the same subject as this "makes the two unities one".[14] If we accept this interpretation of what happens with division -- that there are two simultaneous sets of consciousness -- then, Parfit argues, we are necessarily denying the view that there is a 'necessary unity of consciousness' for all experience.

Parfit explains this position in the following way:

> We can come to believe that a person's mental history need not be like a canal, with only one channel, but could be like a river, occasionally having separate streams. I suggest that we can also imagine what it would be like to divide and reunite our minds.[15]

The division of Parfit's mind during the physics exam is similar to a river with two separate streams. And Parfit regards this analogy as evidence in favor of his conclusion

that, with division, it is not possible to explain the two unities as belonging to a single subject. Since there are two separate streams we cannot say that each stream belongs to the same subject. This in turn serves to refute the existence of a "necessary unity of consciousness" in the case of division.

If it is true that a "necessary unity of consciousness" is not present in division then, Parfit argues, there is no reason to believe that it is *ever* present, even under ordinary circumstances.

While these conclusions are coherent, it is possible to draw opposite and equally coherent conclusions from the same imaginary case. We could accept Parfit's description of the physics exam and still claim that there is a "necessary unity of consciousness", even in the case of division. For example, one might argue that such unity is a requirement for all experience and that it is not possible to make sense of the concept of experience without adopting the concept of unity as an integral part of each experience. According to this view, "unity" does not involve a further fact or a separate substance but, rather, it is a fundamental part of the very notion of an experience. There is no separation between the two concepts: in speaking of an experience we are inevitably referring to a "necessary unity of consciousness".

As Parfit recognizes, if such unity is necessary for the having of experiences under normal conditions, then it must also be necessary for experience in the case of division. On Parfit's view, the very notion of division refutes the concept of unity, for how can there be unity between two separate parts? And it is clear, says Parfit, that his two streams are indeed separate, as neither stream is aware of the other during division. Parfit tells us:

> We need to explain the unity of consciousness within each of my two streams of consciousness, or in each half of my divided mind. We cannot explain these two unities by claiming that all of these experiences are being had by me at this time. This makes the two unities one.[16]

But why does Parfit assume that "unity of consciousness" cannot withstand division? In his review of the idea of the "necessary unity of consciousness" Parfit claims that his case of the physics exam refutes the existence of unity:

> It might be objected that my description ignores 'the necessary unity of consciousness'. But I have not ignored this

> alleged necessity. I have denied it. What is a fact must be possible. And it is a fact that people with disconnected hemispheres have two separate streams of consciousness - two series of thoughts and experiences, in having each of which they are unaware of the other. Each of these two streams separately displays unity of consciousness.[17]

The point of the physics exam is to dispel the unity view of experience and to show that, without unity, there is no reason to hang on to the conventional view of a subject of experience. The unity of different experiences in each stream of Parfit's consciousness cannot, Parfit tells us, be explained by ascribing all the experiences to Parfit. Since unity does not hold between the two streams (he argues), we cannot ascribe unity to a single subject, him.

The question Parfit raises is: Who is the subject of these two streams of experience? It cannot be one person, so is it two? Or three? For Parfit, these answers are inadmissible as there are only two ways of describing division:

1) We may claim that each stream of experience is that of a subject who is not Parfit (and therefore not a person); or

2) we may adopt the Reductionist view and claim that any person's life may be described in impersonal terms.

Parfit tells us that although we feel we ought to be able to define the number of subjects involved, in the case of a divided mind it is not possible to do so. The point is that it is not *necessary* to refer to a subject of experience. It is on these grounds that Parfit adopts alternative (2) and argues for the Reductionist view in his description of what happens with division. He explains this as follows:

> Because we ascribe thoughts to thinkers it is thought that thinkers exist. But thinkers are not separately existing entities. The existence of a thinker just involves the existence of his brain and body, the doing of his deeds, the thinking of his thoughts, and the occurrence of certain other physical and mental events. We could therefore redescribe any person's life in impersonal terms. In explaining the unity of this life, we need not claim that it is the life of a particular person. We could describe what, at different times, was thought and felt and observed and done, and how these various events were interrelated. Persons would be mentioned here only

> in the description of the content of many thoughts, desires, memories, and so on. Persons need not be claimed to be the thinkers of any of these thoughts.[18]

This position hinges on Parfit's argument concerning unity of consciousness. The physics exam is meant to demonstrate that we cannot make sense of unity, and this in turn is meant to support the view that we need not refer to a *subject* of experience. If there cannot be a single unity during division, then there cannot be a single subject. If, however, we can make sense of some concept of unity during division, as I believe we can, then Parfit's claim concerning subjects does not necessarily follow.

The problem is that Parfit addresses these concepts but he does not analyze them. He operates on the assumption that we understand what he means by "unity of consciousness" and "state of awareness" without defining these concepts, and he proceeds to offer arguments using these undefined terms. How can Parfit argue that unity cannot apply to division when he has not provided a coherent explanation of what he means by "unity"?

Parfit assumes that we cannot make sense of unity in the case of division and from this he infers that we have grounds for denying the theory altogether. His argument is based on a literal interpretation of unity and of divided minds. I would argue that to speak of unity of consciousness

is not necessarily to speak of one experience; and to speak of divided minds is not necessarily to speak of two separate minds, each requiring its own unity. It may indeed be possible, under at least one interpretation, to make sense of unity in the case of division, or for unity to withstand division. For example, we are all capable of doing several things at once: we can carry on a conversation while working on a puzzle and, at the same time, have memories of a childhood experience. It would be most illogical to argue that the very fact that we can have several experiences at one time is evidence against the concept of unity of consciousness.

Division could be understood as simply an extension of the ability to think of several things at once. To divide one's mind and separate one stream from another is conceivably no more complicated for someone with the power of division than is having several thoughts at once for the ordinary person. The view that unity of consciousness is present under ordinary circumstances, without division, could support the view that unity is present during division. Unity could be regarded as a film that covers its various parts: with division, its parts (experiences) are separated in two, but continue to be bound and covered by the unity of consciousness that makes each set of thoughts possible.

It is significant that while Parfit claims to deny unity he actually employs the concept of unity. This is how Parfit explains the Reductionist view of what happens in each stream of his divided mind:

> ...what unites my experiences in my right-handed stream is that there is, at any time, a single state of awareness of these various experiences. There is a state of awareness of having certain thoughts, feeling writer's cramp, and hearing the sound of a squeaking pen. At the same time, there is another state of awareness of the various experiences in my left-handed stream. My mind is divided because there is no single state of awareness of both of these sets of experiences.[19]

Parfit is concerned with what it is that makes his experiences his. This, Parfit argues, is explained by way of a "state of awareness". But it is not clear how this differs from the "unity of consciousness" which he claims to have refuted. Parfit decides to re-unite. One could argue that this decision implies at least some sort of unity. On these terms, Parfit himself does not operate within the notion of two entirely separate streams of consciousness.

The physics exam does not in itself speak against unity of consciousness. We could adopt each of Parfit's points and still maintain the position that there is an ever-present unity of consciousness making each specific experience possible. The case of divided minds does not in itself deny the "alleged necessity" of the unity of consciousness. It would not be irrational to accept unity in the face of Parfit's example, and this conclusion would not be in conflict with the terms of his argument. For instance, one could draw two pictures of division, one picture with unity (this would be my picture) and one without (Parfit's picture). While this may tell us that unity is not *necessary*, it does not deny that unity may be a *possibility*.

Regardless of the position that is finally taken, there is nothing in Parfit's argument that forces us to accept his claims about unity. His view is one possible interpretation and conclusion of the case he has presented, but not a necessary conclusion.

If it is true that Parfit's conclusion does not necessarily follow, then we should ask why this is so. Is it the case that Parfit's arguments can support multiple and even conflicting conclusions? Is it the case that Parfit is drawing arbitrary conclusions?

5. Parfit's Division

The next step to Parfit's argument involves a more complicated case of division. In this case Parfit is one of three identical triplets. There is an accident and all three brothers are injured. Parfit's body is fatally injured but his brain remains perfectly healthy. At the same time, the brains of his two brothers are fatally injured, and their bodies remain healthy. In this imaginary world it is possible not only to divide one's mind in half but also to transplant one half of a brain into another body.

As there are two healthy bodies and only one brain, Parfit's brain is divided in half and each half is transplanted into one of his brother's bodies. After the operation there are two surviving persons, both of whom believe themselves to be Parfit. All their memories, their beliefs, and their intentions are connected and continuous with those belonging to Parfit prior to the operation. The question is, who is Parfit?

Parfit considers different possible answers to this question. He concludes that it would be wrong to regard this division as involving his death, since his brain continues to exist and remains psychologically continuous with his brain before division. As it is at least possible to survive with only half of his brain, Parfit claims there is no

reason to believe that he would not survive with half a healthy brain in one of his brother's bodies. And if he can survive under such circumstances in one body, there is no reason to believe that he could not survive with two halves of his brain, each in two separate bodies.

If Parfit does indeed survive division, then does he survive as only one of the two people, or does he survive as both? Should we regard Parfit as having been split in half, and say that both persons are Parfit? It does not make sense to regard only one of the two as Parfit (he argues), as each is psychologically continuous with the original person, and there are no apparent grounds for choosing between one or the other as the sole survivor.

Can we claim, then, that both resulting persons are Parfit? This claim, Parfit explains, may be made as follows:

> What we have called 'the two resulting people' are not two people. They are one person. I do not survive this operation. Its effect is to give me two bodies, and a divided mind.[20]

Parfit argues against this description on the grounds that it greatly distorts our concept of a person. When we speak of persons, we normally speak of one body and one

brain. To think of one person occupying two bodies is too unlike what we mean by 'person' to warrant use of this description in this case.

With division the two resulting persons will become less and less psychologically continuous with one another, to the point that they may eventually not recognize one another. To illustrate this Parfit imagines the two people living in different parts of the world, their memories weakening, and their appearances changing. They might meet each other years later on the tennis court, each one unaware of who his opponent actually is. Would we say:

"What you see out there is a single person, playing tennis with himself. In each half of his mind he mistakenly believes he is playing tennis with someone else"?[21]

For Parfit, this statement cannot make sense, as it fails to employ the ordinary meaning of 'person'. Thus, he concludes, it is incoherent to claim that Parfit would survive as both resulting people.

If we do not regard division as involving Parfit's death, and we do not consider Parfit to have survived as only one of the two resulting people, nor as both resulting people, then how are we to answer the question of who survives division?

The point, we are told, is that the question itself is misplaced. It is misplaced because it is based on the false assumption that there must actually be an answer to the question. This assumption, in turn, is based on the false view that it must always matter whether a particular person survives; *it is based on the false belief that personal identity is what matters in questions of survival.*

Neither of the possible descriptions of who survives division is an accurate description; this, Parfit says, tells us that there is no *necessary* answer to the question of who survives. There is no necessary answer because we can interpret the case in a number of different ways and these descriptions will not tell us anything we did not already know prior to our search for an answer. Each description is just one possible way of describing the same outcome, and we know all we need to know without answering the question of survival. In short, the question of identity is what Parfit calls *an empty question.*

To regard the question of survival as being an empty question is to support the view that we can know all that matters about Parfit's continued existence through time *without* addressing the question of his survival. We know that with division Parfit's brain will be divided in half, that each half will continue to exist in the bodies of his two brothers, and that each half will be psychologically

connected and continuous with his brain prior to its division.

This is all we need to know, Parfit argues, because it is all that matters in questions of survival. That it may be difficult to decide how we should describe the two people that live with each half of Parfit's brain is not in itself important, and we need not have an answer. The above description tells us everything that matters about Parfit and what has happened to him. Whether we call the two people Parfit or call them some other name has little effect on the actual outcome.

It may be that we find it problematic to regard the two resulting people as Parfit, and so decide to think of them not as Parfit but as two separate individuals, each of whom at one time were Parfit. On these terms, the identity relation does not hold between Parfit and those following his division: it is not the case that Parfit prior to division is identical with the two halves of his brain in their new bodies. Although identity does not hold, we could say that what matters continues to hold, regardless of whether or not Parfit still exists in the sense of numerical and qualitative identity.

Given that relation R continues to hold between Parfit before division and the two halves of his brain after division, Parfit regards division as being about as good as

ordinary survival. While Parfit will clearly not survive in the ordinary sense, what matters about his existence will still hold (he argues) and this is equivalent to ordinary survival.

This is the Reductionist view of what happens and of what matters with division. On this view, to know that connectedness and continuity will continue to hold between Parfit before division and the two resulting people after division is to know everything. Questions of identity and survival involve nothing further, so there is no need to insist on the question of survival. According to Parfit, such a question is only relevant if we believe that persons involve separately existing entities. If we cannot defend this belief, then we may as well abandon the question.

Parfit's division, then, is meant to demonstrate two things:

1) That identity is not the relation that matters in questions of survival; and
2) that what matters in questions of survival is psychological connectedness and continuity.

As Parfit agrees, it is indeed coherent to claim that Parfit does survive division, and to claim that he survives as both resulting people. There is, as we have seen, no

reason to believe that Parfit does not survive division: each half of his brain can, we assume, function on its own, and can survive division, and each half is psychologically connected and continuous with Parfit prior to division. The question is whether Parfit can be said to survive as both remaining persons, or, more accurately, how it is that we should describe the remaining persons. What matters, then, is not whether Parfit survives, but rather how and in what form he survives.

Parfit claims that relation R holds in the case of division. It may be true that we can survive, physically and actually, with only one half of our brains, but this does not demonstrate that relation R will continue to hold. Parfit claims that after division each half of his brain will remain psychologically connected and continuous with his brain prior to division, but this claim may be unfounded. It could just as easily be claimed that the shock of division will break all psychological connections, and that after division each half of Parfit's brain will be perfectly blank in the two surviving bodies, with no memories at all. If this were to happen, relation R would not hold. Even if relation R did not hold, we might still regard Parfit as surviving division, to at least some degree. If this were true, then Relation R would *not* tell us all that matters about Parfit's continued survival.

I would argue that Parfit's conclusion that relation R continues to hold is not a necessary conclusion in the case of division. Parfit assumes that relation R will hold, but he does not show this to be necessarily true. Parfit's argument that the identity relation is not the relation that matters in division does not in itself demonstrate that relation R is the significant relation, and that it is the only significant relation, in all cases.

6. The Spectrum

Parfit's claim that the identity relation is not the relation that matters is supported by a further set of arguments, each involving a spectrum. The imaginary spectrum cases are meant to demonstrate that identity may be indeterminate and that, as such, identity is not the relation that matters. If it is shown that we can know all we need to know about survival without answering the question of identity, then this, Parfit argues, will confirm the claim that questions of identity are 'empty' questions.

(i) The Psychological Spectrum

The first case Parfit offers is that of the Psychological Spectrum. This imaginary spectrum represents

all the possible degrees of psychological connectedness that could hold between Parfit as he now is and a person wholly unlike Parfit, such as Napoleon. Parfit and all his present psychological states (memories, intentions, and beliefs) are represented by the extreme right of the spectrum, and Napoleon is represented by the far left. There are no psychological connections between Parfit and Napoleon, or between the two extremes of the spectrum, but there is a broad range between these two ends where varying degrees of psychological connectedness could hold. In this case, a surgeon is able to effect a change in the psychological connections by flipping switches on the spectrum.

To begin with, the surgeon is faced with the extreme right of the spectrum: Parfit as he now is, with all his usual psychological connections. Were the surgeon to flip all the switches at one time, all of Parfit's psychological connections would be severed; he would lose all psychological connectedness to the new individual, and would become fully psychologically connected to Napoleon. It would be clear, under these circumstances, that Parfit would no longer exist.

The surgeon might also flip only a few switches, thus effecting a change that would place Parfit closer to his end of the spectrum, with only a slight reduction of his psychological connections. In this case, Parfit would take

on some of Napoleon's memories, and some of his character, and lose some of his own. He would, however, maintain enough psychological connections to his past self that there would be no question of his survival.

In these two extreme cases Parfit clearly exists at one end of the spectrum and does not exist at the other. However, were the surgeon to flip additional switches so that Parfit held a greater degree of psychological connectedness with Napoleon than with himself, yet maintained some percentage of his own mental states, it might not be so easy to determine whether Parfit could be said to survive.

For example, the surgeon might flip a sufficient number of switches so that Parfit's love of strawberries is replaced with Napoleon's yearning for power, and two years of Parfit's childhood memories are eliminated and replaced by thirty of Napoleon's delusions. Parfit would continue to hold all his memories and desires, except for these relatively minor alterations to his character. While Parfit could be said to survive under these circumstances, his survival would not be as clear were the surgeon to replace, say, an additional sixty memories and desires with those of Napoleon's.

The problem that Parfit presents us with is that of determining the point at which Parfit ceases to exist and

Napoleon regains his existence. Each possible point on the spectrum represents a slight progression in the degree of psychological connections that are held; in order to answer the question of Parfit's survival we must define the exact number of connections that are necessary in order for him to continue to exist. There must be a critical point at which Parfit has lost enough psychological connections that he does not survive the surgeon's changes.

Parfit argues that it is not possible to define that critical point since each change is in itself so mild that no one switch could effect such a dramatic change:

> It is hard to believe both that I would survive in one of these cases, and that, in the next case, I would cease to exist. Whether I continue to exist cannot be plausibly thought to depend on whether I would lose just a few more memories, and have a few more delusory memories, and have my character changed in some small way. If no such small change could cause me to cease to exist, I would continue to exist in all of these cases I would continue to exist even in the case at the far end of this spectrum. In this case, between me now and the resulting person, there would be no psychological connections.[22]

There is no question that Parfit does not survive at the extreme left of the spectrum, and that he does survive at the extreme right. But there cannot be one critical point at which he ceases to exist, and this tells us that on certain points of the spectrum the question of his survival is impossible to answer; there are points at which Parfit's survival is indeterminate.

This, Parfit claims, confirms that the question of survival is not actually the relevant question. We may choose to answer the question of Parfit's survival one way or the other at each particular point, but this choice would be arbitrary and therefore trivial.

Even though it may not be possible to answer the question of Parfit's survival, we can still know the extent to which psychological connectedness holds for him. Depending on where on the spectrum the surgeon decides to settle, we can know the exact degree of psychological connections that hold between Parfit and Napoleon, and so we can fully describe the resulting individual. We can describe that individual as, for example, having three quarters of his mental states continuous with those of Napoleon, and one quarter continuous with Parfit. If someone asks, "Yes, but does that mean Parfit survives or that Napoleon survives?", we can logically reply that the question is irrelevant. We can describe all that matters about survival without

answering that particular question; the question of Parfit's survival is in this case an *empty question*.

The point of the psychological spectrum is to demonstrate that we can know all that matters about questions of survival even where identity may be indeterminate. Parfit's argument is that concerns about survival should be limited to concerns about the degree of psychological connectedness and continuity that holds in any given case. If we know the degree to which relation R holds in a given instance, then we know all we need to know. This must be true, Parfit claims, because it is all that there is in the having of experience: relation R fully describes all our mental states, and therefore tells us everything that matters about our survival.

Parfit's defense of this major claim is, in part, dependent on his spectrum arguments. The psychological spectrum serves to demonstrate that identity may be indeterminate. Since it is indeterminate, identity cannot be the relation that matters in questions of survival. As identity cannot be the relevant relation, what, Parfit asks, are we really looking for when we wonder whether Parfit survives at a given point on the spectrum? We are really concerned with psychological connectedness, or relation R.

There is a major leap in this line of reasoning. I may be convinced in the case of the spectrum that identity can

be shown to be indeterminate. But this does not in any way demonstrate that relation R is the relation that matters in questions of survival. We may not be able to answer the question of whether Parfit survives with only one-third of his memories and intentions and two-third's of Napoleon's memories and intentions, but the fact that we may know the exact degree of Parfit's psychological connections does not in any way help us with this question. Nor does it demonstrate that the question of Parfit's survival is irrelevant, or empty. It is logical to accept that there are some questions that cannot be answered -- that some questions are indeterminate -- but this is not grounds for dismissing the question as empty. We are left with the information we have access to -- the number of mental states continuous with Parfit, and the number continuous with Napoleon -- and are still left wondering whether, in this case, we are looking at Parfit or at Napoleon.

Consider, for instance, another kind of example, such as the life of a fetus. Parfit's spectrum, in fact, serves as an appropriate analogue for this example. The problem that ethicists and others are faced with concerning the fetus is that of determining the point at which the development of a person can be said to begin. Most are in agreement that personhood does not exist prior to fertilization of the egg, and that personhood does exist

after birth; it is hard, though, to find agreement on the question of when exactly it begins.

Each of the various stages of fertilization and gestation are represented by different points on the spectrum between these two extremes. One could argue that each neighboring case produces such a small change in the fetus that it is impossible to claim that there is one critical point at which the fetus moves from not being a person to being a person. We could have at hand all the facts concerning the fetus, and be able to fully describe every aspect of its development, and still not be able to answer the question of whether or not it is a person. And one could argue that if personhood is determined to begin at a certain point on this spectrum, that determined point is necessarily an arbitrary one.

It may be that the question of personhood can, in some sense, be indeterminate. But this does not tell us that it is an irrelevant question. We may not always have an answer to the question: "Is there a person in this case?", but this does not mean that the question does not matter. We will always care deeply about that question, regardless of whether or not we are able to answer it. In the same way, one might argue that the question of identity may not always be answerable, but this does not in itself speak against the relevance of the question.

Another important claim made by Parfit and defended by way of the Spectrum is that views concerning questions of identity must stand or fall together. Parfit argues that if we believe that identity must always be determinate, then we must be able to define that critical point on the spectrum where Parfit becomes Napoleon. If we are unable to do this, the only remaining defense in favor of the identity relation is that identity involves a "further fact" that cannot be defined in psychological terms.

This "further fact" might be a physical criterion of personal identity, such as the one offered by Bernard Williams. Williams claims that the sameness of the brain is the criterion of personal identity, and that when we are concerned with questions of survival, we are actually concerned with the existence of our brains. As long as the brain of the particular individual continues to exist, that individual can be said to exist. If we know whether or not that person's brain survives, then we know all we need to know about that person's survival. Parfit denies this claim by way of his next argument, which is the argument of the Physical Spectrum.

(ii) <u>The Physical Spectrum</u>

The Physical Spectrum provides an argument similar to that of the Psychological Spectrum, with the points on the

spectrum representing all the possible degrees of physical continuity. The extreme right end of the spectrum represents someone who is fully psychologically and physically continuous with Parfit; the other extreme represents an individual that is psychologically continuous with Parfit, but not physically continuous with him. In this case, Parfit's brain and body are replaced with exact replicas of his cells; a very small percentage of his cells are replaced at the near end of the spectrum, and all of his cells are replaced at the far end. At the far end, Parfit's brain and body are completely destroyed and later replaced by a brain and body that is a replica of his own. It is in this way that there is no physical continuity between Parfit and the resulting person at the far left of the spectrum.

For Parfit, each of the stages on the spectrum, even the case at the extreme left, should be regarded as equivalent to ordinary survival. Although the person at the far end will not be physically continuous with Parfit, he will be an exact replica of Parfit. He will have exactly the same memories and intentions and beliefs as the original Parfit, and so he will in all psychological respects be exactly like Parfit. He will not be physically identical to Parfit, but, Parfit argues, this is relatively unimportant. All that matters about Parfit's continued existence will be

true of his replica, and for this reason his replica should be regarded as being Parfit, *in all the important ways*.

This is the Reductionist view of what matters in questions of personhood. Parfit claims that this is the only rational way of understanding the question of his survival in the case of the Physical Spectrum. If, however, we do not accept this interpretation, then we are defending a rival view of what matters in questions of survival. A look at what these views might be will reveal, Parfit argues, that such views are essentially indefensible.

We could easily claim that Parfit does not surive at the extreme left of the spectrum, as his brain and body are completely destroyed. Parfit does not deny this claim, as continuity in this case is interrupted. What, though, of the other points on the spectrum? Parfit argues that he would clearly survive if only one percent of his cells were replaced since he does not need all his brain and body in order to live. Nor does he necessarily need twenty percent, or even forty percent. At what point, Parfit asks, would we say that there is a sufficient number of replica cells to support the claim that Parfit does not survive the change?

As with the Psychological Spectrum, Parfit claims that it is not possible to determine a critical point at which he does not survive, and so the question of his survival is, in all these cases, an empty question. And Williams' defense

of the physical criterion has been addressed by way of the physical spectrum, which demonstrates that it is difficult to argue that an individual's survival is dependent on the continued existence of the brain, for it is not clear where we are to draw the line with this criterion. Williams, then, is faced with the same sort of dilemma as is the psychological spectrum: that of determining the point at which survival begins and ends.

However, Parfit's reply to Williams is subject to the same criticism as that launched against his psychological spectrum: Although it may be true that identity is, in some cases, indeterminate, this does not in itself speak against the relevance of the question of identity.

In case we want to defend what Parfit refers to as a hybrid view of personal identity, and claim that either physical or psychological continuity is sufficient for the continued existence of persons, Parfit offers yet a third version of the spectrum argument. This is the Combined Spectrum.

(iii) The Combined Spectrum

The Combined Spectrum represents all possible degrees of both physical and psychological continuity, with continuity holding in the normal sense at the extreme right, and no continuity, either physical or psychological, at the

extreme left. The person at the far right of the spectrum would be Parfit as he normally is, and the person at the other end would be a replica of someone else, wholly unlike Parfit. In this example, a replica of Greta Garbo as she was at thirty is represented at the far end of the spectrum. At the far left, then, Parfit is completely destroyed and Greta Garbo is created out of entirely new matter. There is no physical or psychological continuity between that replica and Parfit.

On this spectrum, each stage represents a change in Parfit's cells from his own to those of Greta Garbo's. The new cells are not exact replicas of his own, but rather are replicas of Garbo's. This means that connectedness does not hold between Parfit's cells and the new replacement cells. Thus, both physical and psychological connectedness is interrupted.

To begin with, a small number of Parfit's cells are replaced with replicas of Garbo's cells. This might result in Parfit taking on some of Garbo's beliefs and some of her fears, as well as some of her physical characteristics. But these changes would, at this point, be so subtle they would not obviously change Parfit. Those around him might not notice, and might continue to assume that the resulting person is the original Parfit.

After a few more changes, Parfit might visibly come to resemble Garbo; perhaps, though, he would retain enough of his own characteristics that we might still consider him Parfit, though somewhat changed. At some point in this process a sufficient number of Parfit's cells would be replaced so that there would be no question that he was now Garbo. The problem here is the same: At what point has Parfit become Garbo?

Again, Parfit claims that there is no obvious point where he ceases to be himself. If we believe that there is something further to the existence of persons than their physical and mental states, then there would be one critical thing that is present when Parfit is himself, and that is no longer present when Parfit is changed to Garbo. Parfit argues that this critical thing would reveal itself to us on the spectrum and that it would solve the puzzle of Parfit's survival. Although we believe that there is a deep difference between Parfit being himself and his being Greta Garbo, it is impossible to determine the point at which Parfit becomes Garbo.

As Parfit explains:

> ...between neighboring cases in this Spectrum the differences are trivial. It is therefore hard to believe that, in one of these cases, the resulting

person would quite straightforwardly be me, and that, in the next case, he would quite straightforwardly be someone else.[23]

The Combined Spectrum is meant to finally convince us that the question of Parfit's survival is essentially an empty question. We can know all we need to know about the resulting person -- that is, we can know the exact number of physical and psychological connections that are Parfit's and those that are Garbo's -- without having an answer to the question of Parfit's survival. And this is meant to convince us of the validity of Reductionism.

If we accept Parfit's refutation of the physical criterion of personal identity and still deny his Reductionist claims, then, Parfit tells us, we must believe that there is something further to the existence of persons than the continuity of their physical and mental states. We must be defending the view that identity involves a "further fact". This further fact might be a separate entity, that is neither physical nor psychological, such as a soul or a Cartesian ego. If we are not postulating the existence of a Cartesian ego, then we are claiming that there is some other factor that determines the existence of persons.

Parfit argues that the Spectrum cases demonstrate that these views stand or fall together. If we claim that there

is something further to the existence of persons than their physical and mental states, then we must defend the existence of a separate entity, apart from these states. If we are unable to defend the existence of such an entity, then we must be able to show what that "further fact" could be. If there is such a "further fact", then we should be able to determine when that fact is present and when it is not. The Combined Spectrum tells us, Parfit claims, that there is no such further fact, for we are unable to point to the criterion that defines the continued existence of a given individual.

Even if we are convinced by the Combined Spectrum and agree that it is impossible to say exactly where Parfit ends and Garbo begins, the Spectrum is, I would argue, ineffective in convincing us of Parfit's more significant claim: namely, that Reductionism is the "true view" in questions of personal identity. It may be that identity is not *at all times* the relation that matters in questions of survival, but this is an argument against the significance of the identity relation as the *only* relation that matters. It is not necessarily, as Parfit claims it to be, an argument in favor of Reductionism.

Essentially, Parfit's argument is the following: In questions of the continued existence of persons through time, the conventional view is that personal identity is the

relation that matters. The Spectrum throws new light on this established view, and reveals that we may not always be able to determine whether identity holds. Even in such cases, we can fully describe the resulting individual. This tells us that identity is not the relation that matters but, rather, that what we really care about in questions of survival is relation R.

As I have tried to show, the fact that there may be cases where identity is indeterminate does not in itself speak against the value or significance of that relation as a concern in questions of survival. A denial of the identity relation is not in itself a sufficient defense of Reductionism. Parfit has yet to demonstrate that relation R is indeed all that matters in the continued existence of persons through time.

My claim is that there are cases where relation R <u>does not</u> hold and where we <u>do</u> know all that matters about the continued existence of persons. I also claim that the reverse can be shown to be true: there are cases where relation R <u>does</u> hold and we <u>do not</u> know all that matters in questions of survival. The following chapter is a defense of these claims.

CHAPTER THREE

1. Reductionism

The importance of the above discussion is to recognize the progression of Parfit's arguments, and to see where it is he has led us. The four arguments that I reviewed in Chapter Two are designed to convince us of the following:

1. That it is possible to describe experiences in an impersonal way;
2. that there is no such thing as a "unity of consciousness"; and, most importantly,
3. that the question of identity is in fact an "empty" question.

These claims are presented as natural and logical conclusions to Parfit's thought experiments, and each serves as an important defense of the Reductionist view of what matters in questions of identity. I have argued that there are problems with each of the conclusions drawn by Parfit,

and that alternate -- and conflicting -- conclusions may be consistent with the same arguments.

It is evident from the previous chapter that Parfit assumes an all-or-nothing position such that, on his terms, we are either Reductionists or non-Reductionists. To assert that the identity relation must matter in the case of division is, for Parfit, to assert that there must be something further to the existence of persons through time, such as a separately existing entity, or a Cartesian ego. In like manner, to claim that the identity relation is *not* the one that matters is, on Parfit's view, to necessarily support Reductionism.

This clearly leaves no middle ground. However, one might agree with Parfit and claim that there is no obvious answer to the question of who survives division. One might also agree that the identity relation is not (at least in some cases) the one that matters, and that there may be cases in which the identity relation does not hold, but where we can still know all we need to know about what matters in questions of survival. It is possible to make all these claims and still disagree with Parfit and assert that relation R is *not all that matters* -- that there is some other relation that has yet to be considered.

Parfit claims that the existence of persons involves nothing further than a series of mental states and their

psychological connectedness and continuity, or relation R. Since this is all there is to the existence of persons, it is this -- and only this -- that matters in questions of survival. In making this claim, the onus is on Parfit to demonstrate that relation R is indeed all that matters.

But is it true that relation R is all that matters in all cases? Consider, for instance, the prospect of being tortured. If I were to learn that someone exactly like me will be tortured tomorrow, the Reductionist view of personal identity tells me that the rational response to have would be:

"Will this person be R-related to my present experiences?"

As a Reductionist, the prospect of torture means no more than that there will be a given experience that will be related in a certain way to past experiences. This should make it less important to me that someone exactly like me will be tortured tomorrow. All that this fact tells me is that there will be an act of torture taking place tomorrow, and that the experience will be linked to my past and present mental states by way of relation R.

This experience could just as easily be linked to another set of experiences, and so be had by someone else. Parfit claims that what should be more important to me is

simply the fact that torture will be occurring. This should matter more to me than whether it will be me that will be tortured.[24]

The trouble is that when I consider the prospect of being tortured I find that I do care deeply about that experience as an experience that I will have. And I find that I continue to care deeply even after accepting Parfit's relation R. Why is this? Is it because I am actually a non-Reductionist, and so cannot accept the implications of relation R, or is it because there is something missing from relation R? I accept the Reductionist view, so I cannot be accused of holding non-Reductionist beliefs. Yet I am not convinced that relation R takes care of all that matters, as Parfit claims it does.

Let me examine more closely the question of my torture, which I will call the problem of pain. When faced with a choice of either having pain inflicted on me or not having pain inflicted on me, I will usually choose the latter. I have to admit this probably means that, if I were forced to choose between three possible victims of pain -- myself, victim X, or victim Y -- I would choose one of the other two. This of course depends to some extent on who the other two possible victims are. If I were told they were people I did not know, the implications would be more abstract, and it would certainly be an easier choice than if

they were people for whom I cared. Or I might decide that I would rather endure the pain than live with the knowledge of having chosen to inflict that pain on others. This might be especially true if the pain is not too horrible, such as a broken arm. The choice becomes harder when faced with the loss of a limb, or with death.

In any case, if I were to choose the pain for myself it would be because I anticipated some other, greater kind of suffering that I did not want to live with. It would not be because I thought of the pain as being R-related to me.

The point is, regardless of the choice I might make on moral grounds, I will always prefer not to experience pain myself. I am fully aware that others endure pain every day, yet this does not affect me nearly as much as half an hour of torture would affect me. This is obvious because there is a great difference between my experiencing something and my not experiencing it. The fact that a certain experience of torture will take place may matter to me, but the fact that that torture will happen *to me* will always be much more important to me than anything else.

Parfit claims that persons are nothing further than a series of mental states related in a certain way. This should have a positive effect on us, he says. It should make us care less about ourselves and more about others. This is particularly true of our future interests. Since

connectedness between my present mental state and my future mental state will diminish with time, it makes sense that I should care less about my future when such connectedness holds to a lesser degree. Parfit explains that this view can also make us care less about the present:

> Suppose...that I must undergo some ordeal. Instead of saying, 'The person suffering will be me', I should say, 'There will be suffering that will be related, in certain ways, to these present experiences'... The redescribed fact seems to me less bad.[25]

But when I consider the case of my torture I find that *however* I describe it it seems very bad. Nothing makes it better. And this is not because I believe that I have a soul that will be permanently harmed, or because I believe any other non-Reductionist view. It is simply because my being tortured seems very bad to me. It seems likely that Parfit, too, would forget about the psychological criterion in the event of his own torture, and would also find that no matter how he tried to redescribe the fact, it would be very bad for him -- far worse than the knowledge of torture experienced by someone else could ever be for him.

John Perry offers an explanation as to why we care so much about our own pain. Since pain is in itself undesirable, it is rational to hope that experiences of pain will be limited. But it is not rational to care more about our own pain, except in that such pain may interfere with the completion of one's aims, or projects. Thus, it would be rational for me to care about my being tortured tomorrow if that will mean that I will be unable to read a book that I planned to read. Parfit agrees with this view, but this strikes me as a bizarre explanation. When I imagine being tortured I do not expect my reaction to be: "What about all the reading I have to do?" Such thoughts would be far from immediate. Furthermore, what if I had no plans or projects at all? It is doubtful that this would make me care less about being tortured.

Perry, in defending this project-caring explanation, assumes that relation R is all that makes the future experience of torture important to me now. This view claims that what we actually care about is how a given experience will be related to our past experiences, and how it will affect our intentions concerning the near future. I claim that what is most important is not the fact that pain, or any other experience, will be R-related to me, but rather that *I* will be in pain. Parfit would answer this by explaining that my being subjected to pain is, simply, my

having a certain experience that is R-related to a particular set of physical and mental states. This is all that is involved in my experiencing pain, or any other kind of sensation. When I say that I will be in pain, I am therefore saying nothing further than that there will be an experience that will be related to me by way of relation R. Thus, all that matters is relation R.

Parfit's response is, to some extent, an appropriate one. It is fine to say that to have a given experience is to have a certain mental state that is related to past mental states. However, this does not fully explain the problem of pain. What we are trying to explain is what it is that makes one care deeply about the prospect of experiencing pain. This cannot be explained in terms of relation R, as Parfit claims it can. When I fear pain I do not fear the fact that the pain will be R-related to a particular set of (past) experiences. Were the pain not R-related to me, I would still fear it, and it would still matter deeply to me.

That relation R does not explain our deep caring about the prospect of pain can be shown by examining a case where relation R does not hold. Consider, for example, the case of the Korsakov Syndrome.[26] This is an actual disease in which victims are unable to store long-term memories. Memories hold over a very short period of time, such as three or four minute intervals, and are then lost. The victim does,

however, retain long-term memories stored prior to the occurrence of the disease.

For example, a forty-four year old woman suffering from the syndrome since the age of nineteen would be able to refer to all the memories of her childhood and adolescence that she had before her illness, but she would have no memories of the twenty-five year period between the immediate present and the age of nineteen. In this case, relation R holds strongly for the first nineteen years of her life but holds weakly between ages nineteen and forty-four. There is direct connectedness between each particular minute of experience and the next, and this overlapping chain provides continuity.

We can easily imagine another kind of illness that is only a slight modification of the Korsakov Syndrome. This is the case of the five-minute amnesiac. In this case, X has lost all her memories and intentions. This means that connectedness and continuity does not hold between her present mental state and her past mental states; she is no longer R-related to her past self. Under these conditions, we could not say that a present experience is related to her past experiences by way of relation R. Her present experiences are simply immediate experiences, or "now" experiences, with no relation holding between past and present ones.

To make this case even more clear, suppose that X's state of amnesia is such that her memory of present experiences is lost at five-minute intervals, and that the maximum degree of connectedness that can hold is within a five-minute period. This makes her existence very much of a "now" existence. Although there is strong direct connectedness between experiences within each five-minute period, connectedness does not hold between one five-minute period and the next.

Parfit claims that the amount that we care about a given future experience is directly related to the number of connections that hold between our present experiences and that future experience. In the case of the five-minute amnesiac there are no connections between two periods of experience: relation R does not hold in any significant sense. According to Parfit, this means that it should not matter to the amnesiac that she will be tortured in six minutes, given that there will be no connectedness between her present self and her self in six minutes.

Yet, even in this case, if X were told at minute one that her body would be tortured at minute six, I would argue that she would care deeply about this fact. I would argue that she would care just as deeply about this as would someone for whom relation R did hold with the ordinary number of connections. She would care deeply because,

although she would realize that she would have no memory of her present self at the moment of torture, she would also realize that it will be *her* experience of torture. She is the one who will be in pain, and she knows now what it means to be in pain. She also knows that when she will be tortured it will not matter whether or not she has memories of her past selves; what will matter most will be the pain itself, and the fact that she will be the victim of that pain.

This may seem irrational of the five-minute amnesiac, but consider another case that might better explain her response. Suppose that you were to hear, at this moment, that you will be woken up at three o'clock tomorrow morning and subjected to thirty minutes of excruciating torture. Just before you wake up, however, you will be given a pill that will put you in a state of complete amnesia, including a loss of all prior intentions and beliefs, for the thirty-minute period immediately following. When you wake up, prior to the torture, you will not remember anything at all about your past, and after the torture is over you will again lose all memory of that particular experience. The question is, do you care now about the fact that you will be tortured under those circumstances? Given that relation R will not hold, does the prospect of torture matter to you?

My claim is that it still matters deeply. It matters deeply to me because I know that I will be in pain, and this

matters regardless of the number of connections that are said to hold with respect to that particular experience. What I care about is the pain, not its relation to my past or to my future.

As Bernard Williams argues:

> Physical pain...is absolutely minimally dependent on character or belief. No amount of change in my character or my beliefs would seem to affect substantially the nastiness of tortures applied to me; correspondingly, no degree of predicted change in my character and beliefs can unseat the fear of torture which, together with those changes, is predicted for me.[27]

One might object to this position on the grounds that not all persons seem to care deeply about the prospect of pain. What, for example, of those who smoke, the objection might ask. Aren't smokers willingly engaging in something that will eventually lead to some sort of pain? The prospect of such pain does not deter smokers because it is pain that will be experienced well into the future -- at least that is what they're betting on. If and when I get sick, a smoker will say, I will be so old that it probably won't matter; anyway, I will be a different person then. This might be offered as evidence in favor of Parfit's claim that what we care about are only those experiences that are strongly

connected to our present experiences, and that we do not care about experiences that are not so connected.

The answer to this objection is that such rationalizations are not in fact based on the degree of connectedness that will hold between the present person and the future person, and so do not constitute a defense in favor of relation R. Rather, they are based on a typical but irrational tendency to expect that pain ten years hence will be less bad than pain that will be experienced tomorrow. However, the experience of pain itself will be equally bad, whether it is experienced now or in the future. In fact, it might be said to be worse to postpone pain, as then the victim must go through the agony of anticipation, and its realization. The other obvious answer to the smoker's objection is that smokers are not really calculating the degree of connectedness that is likely to hold at the time of illness -- they simply do not believe they will ever be sick from smoking.

The problem of pain, then, demostrates that relation R cannot account for the fact that we care deeply about the prospect of pain. Parfit cannot explain this fact in terms of relation R. What is important is not that pain will be R-related to me, but rather that I will be in pain. This, I argue, speaks against the significance of relation R. Relation R describes the way in which past experiences are

related to present ones, but it does not account for all that matters in a given experience.

The case of pain is significant in that it points to something that relation R does not account for with respect to our own experiences. What, though, of the way in which we view others? Can and does relation R account for all that we consider important about the identity of other persons?

One important consequence of Parfit's position is that our criterion for identifying others is limited to relation R. It follows from this that, if I were faced with a choice between person X whom I love, and a perfect replica of person X, the choice should be a matter of complete indifference to me. If person X has a replica, then relation R will hold between that person's mental states and the replica. All that matters about person X will therefore also be true of the replica. Parfit claims that since strong connectedness holds between the original and the replica, it would be irrational not to love the replica equally. Although the two are not numerically identical, qualitatively they are the same.

The trouble is that I am not convinced by this position. When I imagine the death of person X, I also imagine feeling grief. This grief is not alleviated with the prospect of a replica. Parfit would say that this is irrational, that the only rational grounds for such a

reaction would be a belief in a soul. I might argue that since the replica is not numerically the same as the original, it must be this difference that matters. Being only a copy, the replica is not as good as the original, for it is not the copy that I decided to love, but the original. My emotions cannot simply be transferred from one person to another, and this is what the case of the replica requires me to do.

Parfit's answer to this is that my choice not to love the replica as much as the original is based purely on sentimental grounds. This kind of sentiment, he says, has no rational basis; it is similar to the sort of attachment one might have to an object, such as a ring. If I were to lose a ring that has sentimental value to me, I might not feel the same about a perfect replica of that ring as I do about the original. While the two rings are qualitatively identical, I might prefer to have the original, and in fact I might decide that if I cannot have the original I will do without a ring at all. But this, Parfit says, is not rational. There is nothing *in* or *about* the original ring, there is no non-relational property that the replica ring does not also have. My sentiment is not anything that can be found in the object itself; instead, it is simply a choice that I am making about how I wish to treat a particular object.

In the same way, there is nothing about the original person X that is not also in his replica. If I do not wish to regard them as the same, this is a sentimental decision and is not based on any distinction to be found between the two persons themselves. If I were not such a sentimental person, then I would have no difficulty doing the rational thing and viewing the replica as being the same as the original.

Is this sentimental value all that the replica is missing? According to Thomas Nagel there is another critical distinction between the two and that is their brains. Nagel would argue that the case of the replica and the problem of pain both support his view of what matters in questions of identity, and that is the identity of the brain. Parfit's relation R cannot account for what it is that is missing in the replica and in the problem of pain because relation R does not account for what really matters. What matters, in such cases, is the identity relation, *not* relation R. And this identity relation matters with respect to the identity of the brain.

Nagel is also a Reductionist and agrees with Parfit that persons are simply their brains and their bodies, their thoughts, deeds, and experiences. However, he argues that Parfit over-emphasizes the importance of the relation among

these experiences, and in doing so overlooks what it is that is most important in the identity of persons.

According to Nagel, the word "person" refers to whatever it is that makes psychological continuity possible. As Nagel explains:

> What I *am* is whatever *in fact* makes it possible for the person TN to identify and reidentify himself and his mental states... If certain states and activities of my brain underlie the mental capacity, then that brain in those states...is what I am, and my survival of the destruction of my brain is not conceivable.[28]

What really matters, Nagel argues, is the physical continuity of the brain, since without this there would be no continuity of mental states. With respect to the replica, Nagel would say that it is rational to regard the replica as not being as good as the original. What fundamentally matters for the continued existence of a person is the physical continuity of that person's brain, not psychological continuity. In the case of the replica, the brain of the person that I love does not continue to exist.

This means that what fundamentally matters about that person no longer exists. Thus, I am justified in feeling that there is something missing in the replica: the replica is only a copy -- it does not have the same brain as the person I love.

If Nagel is right and it is the continued existence of the brain that matters, then it is more understandable that we should be deeply concerned about pain. When faced with the prospect of pain, it is clear that this will be a particular experience had by a particular individual. The subject of that experience will be that person's brain. It is this fact that is of deep concern, not relation R. What matters, on this view, is <u>not</u> that experience P will be had and that it will be related to experience E, but rather that experience P will be had by a particular brain.

I agree with Parfit that this position is problematic. The argument that the physical continuity of the brain determines personal identity still leaves many aspects of the problem unresolved. However, in many ways this criterion seems no more arbitrary than Parfit's conclusion that personal identity involves nothing further than relation R.

2. Conclusion

Parfit's position is that the problem of personal identity has remained unresolved because it has mistakenly concerned itself with the wrong relation. What matters in questions of the continued existence of persons through time is not the identity relation; rather, it is relation R. Parfit defends this claim through a series of arguments in favor of the Reductionist view of personal identity. Reductionism, he argues, logically leads to his conclusion that the identity relation must be discarded in questions of survival. Further arguments demonstrate that the relevant element in such questions is actually relation R. It is through this line of reasoning that Parfit offers a solution to the problem of determining the identity of persons through time.

My claim is that Parfit does not fully succeed in providing a solution. I have made this claim on two grounds:

First, Parfit's arguments in defense of Reductionism do not in themselves support that view. Each of the four arguments presented by Parfit do not inevitably lead to the conclusions he has drawn and which he claims necessarily follow. It is just as logical to take these same cases and draw conclusions that are inconsistent with the position Parfit defends. Parfit's own arguments are themselves

subject to many of the same criticisms he launches against the identity relation. This suggests that Parfit's conclusions rely, to at least some extent, on arbitrary interpretations of his own cases.

Secondly, Parfit's claim that relation R is *all that matters* is also open to question. Parfit has not conclusively demonstrated that there is indeed one relation that takes care of all the relevant concerns we have regarding survival, *in all cases*, nor has he shown that the relation must necessarily be relation R.

FOOTNOTES

[1] Hume, David. A Treatise of Human Nature, Book I, Part IV, sects. 1,5,6, "Of Personal Identity."

2 This general formulation of Parfit's argument calls for a qualification. Identity involves something further than relation R when R holds 'uniquely', taking a 'non-branching' form. According to Parfit, the case of division demonstrates that this qualification is minor.

[3] Parfit, Derek. Reasons and Persons, Oxford: Oxford University Press, 1986, page 211.

[4] Ibid., page 211.

[5] Ibid., page 212.

[6] Ibid., page 212.

[7] Ibid., page 220.

8 This refers to Butler's claim that Locke's memory criterion of personal identity leads to a circular argument.

[9] Parfit, Reasons and Persons, page 206.

[10] Ibid., page 199-217.

[11] Ibid., page 221.

[12] Ibid., page 220.

[13] Ibid., pages 245-246.

[14] Ibid., page 249.

[15] Ibid., page 247.

[16] Ibid., page 249.

[17] Ibid., page 245.

[18] Ibid., page 251.

[19] Ibid., page 250.

[20] Ibid., page 256.

[21] Ibid.

[22] Ibid., page 231.

[23] Ibid., page 239.

[24] Ibid., pages 281-282.

[25] Ibid., pages 281-282.

[26] This case was brought to my attention by Ian Gold.

[27] Williams, Bernard. Problems of the Self, Cambridge: Cambridge University Press, 1973. Page 54.

[28] Nagel, Thomas. From Parfit, Reasons and Persons, page 469.

BIBLIOGRAPHY

Ayer, A.J. Language Truth and Logic. New York: Dover, 1946. "The Self and the Common World," Ch. 7.

Ayer, A.J. "Individuals," Mind. 1952.

Ayer, A.J. The Concept of a Person and Other Essays. London: Macmillan, 1964.

Butler, Joseph Bishop. "Of Personal Identity." The Works of Bishop Butler. J.H. Bernard, Ed. London: 1900.

Dennett, Daniel. Brainstorms. Bradford Books, 1978.

Descartes, Rene. Meditations. translated by E.S. Haldane and G.R.T. Ross, Cambridge University Press, 1969.

Flew, Anthony. "Locke and the Problem of Personal Identity," Philosophy. Volume 26, 1951.

Flew, Anthony, ed. Body, Mind and Death. New York: Macmillan, 1964.

Glover, Jonathan, ed. The Philosophy of Mind. 1976.

Hare, R.M. Freedom and Reason. Oxford University Press, 1963.

Hare, R.M. Moral Thinking. Oxford: Clarendon Press, 1981.

Hume, David. A Treatise of Human Nature. edited by L.A. Selby-Bigge, Oxford, 1896.

Kant, Immanuel. Critique of Pure Reason. Norman Kemp Smith, translator. London: Macmillan Press, 1929.

Leibnitz, G.W. New Essays Concerning Human Understanding. Langley, translator. La Salle, Ill.: Open Court, 1949.

Locke, John. Essay Concerning Human Understanding. edited by Campbell Fraser, Oxford, 1894.

Mackie, J.L. Ethics. Harmondsworth: Penguin Books, 1977.

Nagel, Thomas. "Brain Bisection and Unity of Consciousness," Synthese. 1971.

Parfit, Derek. "On 'The Importance of Self-Identity.'" Journal of Philosophy. LXVIII, 20, October, 1971.

Parfit, Derek. "Personal Identity," Philosophical Review. LXXX, I, January, 1971.

Parfit, Derek. Reasons and Persons. Oxford: Oxford University Press, 1986.

Penelhum, Terence. "Hume on Personal Identity," Philosophical Review. Volume 64, 1955.

Penelhum, Terence. "Personal Identity," The Encyclopedia of Philosophy. New York, London: Macmillan, 1967. Paul Edwards, ed. Volumes 5 and 6, pp.95-107.

Perry, John, ed. Personal Identity. Berkeley, Los Angeles, London: University of California Press, 1975.

Reid, T. Essays on the Intellectual Powers of Man. "Of Memory," Edinburgh, 1785. Chs. 4 and 6.

Ryle, Gilbert. The Concept of the Mind. London: 1949.

Ryle, Gilbert. Dilemmas. Cambridge: Cambridge University Press, 1954.

Rorty, Amelie Oksenberg, ed. The Identities of Persons. Berkeley, Los Angeles, London: University of California Press, 1976.

Searle, John. Minds, Brains and Science. Cambrdige, MA: Harvard University Press, 1984.

Shoemaker, Sydney. "Personal Identity and Memory," Journal of Philosophy. Vol. 56, 1959.

Shoemaker, Sydney. "Persons and their Pasts," American Philosophical Quarterly. 7, 1970.

Strawson, P.F. The Bounds of Sense. London: Methuen, 1966.

Swinburne, R.G. "Personal Identity," Proceedings of the Aristotelian Society. 74, 1973-4.

Unger, P. "Why there are no People," Midwest Studies in Philosophy. 4, 1979.

Wiggens, D. Sameness and Substance. Oxford: Basil Blackwell, 1980.

Williams, Bernard. "Personal Identity and Individuation," PAS. Vol. 57, 1956-57.

Williams, Bernard. Problems of the Self. Cambridge: Cambridge University Press, 1973.