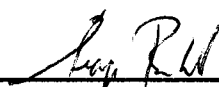


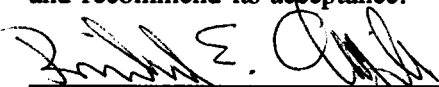
To the Graduate Council:

I am submitting herewith a dissertation written by Thomas G. Buller entitled "Personal Identity and Advance Directives." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Philosophy.

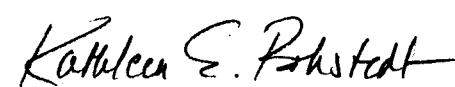


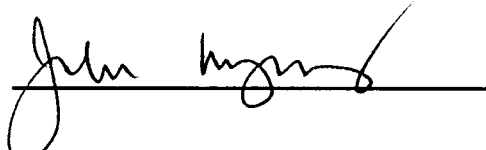
Dr. George Brenkert, Major Professor

We have read this dissertation
and recommend its acceptance:









Accepted for the Council:



Associate Vice Chancellor and
Dean of The Graduate School

PERSONAL IDENTITY AND ADVANCE DIRECTIVES

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Thomas Gilbert Buller

August, 1994

To Kerri

ACKNOWLEDGMENTS

There are many people to whom I am grateful for making my time at the University of Tennessee so rewarding. I have benefitted greatly from knowing the faculty and graduate students in the Department of Philosophy over the last six years, and hope that these relationships will endure. I am particularly grateful to my Dissertation Committee, George Brenkert, Richard Aquila, Kathy Bohstedt, John Dougherty and Michael Lavin for their support and encouragement. They have all read a number of revisions of the dissertation and the merit of this work is due in no small part to their suggestions.

There are a number of other people whose assistance should be recognized. I would like to thank Anne Beardsley and Marie Horton for their help in grim logistical matters and their friendship. I have received generous financial support from the Department of Philosophy, the Graduate School and the Physicians' Medical Educational and Research Foundation.

Lastly, the greatest debt is owed my wife, Kerri. Without her help and encouragement I would never have reached this moment.

ABSTRACT

Treatment decisions for incompetent patients are frequently based on the patient's prior expressed wishes when competent. The merit of this approach is that it is consistent with a respect for patient self-determination and extends the rights of the competent patient to cover incompetence. However, this approach has been challenged on the grounds that the neurological deficit or dementia that has rendered the patient incompetent has also sufficiently impaired or destroyed the patient's psychological states so as to destroy the psychological characteristics that constitute personal identity. Hence the incompetent patient is not the same person as the prior competent person and therefore has no authority to make treatment decisions for the incompetent patient.

In this work I attempt to resolve the objection to this approach to decision-making for incompetent patients stated above. I examine the use of advance directives as a means by which the competent person can make future treatment decisions, and discuss the different approaches to decision-making for incompetent patients described in recent legal cases. This work includes an extensive discussion of the concepts of person and personal identity in order to determine whether the incompetent patient is a person, and whether he or she is the same person as the prior competent person. I examine the connection between the concept of person and the possession of rights, and the necessary conditions of personhood and personal identity. A number of rival accounts of personhood and personal identity are discussed, and I argue that we can present plausible descriptions of these concepts that will include the incompetent patient in all but the most severe cases of neurological deficit. This work investigates the distinctions between the identity and survival relations, and attempt to determine whether it is possible for you or I to survive the loss of

personhood. We discuss the conditions of survival and whether you or I are essentially persons. I conclude that we can describe the incompetent patient in such terms that he or she can be regarded as a survivor of the competent person even though the incompetent patient may no longer be a person, and hence the moral authority of advance directive is upheld.

TABLE OF CONTENTS

CHAPTER

INTRODUCTION	1
1. DECISION-MAKING FOR INCOMPETENT PATIENTS	6
Competence and Incompetence	11
The Rights of Incompetent Patients	16
Three Different Approaches to Decision-Making for Incompetents	27
The Limitations of the Three Approaches to Decision-Making	35
2. THE CONCEPT OF PERSON	44
The Concept of Person	45
The Conditions of Personhood	49
Personhood and Incompetence	99
3. PERSONS AND PERSONAL IDENTITY	108
The Identity Question	108
Personal Identity	114
Persons and Personal Identity	166
4. A MATTER OF SURVIVAL	169
Identity and Survival	171
The Conditions of Survival	180
Sortal Concepts and Relative Identity	202
Relative Identity and the Loss of Personhood	208

CONCLUSION	223
SELECTED BIBLIOGRAPHY	228
VITA	238

INTRODUCTION

Mr. H, a former concert flautist, is sixty-three years old and suffers from cerebral arteriosclerosis. As a result of this condition Mr. has become increasingly agitated and depressed, and already shows symptoms of increasing dementia. He has recently been admitted to University Hospital suffering from pneumonia, and has been intubated. Mr. H has a Living Will which he executed seven years ago. He requested that he not be kept alive by artificial means if he should become incompetent. However, Mr. H's family have requested that the Living Will be ignored on the grounds that since the time of its execution, Mr. H has changed his mind. As evidence in support of this contention, the family point to Mr. H's apparent enjoyment in being read to, that from time to time he seems to take pleasure in being with his family, and his now obvious dislike of music.

The case of Mr. H raises a number of questions concerning the continuation of life-sustaining treatment for incompetent patients. An initial question is to ask how we should make such decisions. Since the *Quinlan* case eighteen years ago the courts have shown a tendency to regard the previously expressed wishes of the patient when he or she was competent to be influential in decisions to continue or cease treatment, including life-sustaining treatment.¹ This approach is referred to as the "subjective" approach on the grounds that the particular patients subjective wishes are influential in treatment decisions. This is distinguished from the "objective" best interest approach where treatment decisions are made with regard to what any "reasonable person" would want. More recently, the *Cruzan* case established a *de facto* right of a person to make treatment decisions should he or she become incompetent by executing an advance directive.² The most well-known version of such a directive is a "Living Will." In this legally-binding document a competent

¹*In re Quinlan*, 70 N.J. 10, 355 A.2d 647 (1976).

²*Cruzan v. Harmon*, 760 S.W.2d 408 (Mo. 1988); *Cruzan v. Director, Missouri Dept. of Health*, 1990 US Lexis 3301 (US June 25, 1990).

person can state the conditions under which he or she requests treatment to be performed. The competent person can request that all life-sustaining treatment be ceased including nutrition and hydration.

The purpose of this work is to attempt to resolve a challenge to the subjective approach to decision-making for incompetent patients. The subjective approach rests on the assumption that the prior competent person and the later incompetent patient are the same person, for if this were not the case, then *prima facie*, the competent person would have no authority to make treatment decisions for the incompetent patient. However, this assumption is regarded by some as questionable. For it has been maintained that the conditions that have rendered the patient incompetent (for example, the patient may be suffering from Alzheimer's dementia) may have also sufficiently impaired or even destroyed the patient's psychological characteristics so as to make it unclear whether we can regard the incompetent patient and the previously competent person as the same person. Thus in the case of Mr. H above we may wonder whether the cerebral arteriosclerosis has caused sufficient neurological damage so as to destroy his psychological characteristics. That is to say, whatever it was that constituted the identity of Mr. H, this has now been lost.

If this objection is successful, then a decision to continue or cease treatment for incompetent patient should be based strictly on contemporaneous best interest considerations, and should ignore the prior wishes of the competent person. Perhaps this is a viable alternative, but it would seem to disregard respect for patient self-determination since it would deny the competent person authority to make treatment decisions should he or she become incompetent.

This objection to the subjective approach I term the "Identity Question." The success of this objection would appear to rest on two prior claims: First, that the previously

competent person and the incompetent patient are not the same person; second, that the incompetent patient is a person. The second claim would appear to be important for if it could be successfully argued that the incompetent patient is not a person, then this patient cannot be the same or a different person, hence the identity question becomes moot.

Thus one substantial question that has to be addressed is whether the incompetent patient is a person. We can put this another way by asking whether there is any account of personhood that could include cases of incompetence, even such cases as persistent vegetative state. If such an account can be described, then the remaining task is to determine whether the conditions of identity have been met in the relevant cases.

However, one can ask whether it is necessary that the incompetent patient is a person? If our concern is to establish whether the incompetent patient is the same or a different person, then quite clearly it is necessary that the incompetent patient is a person. But this concern assumes that the relevant identity relation between the competent and incompetent individual is *personal identity*. Could it not be argued that there are other identity relations that could describe the competent and incompetent individuals? One suggestion is that although the neurological deficit or dementia is sufficient for the loss of personhood, the incompetent patient may still be the same *human being* as the competent person, and still be the subject of self-interested concerns. To take the example of Mr. H, it might be maintained that although the cerebral arteriosclerosis has caused a loss of personhood, Mr. H may still have interests, concerns and may still be someone whose welfare the prior competent person would be interested in.

We can put this matter a different way by asking whether you or I are essentially persons. If we are essentially persons then we would not be able to survive the loss of personhood, since being a person is necessary for our survival. On the other hand, if we are

not essentially persons, then we could survive the loss of personhood. That is to say, our identity and our self-interested concerns are not dependant on our status as persons.

In this work I discuss all of these considerations. The dissertation is divided into four chapters. The focus of the first chapter is on the different approaches to decision-making for incompetent patients. This involves discussion of the notions of competence and incompetence, and whether the incompetent patient retains the right to direct medical treatment. This leads to a discussion of recent legal cases that have dealt with the matter of incompetent patients and a right to direct medical treatment, and the identification of three approaches that have been used to make treatment decision for such patients. We discuss the merits of each of the three approaches, paying particular attention to advance directives. The chapter concludes with an examination of a recent attempted solution to the identity question.

The second chapter examines the conditions of personhood. We begin by distinguishing between two different senses of *person*, and then proceed to a detailed examination of three different concepts of person. One putative condition of personhood is discussed at length -- that to be a person the individual must be capable of self-interested concerns. The chapter concludes by looking at whether these different accounts would regard the incompetent patient as a person.

In the third chapter we turn to the matter of personal identity. Our initial concern is to describe the type of challenge that cases of neurological deficit present to our traditional theories of personal identity. We then turn to a discussion of two main rival views of personal identity, the *reductionist* and *non-reductionist* views, and the analysis of these views with regard to a number of objections that have been raised against them. We discuss different versions of these views, paying particular attention to the question of the unification

of diverse mental states. The chapter concludes with an examination of a functionalist account of personal identity.

In the fourth and final chapter we turn from the question of personal identity to the matter of survival, and examine the question whether you or I can survive the loss of personhood. An answer to this question is dependant on establishing whether you or I are essentially persons. We discuss the notions of strict and relative identity, and the distinctions between the identity and survival relations. The chapter concludes with an examination of three different accounts that address the question whether you or I could survive the loss of personhood. The chapter concludes with the identification of two ways in which the moral authority of advance directives, and of the subjective approach in general, can be maintained.

CHAPTER ONE

DECISION-MAKING FOR INCOMPETENT PATIENTS

There is a general recognition, both in law and in practice, that the competent patient has an almost absolute right to refuse medical treatment, including life-sustaining treatment.¹

As Justice Cardozo said

Every human being of adult years and sound mind has a right to determine what shall be done with his own body.²

The right to refuse medical treatment is founded on our notions of autonomy and self-determination, and constitutional claims of privacy and liberty. The practical application of this right to refuse treatment is the doctrine of informed consent, that in order for treatment to proceed the patient's valid informed consent must be obtained.

We can see the connection between competence and informed consent more specifically by identifying the values underlying informed consent. Following Buchanan and Brock, two central values are: i) that informed consent serves to promote individual well-being and ii) that it respects the patient's self-determination.³ The value of promoting individual well-being comes from the fact that each of us has our own subjective aims and values that make up who we are and are integral to the health care decisions that we make. As

¹There are a number of instances when the courts have ordered treatment for competent patients. See Rebecca Dresser, "Advance Directives, Self-Determination, and Personal Identity," in *Advance Directives in Medicine*, ed. Chris Hackler, Ray Moseley, and Dorothy E. Vawter, (New York: Praeger, 1989), 156.

²*Schloendorff v. New York Hospital*, 211 N.Y. 125, 127, 129; 105 N.E., 92, 93 (1914).

³Allen E. Buchanan and Dan W. Brock, *Deciding for Others: The Ethics of Surrogate Decision Making* (Cambridge: Cambridge University Press, 1989), 29.

Buchanan and Brock state

As an expression of the ideal of shared decision-making, the doctrine of informed consent recognizes that while the physician commonly brings to the physician-patient encounter medical knowledge, training, and experience that the patient lacks, the patient brings knowledge that the physician lacks: knowledge of his or her particular subjective aims and values that are likely to be affected by whatever decision is made.⁴

Thus the importance of informed consent is that it respects the patient's own subjective values, even if the physician (or whoever) disagrees with them. The current opinion is that if the patient is competent he or she has a right to refuse or direct treatment based on his or her own subjective values even if others might disagree with such a refusal.

Patient self-determination is valued both instrumentally as well as for its own sake. It is valued instrumentally in that, in the normal course of events, we are in the best position to determine what is in our own best interest. It is valued for its own sake in that there is value in being able to make health care decisions for ourselves, even if we believe others are more competent to make these decisions.

But what approach should be taken if the patient is incompetent? On what grounds should we decide whether it is morally permissible to cease life-sustaining treatment? *Prima facie*, we cannot appeal to notions of self-determination and individual well-being since the condition that renders the patient incompetent directly affects these two underlying values. To say that the patient is incompetent to refuse life-sustaining treatment is to say that she is not able to determine what is in her own best interest. That is to say objective determinations of what is in the patient's best interest take precedence over the patient's

⁴Ibid., 30.

subjective aims and values.⁵ This would suggest that decisions to withdraw or withhold life-sustaining treatment should be made by others better able to determine what is in the incompetent patient's best interest.

However, there are shortcomings to this "best interest" approach. First, the success of this approach rests on the assumption that physicians, family or others are able to determine what is in the patient's best interest. Furthermore, in order to make clear and uncontroversial determinations of what is in another person's best interest, we need to be able to accurately assess the "quality of life" of another person.⁶ This is particularly important when we are concerned with the decision to cease life-sustaining treatment. However, determinations of the quality of life of another are very difficult, particularly so when we are attempting to make these determinations on the basis of objective considerations of what a "reasonable person" would want in conjunction with our knowledge of the wishes of this particular patient. For we might be skeptical not only of the idea of a "reasonable person" but also of our ability to determine what such a person would want.⁷ Furthermore, the incompetent patient's own interests may run counter to our objective notions of best interest. Although the fact of the patient's incompetence undermines the authority of the patient's own values it does not make them irrelevant. Thus although objective notions of best interest

⁵To term these determinations "objective" is not to accord them any particular respect or neutrality, but rather that these determinations are made by others apart from the patient on the basis of what a "reasonable person" would want.

⁶See John A. Robertson, "Involuntary Euthanasia of Defective Newborns: A Legal Analysis," *Stanford Law Review* 27 (1975): 246-261.

⁷Although existing prior to these cases, the term "reasonable person" became the most frequently used standard for determining how much information should be disclosed to the patient following *Canterbury v. Spence*, 464 F.2d 772 (1972), and *Cobbs v. Grant*, 502 P.2d 1 (1972). Subsequent to these decisions, the "reasonable person" standard has been used more broadly to refer to the interests that would belong in the normal course of events to a relevant party.

take precedence over the patient's own values (this will be most clearly seen where there is conflict between the objective and subjective values) we should, nevertheless, pay attention to the patient's values in making treatment decisions. For example, the incompetent patient may value watching television or going for walks, and this information should be included in treatment decisions if the result of the treatment would be to deprive the patient of these opportunities. Conversely, by paying attention to these values we might be able to better care for the patient. Another way of saying this is to say that best interest determinations have to pay some attention to what this *particular* person would want, for an illness that might be unduly burdensome to one patient may not be for another.

Second, the best interest standard focuses only on the incompetent patient's present condition and present interests, rather than taking into account any interests or desires that the patient might have expressed while competent. Behind the best-interest approach is the idea that the beginning of incompetence clearly demarcates the beginning of the end of patient self-determination. And this idea is based on the view that since the incompetent patient is no longer able to understand what is in her own best interest, so others have to make this determination.⁸

The alternative is to make decisions regarding the cessation of life-sustaining treatment on the basis of the expressed interests of the person while competent, what Ronald Dworkin terms *precedent autonomy* and Rebecca Dresser terms *anticipatory choice*.⁹ This

⁸It should be pointed out that "incompetence" is a legal, not a medical term. Determinations of incompetence are made by courts on the basis of the person's cognitive abilities.

⁹Ronald Dworkin, "Autonomy and the Demented Self," *Millbank Quarterly* 64 supplement 2, (1986) 4-16; Rebecca S. Dresser, "Autonomy Revisited: The Limits of Anticipatory Choices," in *Dementia and Aging: Ethics, Values and Policy Choices*, ed. Robert H. Binstock, Stephen G. Post, and Peter J. Whitehouse, (Baltimore: The Johns Hopkins University Press, 1992) 71-86.

approach expands the notion of patient self-determination so that it is retained, despite the loss of competence. The consideration behind this approach is that the beginning of incompetence does not point the end of the person's ability to make decisions regarding life-sustaining treatment, for the person's expressed interests made while competent enable a decision to be made. Thus the best interest approach should only be used as a default when there is no evidence of the previously competent person's wishes.

But there are problems with this approach too. As Dresser argues, this approach rests on two assumptions: First, that the competent's person's notion of well-being applies equally well to the incompetent patient. Second, that the incompetent patient is the one who previously expressed those interests while competent. The first of these assumptions is questionable since what might seem unduly burdensome from the perspective of the competent person, might not do so later on when incompetent. The second of these assumptions is questionable if the properties that constitute the identity of the competent person are missing when the patient is incompetent.

The purpose of this chapter is to examine these different approaches to decision-making for incompetent patients. The structure of the chapter will be as follows. First, we begin by examining the notions of competence and incompetence. Second we look at the question of whether the incompetent patient retains the right to refuse life-sustaining treatment, or the right not to receive life-sustaining treatment that he or she possessed while competent. Third, we examine the three main approaches to decision-making and examine their merits, and look at recent cases and the different judgements that were reached. Fourth, we discuss limitations to these three approaches. Fifth, we look briefly at the role personhood plays in questions of decision-making for incompetent patients.

Competence and Incompetence

There is a certain imbalance in our notions of competence and incompetence. As Buchanan and Brock argue, "...the statement that a particular individual is (or is not) competent is incomplete."¹⁰ Rather, competence is decision-relative; a person is competent to perform a specific task, to consent to a specific treatment. An implication of this view is that even if we might question the person's competence to perform one action, this need not suggest that the ability to perform another is in doubt. Furthermore, there is a tendency to regard competency as a "threshold concept". That is to say that, for a particular decision, if the person meets a certain threshold level of cognitive ability then her ability to make that decision should not be questioned. This approach is in opposition to the treatment of competence as something that can be possessed in degrees. On this latter approach, we would talk of a person being "more" or "less" competent than is required, or of being "insufficiently" competent. But this latter approach is rejected since, according to Buchanan and Brock, it threatens the *function* of the competence determination.

That function is, first and foremost, to sort persons into two classes: (1) those whose voluntary decisions...must be respected by others and accepted as binding, and (2) those whose decisions, even if uncoerced, will be set aside and for whom others will act as surrogate decision-makers.¹¹

If we abandon our support for the threshold approach and regard competence as a comparative notion, then we would be obscuring the distinction between those whose decisions should be respected and those whose decisions should not. For we would not be able to base this distinction on the distinction between competent and incompetent individuals.

¹⁰Buchanan and Brock, *Deciding for Others*, 18.

¹¹Ibid., 27.

A further point is that the presumption is in favor of competence; that is, we presume the person to be competent unless and until there is a formal ruling of incompetence. This applies to all of the person's actions and thus we speak of a "global presumption" of competence. This might be thought to be at odds with the prior notion that competence is decision-relative but it is not. To say that competence is decision-relative is to say that the statement "x is competent" is incomplete unless we specify the action that x is competent to perform, and that the person's competence to perform one task need not imply his or her competence to perform a different task. However, to say that there is a global presumption of competence is to say that x is presumed to be competent to perform an action or make a decision unless and until there is reason to think otherwise.

In this regard, we can see a parallel between competence and "innocence." The statement "x is innocent" is incomplete - one is innocent *of* doing some particular act. Second, a person is not more or less innocent - the defendant is either innocent or guilty. Third, even if we question a person's innocence with regard to one act, this need not imply that we should question their innocence across the board. Fourth, there is a presumption of innocence - one is presumed to be innocent until proven guilty. The same points can be made with regard to competence: To say that a person is competent to make a particular decision need not imply that she is competent to make every decision. One is either competent or incompetent to make a particular-decision -- there are no degrees of competence. We presume the person to be competent unless and until there is evidence to the contrary.

The matter is different with incompetence. On the one hand, incompetence, like competence, is decision relative; the fact that a person is incompetent to perform one action (for example, Cardiopulmonary Resuscitation) does not mean that he or she is incompetent

to perform another. On the other hand, incompetence, unlike competence, is a legal determination (a court has to declare a person incompetent) and this determination is based on the person's inability not just to perform a range of actions but also actions that we would regard as basic for the person's well-being. Thus incompetence tends to be a more permissive notion - we question the person's ability to make *any* decision, whether it be trivial or important, even if there is reason to think that the person is competent to perform a particular action.

As Roth et al. point out, the concept of competency "...like the concept of dangerousness, is social and legal and not merely psychiatric or medical."¹² There is no explicit legal, social or medical notion of competency. In fact, it is incompetence that is the explicit legal and social term; a court has to make a ruling on incompetence. Nevertheless, to determine whether a person is competent, we can identify a number of features that are basic. As identified by Roth et al. there are a number of different tests for competency: 1) evidencing a choice, 2) "reasonable" outcome of choice, 3) choice based on "rational" reasons, 4) ability to understand, and 5) actual understanding.¹³ (These tests are progressively more complex; one needs a higher degree of cognitive ability to "actually understand" the relevant information than simply to make a choice, for the person may base her decision on irrelevant information). If the person can make a choice, or actually understand, then according to the different standards she would be thought to be competent.

Buchanan and Brock identify the following three conditions as necessary for decision-making capacity: (1) understanding and communication; (2) relatively stable values or a

¹²Loren H. Roth, Alan Meisel, and Charles W. Lidz, "Tests of Competency to Consent to Treatment," *The American Journal of Psychiatry* 134 (1977): 279-284, reprinted in *Psychiatry and Ethics*, ed. Rem B. Edwards, (Buffalo: Prometheus Books, 1982), 201-211.

¹³Edwards, *Psychiatry and Ethics*, 201-209.

conception of the good life; (3) reasoning and deliberation.¹⁴ It is worth pointing out the differences between Buchanan and Brock's account and that of Roth et al. The second of Buchanan and Brock's conditions is important and relates to our previous discussion of individual well-being. For Buchanan and Brock, the person's own subjective aims and values must be stable and be consistent with any decision made. Thus we might have a *prima facie* reason to question a person's decision-making capacity if his subjective aims and values were unstable or inconsistent. For example, we might question a person's decision to forego a life-sustaining blood transfusion on the grounds such treatment is prohibited by the Bible if the person had previously been an avowed agnostic. The tests cited by Roth et al. do not concern themselves with the individual's own subjective aims and values. For two of the tests (2 and 3), the values are those of a "reasonable person", and for none of the tests is it necessary the person have a "relatively stable conception of the good".

Buchanan and Brock support what they term a "process standard" of decision-making, that is they are primarily concerned with how the person reaches the decision based on his or her own values. This is to be contrasted with an "outcome standard" (test 2) and with a standard whereby the decision should be based on a set of objective values (test 3). Relatedly, Buchanan and Brock endorse the idea that the more serious the decision, so the more sophisticated the reasoning and understanding. That is to say that treatment choices where the condition is life-threatening require a higher level of decision-making competence than those conditions that are less serious. This is consistent with the view that competence is decision-specific.

It might be thought that to claim that some treatment choices require a higher level of decision-making competence than others is to claim, in effect, that there are degrees of

¹⁴Buchanan and Brock, *Deciding for Others*, 23-29.

competence, a claim that was previously denied.¹⁵ However, a distinction needs to be drawn between claiming the competence is decision-specific and claiming that there are degrees of competence. Consider, for example, a patient who refuses treatment for a ruptured appendix. In such a case it would be appropriate to demand that the patient fully understands the implications of his or her decision, for given the seriousness of the consequences, this is not a decision that should be made without reflection. Thus in this case we would be demanding a high-level of decision-making competence. Conversely, a patient's decision to forego pain medication after suffering a bone fracture requires a low-level of competence, since the consequences of the decision carry little risk of harm. However in neither case would it be appropriate to describe the patient as "more" or "less" competent, rather the patient is competent to make this decision or she is not.¹⁶

A person is deemed to be incompetent if he or she cannot meet any of the previous tests of competency. There is considerable debate as to the merits of these tests, as well as the more formal tests of competency. Certainly, one would not want to claim that there is a clear, fixed notion of competency, and that our notions of competency and incompetency do not reflect some rather vague notions of "reasonableness" and "rationality". Furthermore, one needs to further examine the implications of the distinction between competence and incompetence. From the fact that a person is found to be incompetent it does not follow that this person does not present claims or obligations against others. Thus someone suffering from Alzheimer's and deemed to be incompetent is not, therefore, at the whim and the suffrage of others. For this person still has strong enough interests (for example, in not being

¹⁵See p. 6 above.

¹⁶For further discussion of these issues see James F. Drane, "Competency to Give an Informed Consent: A Model for Making Clinical Assessments, *JAMA: Journal of the American Medical Association*, (August 17, 1984, vol. 252): 925-927.

harmed) that can present claims against others.

The Rights of Incompetent Patients

To recall, the general focus of this chapter (and of the entire work) is to examine how we should make treatment decisions for incompetent patients. In this section I want to examine the question "Does the incompetent patient have a right to refuse life-sustaining treatment?" This question is important for if it can be successfully argued that the incompetent patient does have this right, and that this right is analogous to the right of the competent patient to refuse treatment, then decisions to withhold or withdraw life-sustaining treatment from the incompetent patient would be easier in those cases where we have a clear idea that this is what the patient wants. Conversely, if the incompetent patient does not have a right to refuse life-sustaining treatment, the patient's interest in the withholding or withdrawing of such treatment must be balanced against the interests of others, for example, the State's interest in protecting the incompetent patient from harm.

If we think that the incompetent patient has a right to refuse life-sustaining treatment, then we need to examine further what is the basis for this right and how this right is exercised. To this end, it is useful to distinguish between two different rights that the incompetent patient might possess; the right to refuse life-sustaining treatment and the right not to receive life-sustaining treatment. The distinction between these two rights parallels the distinction between positive rights (a right to do x) and negative rights (a right not to have x done to you). On this analysis the right to refuse treatment is a positive right and the right not to receive treatment a negative right.

The distinction between these two rights would seem to be useful for it does not seem plausible to say that the incompetent patient can actually refuse treatment. If we think that

a necessary condition of a person having a right to x is that the person is able to exercise that right, then the incompetent person cannot be said to possess a right to anything. However, these same considerations need not rule out the incompetent patient's right not to receive life-sustaining treatment. For if the above analysis of positive and negative rights is correct, one can only exercise positive rights since a negative does not entail any action by the possessor of the right.

However, we need to be clear as to whether being able to exercise a right is a necessary condition of having a right, and what it means to exercise a right. If in order to exercise a right the person must be able to understand what the right entails, and what it claims with respect to the actions of others, then neither of the above rights could be claimed by the incompetent patient. For such a patient is, *prima facie*, incapable of such understanding. And even if it could be claimed that the patient has some understanding of the matter (for example, a patient who is moderately demented) we would probably be wary of allowing the person to exercise this right on the grounds that the refusal of life-sustaining treatment is a decision that should only be made with full understanding.

Alternatively, we might think that being able to exercise the right is not a necessary condition of having a right. There are two reasons for thinking this: First, it is not clear what it means to exercise some rights. For example, how does a person exercise the "right to life"? (and this is a right that is often claimed as the paradigmatic positive right). Second, some rights can be exercised on a person's behalf by another. For example, children have the right to initiate legal proceedings although incapable of exercising that right themselves.

There is a further complication here in that we are discussing the rights of the incompetent patient that were uncontroversially held by the prior competent person. Thus an alternative question to the one that we began with would be to ask "What rights (if any)

does the incompetent person retain?" For it is certainly not the case that all rights held by competent patients continue if the patients should become incompetent; the incompetent patient is not permitted to vote in elections, run for office or engage in financial transactions. Thus if the incompetent person does retain the right to refuse life-sustaining treatment or the right not to receive life-sustaining treatment, we need to explain why this is so for we have to determine what sort of entities can have rights, and what feature or property it is that establishes this right.

For example, Alan White examines the necessary and sufficient conditions for being capable of having a right.¹⁷ A number of candidates are presented as conditions for having a right: having an interest; being beneficiary of a duty (or obligation); and being a person.

There is considerable support for the position that the possession of rights is tied to the capacity to have interests. Feinberg (see below) maintains that having an interest in *x* is both a necessary and sufficient condition of having a right to *x*.¹⁸ However, White rules out having an interest in something as either a necessary or sufficient condition for having a right.

When we say that someone (or something) has an interest in something *x* we mean that either that: 1. *x* is interested in *y*, (in other words, *x* is concerned with, entertained or provoked by *y*), as in the case of someone being interested in Etruscan Pottery, or that 2. *x* can be benefitted by *y*, as in the case of having an interest in not being robbed. The first of these formulations seems irrelevant to the question of rights but the second does not. For although few would want to make the possession of a right dependant on the person's intellectual involvement with something, the view that the possession of a right to *x* is

¹⁷Alan R. White, *Rights*, (Oxford: Oxford University Press, 1984).

¹⁸Joel Feinberg, "The Rights of Animals and Unborn Generations," in *Philosophy and Environmental Crisis* ed. William Blackstone, (Athens: University of Georgia Press, 1974).

dependant on the person being able to be benefitted by x is held by Feinberg among others.

The position that in order for an entity to have a right to x the entity must be capable of being benefitted by x gains its support from the understanding that if the loss of x is irrelevant to the entity's fortunes then he, she or it cannot present a plausible claim against another for the loss of x . Put differently, one might say that the purpose of a saying that an entity has right to x is to protect the entity's interest in x .

Furthermore, this notion of harm and benefit is itself tied in part to the view that the entity in question is capable of experiencing the harm or benefit. If we are talking about the interests and rights of living organisms (as we are here) as opposed to the rights and interests of corporations, States or governments then if we maintain that the entity's right to x is tied to the entity's interest in x (which, in turn, is tied to the entity's capacity to be benefitted or harmed by the presence of loss of x) then it would seem necessary that the entity is capable of experiencing that benefit or harm. For what would it mean to say that entity could be harmed by the loss of x if the entity was incapable of experiencing or being affected by the loss of x ? It is for this reason that we are led to question whether it is meaningful to ascribe interests to those profoundly incompetent individuals who are incapable of experience anything except perhaps the most rudimentary sensations.

White maintains that having an interest in x or having the capacity to have an interest in x is neither a necessary nor a sufficient condition for having a right to x . It is not a sufficient condition as there are many things that people have an interest in but this does not mean that they have a right to them. For example, I have an interest in going on a cruise around the Greek Islands, but regrettably, this does not mean I have a right to such a excursion. It is not a necessary condition since I can have a right to x although I might not have any interest in x . For example, I have a right to vote in the next General Election

though I might have no interest in doing so.

However, this might be questionable for as was stated above, the phrase " x has an interest in y " is vague; it can mean either that x is interested in y (in other words, x is concerned with/intellectually involved with y) or that x can be benefitted by y . The above example matches the first of these two formulations; that I am concerned with/intellectually involved with the result, and it seems implausible to maintain that an intellectual concern with x implies a right to x . The more important question is whether White's argument works if we take the second of the formulations to be the relevant one. Thus we are trying to imagine a case where a person could have a right to x even though she can be neither benefitted nor harmed by x . And it seems difficult to imagine such a situation which suggests that if we do take the second of the formulations to be the relevant one, then White's objection to the connection between interests and rights seems questionable.

Secondly, White rules out being a beneficiary of a duty as necessary or sufficient for someone else having a right. His argument is that one can have obligations where there are no corresponding rights. One may have a duty to complete a certain task, though it not be the case that others can regard this as a right. For example, I might have a duty to sweep snow from my drive though, perhaps, no one can claim this as a right against me. Additionally, one may think that one has a duty not to torture kittens or infants not because they have the right not to be tortured, but because of a more general duty of beneficence.

Thirdly, White endorses the notion that being a person is a necessary and sufficient condition of having a right. His argument is that we cannot present any "substantive characteristics" that are necessary or sufficient for the having of rights. If the above arguments are correct, then a person's right to x is not determined by either of the "substantive characteristics" that either a person has an interest in x or others have a duty (to

do x or refrain from doing x). As White says,

...it follows from all this that no valid argument can be given either for including or for excluding, for example, children, imbeciles, foetuses, animals, natural objects, unborn generations, as holders of rights on the grounds that being capable of having a right ensures or necessitates being capable either of having something in one's interest or of being interested in something. Hence the question whether animals, natural objects, etc., can or cannot have interests...is irrelevant to the question whether they can have rights and need not, therefore, be decided.¹⁹

For White, the important question is whether the subject is "logically capable of having a right", that is to say, whether that person can exercise, earn or claim that right.²⁰ It is irrelevant whether the entity is capable of having interests since, as above, being capable of having interests is neither necessary nor sufficient for the possession of rights. In order to be the possessor of a right the individual must have the capacity to understand what the right claims and know when that right is being respected or not. For this reason White states,

...only a *person* can logically have a right because only a person can be the subject of such predications. Rights are not the sorts of things of which non-persons can be the subjects, however right it may be to treat them in certain ways.²¹

White's endorsement of the view that only persons are capable of understanding and exercising rights is closely tied to his definition of a right. White defines a right as

...something one can have or be given, earn, enjoy, or exercise. It is something one can claim, demand, assert, insist on, fight for, or secure or what one can waive, surrender, relinquish, or forfeit.²²

Clearly, only those with sophisticated cognitive and social skills are capable of doing any or all of the above. Thus White would, in fact, deny that children, animals and incompetent adults are capable of having moral rights (though competent persons would retain rights when

¹⁹White, *Rights*, 82.

²⁰Ibid., 89-90.

²¹Ibid., 90.

²²Ibid., 17.

asleep since the capacity to "claim, demand.." the right is not lost while the person is asleep). Moreover, White would seem to conclude that the incompetent patient has neither the right to refuse life-sustaining treatment nor the right to not receive such treatment, since the incompetent patient would seem to be incapable of being able to "claim, demand, assert..." any right.

An alternative view is put forward by Feinberg.²³ Feinberg disagrees with the claims made by White that: (1) only persons can have rights, and (2) that in order to have a right one must be able to "claim, demand, assert...the right." Feinberg points out that young children and the demented have "legal rights recognized and easily enforced by the courts." Although these two groups cannot exercise these rights on their own behalf, we "accept the actions of proxies or attorneys who are empowered to speak in their names." As a positive argument, and in opposition to the account proposed by White, Feinberg argues that the possession of rights is tied to the capacity to have interests,

...the sort of beings who *can* have rights are precisely those who have (or can have) interests. I have come to this tentative conclusion for two reasons: (1) because a right holder must be capable of being represented and it is impossible to represent a being that has no interests, and (2) because a right holder must be capable of being a beneficiary in his own person, and a being without interests is a being that is incapable of being harmed or benefitted, having no good or "sake" of its own.²⁴

On Feinberg's account, a necessary condition of having a right to *x* is that the entity is capable of having an interest in *x*. This would seem to be implied by Feinberg's claim that "the sort of beings who *can* have interests are *precisely* (emphasis added) those who have...interests". In other words, all and only those who are capable of having interests have rights.

As we saw above, White denies that having an interest in *x* is either a necessary of

²³Joel Feinberg, "The Rights of Animals and Unborn Generations," in *Philosophy and Environmental Crisis* ed. William Blackstone, (Athens: University of Georgia Press, 1974).

²⁴*Ibid.*, 264.

sufficient condition of having a right to *x*. However the apparent disagreement between Feinberg and White may not be as stark as first appears, for Feinberg is not claiming that having an interest in *x* is a necessary and sufficient condition of having a right to *x*. Rather, he is claiming that in order for an entity to have rights the entity must be capable of having interests - that is to say, rights are granted on account of the entity's capacity to have an interest in *x* rather than an actually having an interest in *x* itself. Thus Feinberg need not admit that an interest in going on a cruise around the Greek implies a right to such a cruise, since rights are not granted according to interests *simpliciter*. The claim would be that the person could not have right to the cruise if the person was not capable of having an interest in it.

Although Feinberg grants rights to children he does not grant rights to all humans.

Rights are denied to those entities who have no "conative life":

...no conscious wishes, desires, and hopes;...or unconscious drives, aims, and goals; or latent tendencies, direction of growth, and natural fulfillments. Interests must be compounded somehow out of conations; hence mere things have no interests. *A fortiori*, they have no interest to be protected by legal or moral rules.²⁵

Conversely, Feinberg maintains that

...many of the higher animals at least have appetites, conative urges, and rudimentary purposes, the integrated satisfaction of which constitutes their welfare of good.²⁶

With regard to the incompetent, in most cases, the incompetent patient will have a higher "conative life" than these animals, and so should be accorded rights. Although it might seem offensive to make such comparisons, an Alzheimer's patient who is severely demented

²⁵Ibid., 263-4.

²⁶Ibid., 264.

will still have wishes, desires and hopes, and some rudimentary goals.²⁷

Feinberg admits that there is a certain class of incompetent individual who is incapable of a "conative life" and thus cannot be the possessor of interests. The example would be a patient who is in a persistent vegetative state (hereafter abbreviated to "PVS").²⁸ It seems to be generally agreed upon that pvs patient can neither be harmed nor helped by continued treatment since he or she cannot experience harm or benefit. Thus any claims as to the "best interest" of such a patient would have to be made in reference to former interests or in terms of a vaguer notion of human dignity.

The above discussion seems to show that there are problems with White's contention

²⁷"The major clinical feature of Alzheimer's disease is a progressive dementia. The term *dementia* refers to a global decline in cognitive function. Depending on the severity of the dementia, there are defects in several, if not all, areas of intellectual function. In the typical patient, the earliest symptoms are those of memory loss and difficulty with complicated intellectual functions. These symptoms are then followed by more generalized confusion and mild problems with speech. The patient will ultimately have difficulty in carrying out a simple conversation, reading, writing, or following simple directions. Changes in social behavior and personality, including psychotic and aggressive behavior can be seen at any stage of the illness..." Jacob H. Fox, "Alzheimer's Disease and Other Dementias," in *Neurology for the Non-Neurologist*, ed. William J. Weiner, and Christopher G. Goetz, (Philadelphia: J.B. Lippincott Company, 1989), 198.

²⁸The term "Persistent Vegetative State" (PVS) is a recent term, coined by Dr Fred Plum. Dr Plum describes the "vegetative state" in the following terms: Vegetative state describes a body which is functioning entirely in terms of its internal controls. It maintains temperature, it maintains heart beat and pulmonary ventilation. It maintains digestive activity. It maintains reflex activity for muscles and nerves for low level conditioned responses. But there is no behavioral evidence of either self-awareness or awareness of the surroundings in a learned manner". *In re Jobes*, 108 N.J. 394, 529 A 2d 484, 488 (1987). See also Fred Plum and Jerome B. Posner, *The Diagnosis of Stupor and Coma*, 3d. ed. (Philadelphia: F.A. Davis Company, 1980), and Walter B. Jennet and Fred Plum, "The Persistent Vegetative State: A Syndrome in Search of a Name," *Lancet* 1 (1972): 734-737, where the term "persistent vegetative state" appeared for the first time. The most recent (1993) report of the ANA Committee on Ethical Affairs defines the term *vegetative state* as, "a chronic neurological condition characterized by lack of awareness of external stimuli, with preservation of vital vegetative functions, such as cardiac action, respiration, and maintenance of blood pressure. As far as can be determined, the patient remains unaware of both self and environment. All cognitive functions appear to be absent," *Annals of Neurology*, vol. 33 no. 4 (1993).

that having an interest in x is neither a necessary nor a sufficient condition of having a right to x . For a plausible case can be made for maintaining that having an interest in x is a necessary but not a sufficient condition of having a right to x . As such, there seems more to recommend Feinberg's account than White's for there does seem to be an important connection between being capable of having interests and the possession of rights. Additionally, as Feinberg points out, there are clear precedents for the granting of rights to those entities who have interests but who cannot actually claim or present claims on their own behalf.

Moreover, there would seem to be other problems with restricting rights to persons. If we agree with White and deny that incompetents can be the possessors of rights since they are not persons, we still have to take into account the claims that: (1) except in the most profound cases of incompetence, incompetent patients have interests, and (2) other things being equal, we do have obligations to respect the interests of other entities. For example, even if we might want to deny that animals have a right not to be harmed since they cannot claim such a right, my cats have an interest in not being harmed and, other things being equal, I should not harm them. Furthermore, even if we might conclude that such entities cannot present fully-fledged rights claims against the actions of others there may still exist obligations to respect the interests of such entities. Thus even if we accept that the incompetent person does not have the right to refuse life-sustaining treatment or the right not to receive life-sustaining treatment on the grounds that he or she does not meet the conditions identified by White, one could still claim that the incompetent person has an interest in the cessation of such treatment, and that, other things being equal, this treatment should be ceased. That is to say, we might decide that the question of whether the incompetent patient has a *right* to refuse or not to receive life-sustaining treatment is of negligible importance, if the practical

impact of denying him or her these rights is minimal.

An additional reason why one might criticize White's account is that even if we agree with White that only persons are capable of having rights, then the incompetent patient will lose all of the rights he or she held when competent if the loss of competency is somehow linked to the loss of personhood. That is to say that if we want to maintain that the incompetent person retains the rights he or she possessed when competent we must reject White's account. Or differently, we would need to modify White's account to permit the incompetent patient to retain the right to refuse medical treatment that she held when she was competent.

One can claim (as, in fact, Buchanan and Brock do) that the competent person's rights extend to the later time if and when he or she is incompetent. On Feinberg's account, no distinction need be made between whether the patient is competent or incompetent so long as the patient's capacity for having interests is preserved. If the patient's capacity for having interests is preserved while incompetent, and it can be shown that the patient has an interest in the cessation of treatment, then the patient retains the right to refuse or not to receive medical treatment. One way in which we can understand the interests of the patient continuing through incompetence is to talk in terms of *precedent autonomy*. In this case, the patient's desire for the cessation of treatment that was expressed while he or she was competent would dictate treatment decisions when the patient is incompetent. In this way, the right of self-determination applies not just to the present time when the person is competent, but continues through incompetency.

Three Different Approaches to Decision-Making for Incompetents

In this section we will identify three different approaches to decision-making for incompetent patients and then examine some of the recent relevant legal cases. In the following section we will present a number of criticisms against each of these approaches.

The first approach is to make a decision according to what is in the incompetent patient's best interest. This approach is termed "objective" since we do not have any evidence of this particular person's wishes while competent, so the standard used would be that of "reasonable person" - what would a reasonable person want in these circumstances? This was the standard used in *Storar*²⁹ and *Eichner*. The second approach, "substituted judgment," is for a guardian of the patient to make the decision "as the patient would have decided". This is regarded as a subjective standard since there is some evidence as to what the patient would have wanted, through informal comment and conversation. This approach was used in *Quinlan*,³⁰ *Conroy*,³¹ and most controversially *Sakiewicz*.³² The third approach, "advance directive," enables the incompetent person to make the decision for herself through the use

²⁹*In re Storar/Eichner*, 52 N.Y. 2d 363 420 N.E. 2d 64, 438 N.Y.S 2d 266 (Ct. App. 1981).

³⁰*In re Quinlan*, 70 N.J. 10, 355 A.2d 647 (1976).

³¹*In re Conroy*, 98 N.J. 821, 486 A.2d 1209 (1985).

³²Joseph Sakiewicz was a profoundly retarded sixty-seven year old man with the I.Q. of ten suffering from leukaemia. The appropriate treatment of chemotherapy would cause him significant side effects and discomfort with a no greater than a fifty-percent chance of prolonging his life for two to thirteen months. Without the chemotherapy he would die within weeks or months without discomfort. The Supreme Judicial Court of Massachusetts relied on both the right of privacy and the right of informed consent to permit the withholding of chemotherapy in this case. Reasoning that an incompetent person retains the same rights as a competent individual "because the value of human dignity extends to both," the court adopted a "substituted judgment" standard whereby courts were to determine what an incompetent individual's decision would have been under the circumstances. However, it is controversial how this determination could be made of a person who was never competent. *Superintendent of Belchertown State School v. Sakiewicz*, 378 Mass, 728, 870 N.E. 2d 417 (1977).

of the directive written when she was competent. The second and third approaches are termed "subjective" since we have specific information what treatment, or lack of, this particular patient requests. According to this standard, we should respect the particular patient's wishes whether or not they are in accordance with what a reasonable person would want. The *Cruzan*³³ court recognized the use of advance directives as the best way to present "clear and convincing" evidence of the competent person's wishes.

The three approaches can be described in different ways. With regard to the incompetent patient's right to refuse treatment it would seem that, *prima facie*, only the two subjective standards would support such a right. For as Dresser suggests in her term *precedent autonomy*, these two approaches rest on the principle of autonomy and the understanding that either directly or indirectly, the incompetent patient does have this right to refuse treatment. The best interest approach, on the other hand, looks simply at the condition of the incompetent patient, and decisions to withdraw or withhold treatment are made with respect to notions of beneficence (and perhaps justice), rather than with regard to notions of autonomy. Since we are dealing here with patients who are, *ex hypothesi*, non-autonomous, any obligation to respect the patient's autonomy is diminished.

If we look at recent legal cases dealing with the rights of incompetents to refuse life-sustaining treatment we can see that up till *Cruzan* the courts ruled that this right is extended to incompetents. The courts based the claim that the incompetent patient has a right to refuse treatment either on the common law right to informed consent, (as evidenced by Cardozo's remark in *Schloendorff* that competent persons have the right to refuse unwanted medical treatment) or on both this right and a constitutional privacy right (the "liberty

³³*Cruzan v. Harmon*, 760 S.W.2d 408 (Mo. 1988); *Cruzan v. Director, Missouri Dept. of Health*, 1990 US Lexis 3301 (US June 25, 1990).

interest" of the Fourteenth Amendment, that no State shall "deprive any person of life, liberty, or property, without due process of law"). The courts gave different reasons why this right is extended to incompetents. In *Quinlan* the court ruled that the "only practical way" to prevent the loss of Karen Quinlan's privacy right due to her incompetence was to allow her guardian and family to decide "whether she would exercise it in these circumstances." In *Conroy* the New Jersey Supreme Court argued that the right of self-determination should not be lost merely because an individual is unable to sense a violation of it. Whereas in *Sakiewicz*, the Supreme Judicial Court of Massachusetts reasoned that an incompetent person retains the same rights as a competent individual "because the value of human dignity extends to both". In *Storar/Eichner* the courts based their decision on respect for informed consent. In *Storar* the court found the right to refuse life-sustaining treatment "adequately supported" by the informed consent doctrine, and in *Eichner*, the court found clear evidence from statements that the patient made while competent that he "did not want to be maintained in a vegetative coma by use of a respirator."

The matter is different, though, in *Cruzan*. Here, the U.S. Supreme Court agreed that a competent person does have a constitutional right to refuse treatment and this included the refusal of lifesaving nutrition and hydration. However, the Court did not think that this right extended to those persons who are incompetent and "unable to make an informed and voluntary choice to exercise that hypothetical right or any other right." However, the court ruled that life-sustaining treatment could be ceased if there was "clear and convincing" evidence of the previously competent persons wishes.

It is worth examining the main differences between *Quinlan* and the other earlier cases, and *Cruzan*. There are five main differences: First, the courts differ in their answer to the question of whether an incompetent person has the right to refuse life-sustaining

treatment. The *Quinlan* court argued that Nancy Quinlan had the right to refuse such treatment based on her constitutional right to privacy and that this right was not affected by her being incompetent. According to the court, the state does have an interest in the "sanctity of human life" but this interest weakens as the degree of bodily invasion of the treatment increases and the prognosis dims.

The *Cruzan* court, on the other hand, decided that Nancy Cruzan did not have the right to refuse such treatment. It based its decision on the arguments that although a competent person has a constitutional right to refuse treatment it doubted that this right can be equated with "a broad right to privacy that would support an unrestricted right to refuse treatment..." Such a right has to be balanced against relevant state interests, and one of these interests being the "protection and preservation of human life." Furthermore, although the court recognized a right to refuse treatment "embodied in the common law doctrine of informed consent," the court doubted whether such a right applied in this case since Nancy Cruzan was "unable to make an informed and voluntary choice to exercise that hypothetical right or any other right." In other words, the right to direct treatment based on the doctrine of informed consent is dependant on the person's being able to exercise that right by making an "informed and voluntary choice."

Second, the courts differed with regard to how the right to refuse treatment can be exercised. The *Quinlan* court decided that Karen Quinlan's right to privacy that would permit the withdrawal of life-sustaining treatment can be exercised on her behalf by her guardian through the use of a "substituted judgment." In other words, contra *Cruzan*, an incompetent person can still exercise the right to have life-sustaining treatment withdrawn even though he or she is unable to make an "informed and voluntary choice." However, although the *Cruzan* court accepted that in some cases a surrogate may act for the patient, it denied that a

person's constitutional right to direct treatment extended to a surrogate. The *Cruzan* court rejected the assertion of the right of privacy on the person's behalf on the grounds that, unlike Nancy *Cruzan*, not all incompetent person's have loving parents.

Close family members may have a strong feelings-a feeling not at all ignoble or unworthy, but not entirely disinterested either....there is no automatic assurance that the view of close family members will necessarily be the same as the patient's would have been had she been confronted with the prospect of her situation while competent.

In effectively removing the family of the incompetent person from the decision to withdraw life-saving treatment, the *Cruzan* court was distancing itself from the *Quinlan* court's recognition that the incompetent person's family are particularly able to decide what their child would have wanted.

Third, the courts differed with regard to the evidentiary standard necessary for determining what the incompetent person would have wanted. The *Quinlan* court adopted a low standard; it permitted the use of a substituted judgment by Karen Quinlan's father even though there was no evidence of her wishes regarding the withdrawal of life-sustaining treatment. The *Cruzan* court supported a higher evidentiary standard, defending Missouri's requirement for "clear and convincing" evidence of Nancy *Cruzan*'s wishes that such treatment be withdrawn. We can explain this difference in evidentiary standard by making use of a distinction used in *Conroy*. The distinction between "subjective" and "objective" standards of determining whether life-sustaining treatment should be withdrawn. The *Quinlan* court supported an "objective" standard; the decision to withdraw treatment was based primarily on best interest considerations of the burden of prolonged life.³⁴ The *Cruzan* court supported

³⁴We can further distinguish between "limited-objective" and "pure-objective" standards; on one interpretation we might say that the court used a "limited-objective" standard since the decision to withdraw treatment was based on a determination of what Karen Quinlan herself would have wanted as well as best interest considerations of the burden of prolonged life. However, it might be more accurate to say the decision was based on a "pure objective"

a restricted use of the "subjective" standard; life-sustaining treatment can be withdrawn only if we have clear and convincing evidence of the person's own expressed wishes. In the absence of such evidence the court held that it was best to err on the side of preserving life, in this case using best interest arguments to support the continuation of treatment.

Fourth, the courts differed with regard to treatment itself. The *Quinlan* court distinguished between "ordinary" and "extraordinary" treatment; a treatment is "extraordinary" if it is highly invasive and has little or no therapeutic effect and serves merely to prolong the dying process. The court made use of this distinction to the effect that the state's interest in the "sanctity of life" diminished if the treatment was extraordinary. The *Cruzan* court rejected such a distinction; a competent person had the right to reject any treatment, and this included nutrition and hydration which the court now recognized as constituting treatment.

Fifth, the courts disagree as to the mechanism by which it is permissible to withdraw life-saving treatment from an incompetent person. To meet the standard for clear and convincing evidence the *Cruzan* court recommended the use of advance directives or Living Wills. In this way, the court affirmed a competent person's right to refuse such treatment and recognized that this includes the refusal of life-sustaining treatment at a later time when the person has become incompetent. Thus the court accepted that a competent person has the right to refuse treatment both when he or she is competent and, through the use of an advance directive, when he or she is incompetent. But the court did not accept that an incompetent person has the right to refuse treatment, for this right is lost when the person becomes incompetent since he or she is no longer able to exercise that right. The *Quinlan* court made no such distinction; a person has the right to refuse life-sustaining treatment

standard since, as the court admitted, there was no evidence of what Karen Quinlan would have wanted.

whether competent or incompetent.

Prima facie, there would seem to be a difference between those cases, like *Quinlan* and *Cruzan*, that deal with the withdrawal of life-sustaining treatment from an incompetent person who was once competent and those where the person was never competent, as was the case in *Sakiewicz* and *Storar*. The difference would be that in those cases where we are dealing with a previously competent person we would have some subjective evidence of what that person would have wanted. This evidence might be clear and convincing or perhaps only enough to suggest what the person would have wanted. In those cases where the person had never been previously competent the problem of trying to decide what the person would have wanted based on his or her previously expressed wishes would not arise as we have no evidence of the person's wishes. Thus in *Storar* the court reasoned that since the individual concerned had been profoundly retarded during most of his life, "it is unrealistic to attempt to determine whether he would want to continue potentially life prolonging treatment if he were competent." In such cases it would be appropriate to use a "pure objective" standard for determining what is in the person's best interest since there would be no "subjective" evidence. However, in *Sakiewicz* the court did appoint a guardian *ad litem* to make a "substituted judgment" based on this pure objective standard. The court reasoned that (contra *Cruzan*) the incompetent person retained the same rights as the competent person "because the value of human dignity extends to both."

Clearly, in the use of an "objective standard" to determine what the person would have wanted the *Sakiewicz* court took a very different approach from that adopted by the *Cruzan* court. There would have been no need to require clear and convincing evidence of Nancy Cruzan's wishes if the decision to withdraw life-sustaining treatment could have been made on this objective standard. The *Cruzan* court did not adopt the objective standard on the

grounds that (a) the right to refuse treatment is not extended to an incompetent person, and (b) substituted judgments cannot be made where we have no evidence of the person's wishes. Put differently, in using an objective standard to authorize the removal of life-sustaining treatment we are by-passing the issue of what this particular person would have wanted, and in its place we are adopting a standard that enables us to conclude that any person given a similar condition to that of Joseph Sakiewicz would request (if possible) the removal of the treatment. This is based on the contention that we would be able to say that the continuation of chemotherapy for someone in Joseph Sakiewicz's condition would not be in their best interest.

But this was the same type of reasoning as that used by the *Quinlan* court. In *Quinlan* the court adopted an objective standard in conjunction with a "substituted judgment." However, in neither *Quinlan* nor *Sakiewicz* was this substituted judgment based on the "subjective" expressed wishes of the person but on an objective "best interest" standard.

The difference between *Cruzan* and the earlier cases might be less than first appears, for although *Cruzan* specifically states that the incompetent person does not have a right to refuse treatment, it seems to accept that a competent person does have this right, and if the evidence of the person's wishes is clear and convincing, then the right can be extended through incompetence.

In the above description of recent the case law there is little inquiry as to why or why not legal rights to refuse treatment should be extended to incompetents. Thus although in *Sakiewicz* the court stated that competent and incompetent patients have the right to refuse treatment because the "value of human dignity extends to both," the question remains whether having this value is a necessary and sufficient condition of the right to refuse life-sustaining treatment.

The Limitations of the Three Approaches to Decision-Making

As is clear from the earlier section, the courts have shown a marked preference for the two subjective approaches to decision-making for incompetent patients. These two approaches have clear merits in that they do attempt to respect the wishes of the former competent person, and thereby endorse the notion of patient self-determination. However, there are problems with these two approaches. These problems can be divided into two camps: 1) problems of accuracy and compliance; 2) problems of authority and benefit.

Problems of accuracy and compliance. Despite the fact that nearly every state has some legislation on the use of advance directives, and their endorsement following *Cruzan* and the Patient Self Determination Act (1990), the use of advance directives has not been widespread. In either of its forms, Living Will or Durable Power of Attorney for Health Care, the advance directive has not been very successful as a means of decision-making for incompetent patients. Two main reasons have been presented for this. First, with regard to accuracy there are problems both with regard to the competent person's ability to predict what treatment decisions he or she would want to be made if he or she should become incompetent, and with being able to write a directive that is not so general so as to be unworkable. In the case of persistent vegetative state one might be relatively clear that to be in such a state would be good as being dead, and hence any life-sustaining treatment would be unwarranted. But we can imagine cases of dementia or of other neurological deficit in which we are not so able to describe the "mental life" of the incompetent patient. Furthermore, the person writing the directive may agree to the cessation of life-sustaining treatment in some cases but not others, even if both involve neurological deficit. Buchanan and Brock provide the following example.

An apparently valid advance directive specifies that no life-sustaining treatments, including antibiotics, are to be used on the individual if he or she suffers a serious loss

of cognitive function. The patient does suffer a serious loss of cognitive function due to a discrete neurological injury, such as a stroke: The person is now moderately mentally handicapped. However, the patient is otherwise healthy, and apparently quite happy. Then the patient develops a life-threatening pneumonia. Following the advance directive will result in the easily avoidable death of a happy and healthy, though mentally handicapped individual.³⁵

Second, with regard to compliance, there are problems with advance directives both with the ease of their revocation and their being respected. The case described by Buchanan and Brock above is testimony to the claim that we might not want advance directives to have absolute authority, but at present they can be revoked by ambiguous statements from questionably competent patients. And in cases where the treatment team might think that, in good faith, the patient's wishes for withdrawing or withholding life-sustaining treatment are unreasonable, it would not be unheard of for treatment to be continued nevertheless.

There are analogous problems with substituted judgment. In this case there are questions as to the ability of someone else to determine what the patient would have wanted. As was pointed out above, such a strategy would only work in cases where the patient was previously competent. In both *Quinlan* and *Cruzan* we saw the problems in trying to determine what the patient would have wanted when there is a questionable sufficiency of information to do so. Furthermore, there is the additional problem of choosing proxy decision-makers who are both knowledgeable as to what the patient would have wanted and who have no specific interest of their own that could influence their decision. As Chief Justice Rehnquist argued in *Cruzan*, one should not assume that family-members are always the best arbiters of what the incompetent patient would want.

Problems of authority and benefit. There are other and more fundamental problems with the two subjective approaches, and these have to do with the identity of the incompetent

³⁵Buchanan and Brock, *Deciding for Others*, 108.

patient and what is in that patient's best interest. The matter is stated succinctly by Dresser.

According to the accepted approach, treatment choices on behalf of incompetent patients - those who are unable to choose for themselves - should incorporate whatever relevant preferences and values the patient expressed while competent; that is, treatment should be what the patient "would want" if competent to choose. The underlying assumption is that the competent individual's concept of personal well-being applies equally to the patient as a demented person. Yet this assumption is vulnerable to challenge, for in significant respects the demented patient is no longer the one who formerly expressed the treatment preferences. Indeed, the patient's well-being now could be quite different from that of the former competent person.³⁶

These two assumptions can be summarized as follows:

- i) The competent person's concept of personal well-being applies equally to the patient as a demented patient.
- ii) The demented patient is the same person as the one who formerly expressed the treatment preferences.

With regard to the first assumption, the problem here can perhaps be termed one of perspective. From the perspective of the competent person, the life of an incompetent patient (for example, someone suffering from Alzheimer's dementia) may appear to be a life not worth preserving. If one values reading books, travelling, going to the cinema, then no longer being able to do these things may seem to deprive life of much, if not most, of its value. But, on the other hand, one could say that the life of the demented patient, although not as full or as varied, still contains value and enjoyment for the demented patient. Thus the use of advance directives to make treatment decision for the later incompetent individual might deprive the incompetent individual of a worthwhile life.

There are similar problems with substituted judgement, in fact it could be claimed that the problems are more obvious. If we have reservations about a competent person determining what would be in her best interest were she to later become incompetent, these problems would seem to be exacerbated when other people attempt to make this

³⁶Dresser, "Autonomy Revisited," 71.

determination on her behalf. We can see this problem in the *Spring*³⁷ case. In this case, a demented nursing home patient required renal dialysis, but had no other serious health problems. At times, the patient physically resisted the dialysis and had to be sedated. The patient's family sought a court order to have the treatment stopped, and this was granted on the evidence of the family's claim that this is what the patient would have wanted were he competent. However, there was no specific evidence to this effect apart from the claim that the patient's formerly active physical life-style suggested that he would refuse treatment if he were competent. Let us grant that this is what the patient would have wanted were he competent, but it seems questionable whether this can be inferred from the evidence presented - many people live an active life-style but this does not mean that they do not want to forego all life-sustaining treatment. And furthermore, this does not address the question of whether this is what is best for the incompetent individual.

A challenge to the second of these assumptions, that the incompetent patient is the one who previously expressed the treatment preference, has more serious consequences. For if it can be shown that the incompetent patient is not the one who previously expressed a treatment preference, then this would undermine the "moral authority" of advance directives (to use Buchanan and Brock's phrase). It will be useful here to examine Buchanan and Brock's presentation of this challenge to the moral authority of advance resolution and their proposed resolution.

In Chapter 3 of *Deciding for Others* Brock and Buchanan examine the following challenge to the "moral authority" of advance directives.

Advance directives only have moral authority if the person who issued the directive and the person to whom the directive would be applied are the same person; but the very circumstances which would bring an advance directive into play are often those

³⁷*In re Spring* 380 Mass. 629, 405 N.E. 2d 115 (1980).

in which one of the necessary conditions for personal identity is not met.³⁸

The authors present this challenge in another form, in what they term the "slavery argument."

- (i) One person's advance directive has no moral authority to determine what is to happen to *another person*.
- (ii) In some cases of severe and permanent neurological damage, for example, that due to advanced Alzheimer's dementia, psychological continuity is so disrupted that the person who issued the advance directive no longer exists.
(Therefore)
- (iii) In such cases the advance directive has no moral authority to determine what is to happen to the individual who remains after neurological damage has destroyed the person who issued the advance directive.³⁹

The argument is invalid, contend Brock and Buchanan, since (iii) does not follow from (i) and (ii). For the argument to be valid we would have to replace (ii) with (ii')

- (ii') The individual who remains after neurological damage has destroyed the person who issued the advance directive is a (different) *person*.

Brock and Buchanan reject the "slavery argument". They argue that the slavery argument does not challenge the moral authority of advance directives since "the living being who remains is *not a person at all* and hence, *a fortiori*, not a different person." Furthermore Buchanan and Brock argue that it is the case that a competent person has the right, through the use of an advance directive, to determine treatment for his or her later non-person self.

The central argument that is presented to support the first of these claims is that the neurological damage that is suffered by the individual that would bring the advance directive into play not only destroys personal identity but destroys personhood as well. This is to say that the degree of neurological insult is sufficient not only to deprive the individual of those psychological features that go to define the identity of the person but also to deprive the

³⁸Buchanan and Brock, *Deciding for Others*, 155.

³⁹*Ibid.*, 157-8.

person of those cognitive capacities that constitute personhood. Brock and Buchanan present the following as some of the "strongest candidates for being included among the necessary conditions" for being a person.

- (a) The ability to be conscious of oneself as existing over time - as having a past and a future, as well as a present.
- (b) The ability to appreciate reasons for or against acting...
- (c) The ability to engage in purposive sequences of actions.⁴⁰

In the following chapter we will examine in detail whether these criteria of personhood are acceptable, and therefore it would be redundant to examine them here. However, we can present a number of criticisms to the general claim that in most if not all cases the later incompetent patient will be a nonperson. We can certainly imagine cases, similar to those that Brock and Buchanan imagine, where it does seem as though the individual fails to meet any of the conditions of personhood that have been put forward. The standard example would be that of a patient in a persistent vegetative state. However, even if it could be argued that it was justifiable and useful to categorize certain individuals as nonpersons, it is not clear that the number of cases where this would be appropriate is as great as Brock and Buchanan imagine. In the case of an individual in a persistent vegetative state we are looking at the most severe degree of insult to the brain short of actual brain death. This might be a case where it is justifiable and useful to regard such an individual as a nonperson. But there are many other cases where such a categorization cannot be so easily justified. Consider the following case: A patient, Ms. T, writes an advance directive that states that she does not want life-sustaining treatment to be performed if it merely serves to prolong the dying process. Some time later she develops Alzheimer's dementia. There will

⁴⁰*Ibid.*, 160.

be a time when Ms. T is incompetent to look after herself but still has some or all of the cognitive capacities described above. It may be the case that we would want to say that at some point in her progressive dementia Ms. T no longer meets the conditions of psychological continuity (what ever they might be) but until that point we will have a patient who although incompetent may still meet the conditions of personhood.

Additionally, there is the point that persons are not the only entities that have interests. This was a point that came up in our previous discussion of the accounts of White and Feinberg. To recall, Feinberg argued that entities that do not qualify as persons, for example young children and animals, still have interests that should be respected. Thus even if it could be successfully argued that the incompetent patient is not a person this does not mean that such a patient does not have interests. And as we saw in our discussion of the *Spring* case above, the interests of the incompetent patient may conflict with the previously expressed wishes of the competent person as stated in the advance directive.

Brock and Buchanan regard cases of PVS "wholly unproblematic" in the sense that, since the PVS patient is incapable of having interests, so the patient does not present any claims against us. Therefore, apart from the interests of a third-party, pvs patients would have no interests in any treatment decisions made. But this is not the case if the incompetent patient retains the capacity to have interests, and I think that this will be true in the majority of cases of incompetence.

Our conclusion here is that Buchanan and Brock's proposed resolution to the challenge to the moral authority of advance directives is not successful. There seem to be plausible reasons for denying the claim that in most cases where the use of an advance directive (or substituted judgment) would be appropriate the incompetent patient is a nonperson. And even if this claim can be defended, there is still the further problem that the

nonperson may have interests that are odds with those of the previously competent person.

Does this mean that we should abandon the subjective standard and instead base our decision-making simply on what is in the best interest of the incompetent patient? This conclusion is perhaps unwarranted. What is required is to pay closer attention to the interests of the incompetent individual as a means of putting boundaries on precedent autonomy. As Dresser says, we need

...limits on the range of choices competent persons have concerning their future treatment and the choices available to proxy decision-makers in general. In short, we as a society must take a moral and legal stand on which categories of incompetent patients must be treated and which may have treatment foregone.⁴¹

However, although we would not want to say that the well-being of the competent will always be that of the incompetent patient, it would seem unwarranted to ignore the previously expressed wishes of the competent person. As Dresser says, one possible solution to safeguard the interests of the incompetent patient is the development of standards as to when life-sustaining treatment should be continued, even if this goes against the expressed wishes of the previously competent person. However, this approach would undermine concern for patient self-determination. Furthermore, it sees the competent and incompetent individuals as distinct, and having contrary interests. Now in some cases this is defensible, as in the case of PVS. Here, by default (and by definition) the competent and incompetent individuals are distinct (though by default there is no conflict of interests). But in other cases, there is not such a clear demarcation between the two individuals. If we think of a case of progressive dementia, then we will see the gradual and tragic destruction of the mind of the patient. Here, there is no clear division between present and past interests, and it is less clear when we would be able to say that the interests of the later incompetent patient are not those of

⁴¹Dresser, "Autonomy Revisited," 83.

the competent person.

One way that we could attempt to cast light on these issues is by looking at the concept of personhood. As we saw in our discussion of the conditions for the possession of rights, the position supported by White was that only persons can be the possessors of rights. However, this does little more than put the question back one stage, for we need to know what it is about personhood that is necessary and sufficient for the possession of rights. Furthermore, we need to know just what we mean by personhood, for even if we were to agree with White that all and only persons have rights, we would need to know be clear as to the class of entities that the term "person" covers. That is to ask, what are the features of personhood?.

CHAPTER TWO

THE CONCEPT OF PERSON

In this chapter we will examine the necessary and sufficient conditions of the concept of person.¹ The main question of this chapter is the question of *class differentiation* -- what separates the class of persons from the class of nonpersons.² An answer to this question will enable us to evaluate Buchanan and Brock's claim that the incompetent patient is not a person. Our search for the conditions of personhood will not be easy, for there are a number of rival accounts each having a different set of necessary and sufficient conditions. Thus we will have to formulate some set of criteria by which we can claim that one account (or even perhaps, no account) of personhood is superior to the others.

The chapter will proceed as follows. In the first section we will make some preliminary remarks about the concept of person, and suggest two criteria by which we can judge the success of any account of the conditions of personhood. In the second section we will turn to a discussion of the necessary and sufficient conditions of personhood, and look at three rival accounts of personhood. In the third section we will attempt to determine whether any of the accounts can provide a plausible answer to the question of when and whether the incompetent patient can no longer be regarded as a person.

¹ Throughout this chapter I will use the terms "personhood" and "concept of person". Personhood refers in a loose sort way to the state of being person, roughly in the same way that we might talk about "childhood". The concept of person is a more formal term, and in our discussion of this concept we are investigating the features associated with this concept. Like the authors what we will be discussing, I will use these terms almost interchangeably, since I do not think that any crucial distinctions are muddled by doing so.

²Amelie O. Rorty, *The Identities of Persons*, (Los Angeles: University of California Press, 1976), 1.

The Concept of Person

As the discussions of the previous chapter showed, the concept of person is used to enable us to determine our moral responsibilities to different entities. As we saw, Buchanan and Brock argued that the different obligations that we have towards the prior competent individual and the present incompetent patient rests on the claim that whereas the former is a person, the latter is a nonperson. Crudely put, the person/nonperson distinction is meant to "carve the world up in a morally significant way." There are two ways the person/nonperson distinction could do this. First one might claim that persons possess certain properties lacked by nonpersons. For example, one might claim that a necessary condition of personhood is consciousness, and thus this property distinguishes persons from nonpersons. Second, there is the weaker claim that persons possess certain properties to a greater degree than nonpersons. For example, one might claim that although consciousness is a property of both persons and nonpersons, persons have a higher level of consciousness. Stated in this form, this weaker claim makes no assertion that there is a moral distinction between persons and nonpersons, thus we must add the claim that this difference in degree is itself morally relevant.

Thus in order for the person/nonperson distinction to carve the world up in a morally significant way we need to be able to show the following:

- i) There is a set of morally relevant properties possessed by persons and lacked by nonpersons.

or

- ii) There is a set of morally relevant properties possessed to a higher degree by persons than by nonpersons, and that this difference of degree is itself morally relevant.

This then suggests our first criterion for any account of personhood:

- 1) The distinction between persons and nonpersons must be a morally significant one.

One might object to this condition on the grounds that it implies that person is exclusively a moral concept. But this is not what we are claiming. Rather, we are claiming that the concept is *foundational* to morality; it permits us to distinguish different moral responsibilities to different entities.³ To understand this foundational nature it is helpful to distinguish between the *metaphysical* and *moral* senses or notions of person.⁴ Dennett writes,

Does the metaphysical notion - roughly, the notion of an intelligent, conscious, feeling agent - *coincide* with the moral notion - roughly, the notion of an agent who is accountable, who has both rights and responsibilities? Or is it merely that being a person in the metaphysical sense is a necessary but sufficient condition of being a person in the moral sense?⁵

To speak of the *metaphysical* or non-moral sense of person is to speak of a set of features by virtue of which a certain class of entities is constituted and which distinguish that class from all others. Thus, for example, on Dennett's account, the class of persons is constituted of all those entities who are "intelligent, conscious and feeling agents." To speak of the *moral* sense of person is to speak of the entity as the locus of moral value; as an entity that has "rights and responsibilities." For Dennett, persons have moral standing since they are the possessors of rights and obligations, a view that is echoed by Engelhardt.

Not all humans are self-conscious, rational, and able to conceive of the possibility of blaming and praising. Fetuses, infants, the profoundly mentally retarded, and the helplessly comatose provide examples of human non persons. Such entities are members of the human species, They do not in and of themselves have standing in

³Peter Carruthers, *Introducing Persons: Theories and Arguments in the Philosophy of Mind*, (London: Routledge, 1989), 228-9.

⁴Dennett makes use of this distinction in "Conditions of Personhood," in Rorty, *The Identities of Persons*, 176, and the distinction is discussed by Buchanan and Brock, *Deciding for Others*, 160-161.

⁵Rorty, *The Identities of Persons*, 176.

the moral community. They cannot blame or praise or be worthy of blame or praise. They are not prime participants in the moral endeavor. Only persons have that status.⁶

Relatedly, the objective that Tooley sets himself in the account of the concept of person that we discuss in this chapter

...is to formulate a concept of a person that is itself purely descriptive, and free of all moral and evaluative elements. The choice of content is to be determined, however, by moral considerations. In particular, it is to be guided by the answer to the following question:

What relatively permanent, non-potential properties, possibly in conjunction with other, less permanent features of an entity, make it intrinsically wrong to destroy an entity, and do so independently of its intrinsic value?⁷

Each of the accounts that we will look at presents certain properties that the author thinks are morally relevant. And in each case these properties are presented in the first instance as non-moral properties. For example, as the quotation from Tooley above suggests, Tooley is attempting in the first instance to describe the concept of person descriptively, although this description is influenced by moral considerations. The same is true of the accounts of Noonan and Gillett -- in each case the properties are deemed to be non-moral but necessary for moral standing.⁸

It might be thought that this distinction between the *metaphysical* and *moral* senses of personhood complicates our task even further. For in attempting to discover the necessary

⁶H. Engelhardt does refer to such individuals as persons in the "social sense." Tristram Engelhardt, *The Foundations of Bioethics*, (New York: Oxford University Press, 1986), 107.

⁷Michael Tooley, *Abortion and Infanticide*, (Oxford: Clarendon Press, 1983), 87.

⁸John Noonan, "An Almost Absolute Value in History" in *The Morality of Abortion: Legal and Historical Perspectives*, (Cambridge: Harvard University Press, 1970), reprinted in *Bioethics*, ed. Rem B. Edwards and Glenn C. Graber (New York: Harcourt Brace Jovanovich, 1988), 598-603. "Why Let People Die?" *Journal of Medical Ethics* 12 (1986) and "Reasoning about Persons," in *Persons and Personality*, ed. Arthur Peacocke and Grant Gillett (Oxford: Blackwell, 1987). See also "Consciousness, the Brain and What Matters," *Bioethics*, vol. 4, no. 3, (July 1990): 181-198.

and sufficient conditions of personhood, we need to clarify whether we are talking about metaphysical or moral personhood. Furthermore, as Buchanan and Brock contend, one should not assume that these two senses of person can be neatly distinguished

The fact that philosophers typically appeal to our intuitions about responsibilities and commitments - that is to basic *moral* concepts - in order to support their metaphysical theses..., casts serious doubt on this assumption....In other words, if the concept of a person...is a purely metaphysical nonmoral concept, then it is hard to understand why certain marks on paper made by a person in *this* (metaphysical) sense should *ever* be thought to create obligations or confer authority, since obligations can be created and authority conferred only by persons in the moral sense.⁹

There is perhaps some truth to the claim that it not always easy to distinguish the metaphysical sense from the moral sense, but this does not warrant the claim that there are not two senses. Nor is it being claimed that the concept of person is a *purely* metaphysical concept, for as Buchanan and Brock suggest, this would make the concept of person irrelevant to questions of moral authority and obligation.

But it is equally mistaken to claim that the concept of person is a *purely* moral concept. For if this were the case, then our moral categories would appear arbitrary. For instance, we might contend that since animals are not persons they have less moral worth or standing than you or I. But to make this claim we need to point to some feature that you or I possess and animals lack, for example, a certain level of cognitive ability or a soul. Thus the claim that *x* is a person and thereby the bearer of rights and responsibilities can avoid the appearance of arbitrariness only if we can identify the non-moral properties that the entity must possess and which we think are constitutive of persons in the moral sense. If the concept functions only in this way, then it would be hard to understand why certain entities are granted authority and have obligations.

Notwithstanding the contention that it is sometimes difficult to distinguish the

⁹Buchanan and Brock, *Deciding For Others*, 161.

metaphysical from the moral senses of person, the claim that will be defended here is that at root metaphysical personhood is a necessary and perhaps sufficient condition of moral personhood.

This provides us with a purported second condition of any account of personhood; that we should be able to differentiate metaphysical and moral senses of personhood on that account, and that the moral sense is in some way based on the metaphysical.

We can now summarize our discussion. Our first two conditions require of any proposed account of personhood that:

- i) The distinction between persons and nonpersons must be a morally significant one.
- ii) We should be able to differentiate metaphysical and moral senses of personhood and the moral sense must be based on the metaphysical.

Conjoined, these two conditions require that on any account of personhood the distinction between persons and nonpersons is a morally relevant one, and that this distinction is based in some way on metaphysical or non-moral features.

The Conditions of Personhood

Noonan

Noonan's account is examined here both for its own sake and because it is representative of a more general type of account of personhood; an account that regards either the actual *or* the potential holding of properties as necessary and sufficient for moral standing. Accounts of this sort are contrasted with those that claim that only the *actual* possession of these properties are necessary and sufficient. We will discuss Noonan's account specifically, but our more general goal is to determine whether moral standing can be granted through the potential possession of these properties. Thus we should not be satisfied with

a rival account to Noonan's if that account simply presents a different set of properties that can be held potentially.

Noonan does not speak of personhood as such, preferring to speak of the "humanity" of the individual. Nevertheless, the function of "humanity" on Noonan's account is consistent with the role ascribed to personhood, thus it is legitimate to regard this as a terminological rather than a functional difference.

Noonan's account parallels what he terms the "Christian position" of "ensoulment"; that the possession of a soul grants human beings moral standing. Noonan sees himself as presenting a "secular" version of this position. He argues that

...the theological notion of ensoulment could easily be translated into humanistic language by substituting "human" for "rational soul"; the problem of knowing when a man is a man is common to theology and humanism.

If one steps outside the specific categories used by the theologians, the answer they gave can be analyzed as a refusal to discriminate among human beings on the basis of their varying potentialities. Once conceived, the being was recognized as man because he had man's potential. The criterion for humanity, thus, was simple and all-embracing: if you are conceived by human parents, you are human.¹⁰

On Noonan's account, the humanity of the entity is determined by the fact that the entity is a human being and has the capacity for human wisdom, and it is this capacity that grants the human being moral standing. And so far as can be determined from Noonan's account all human beings have the capacity for human wisdom.¹¹

¹⁰John Noonan, "An Almost Absolute Value in History" in *The Morality of Abortion: Legal and Historical Perspectives*, (Cambridge: Harvard University Press, 1970), reprinted in *Bioethics*, ed. Rem B. Edwards and Glenn C. Graber (New York: Harcourt Brace Jovanovich, 1988), 598-603.

¹¹Support for this claim can be derived from the fact that (as we see below) Noonan regards conception as a morally important event, and that after conception the fetus has a much higher chance of developing into fully-fledged human being than does the sperm before conception. Moreover, since at conception one cannot tell whether the *conceptus* will eventually develop into a fully-fledged human being, one must conclude that on Noonan's account, a necessary and sufficient condition of having the capacity for human wisdom is that the entity has a complete set of human genetic material.

In addition to this condition of humanity, Noonan presents a "buttressing" consideration based on probability. An important reason why we might regard conception as being a morally relevant moment is that it is at conception that we see a substantial change in the probability that the entity will develop into a fully-fledged human.

If a spermatozoon is destroyed, one destroys a being which had a chance of far less than 1 in 200 million of developing into a reasoning being, possessed of the genetic code, a heart and other organs, and capable of pain. If a fetus is destroyed, one destroys a being already possessed of the genetic code, organs, and sensitivity to pain, and one which had an 80 percent chance of developing into a baby outside the womb who, in time, would reason

The positive argument for conception as the decisive moment of humanization is that at conception the new being receives the genetic code. It is this genetic information which determines his characteristics, which is the biological carrier of the possibility of human wisdom, which makes him a self-evolving being. A being with a human genetic code is a man.¹²

According to Noonan it is human wisdom that gives humanity moral relevance and grants the human being ("man") moral standing. It is not necessary that the entity actually possess human wisdom in order to have moral standing, it is sufficient that the entity has the capacity to develop human wisdom. Noonan does not tell us what he means by "human wisdom," but from the evidence of his use of the term "reasoning being" let us assume that human wisdom can be understood in terms of rationality, reason and deliberation. In any case, whether or not this is an accurate description of what Noonan means by "human wisdom" is not the main concern, for the success of Noonan's account rests on the claim that the capacity to develop human wisdom (whatever that might be) is morally relevant. If it is not morally relevant, then Noonan is reduced to claiming that moral standing is granted only on the actual possession of the relevant properties. This would challenge the claim that conception itself is morally relevant, and undermine Noonan's conservative position on abortion.

¹²Edwards and Graber, *Bioethics*, 601-602.

What does it mean to grant moral standing according to an entity's capacity to develop a certain property? Broadly speaking, it means that: (i) the property is regarded as morally relevant, and (ii) that in the normal course of events, the entity will develop this property. Noonan's position is thus a version of the potentiality argument, a type of argument that has been described by Tooley in the following way. Tooley asks whether there is a property *J* which satisfies the following three conditions:

1. There is a property, *K*, such that any individual possessing property *K* has a right to life, and there is a scientific law, *L*, to the effect that any organism possessing property *J* will, in the normal course of events, come to possess property *K* at some later time;
2. Given the relationship just described between property *J* and property *K*, it is seriously wrong to kill anything possessing property *J*.
3. If property *J* were not related to property *K* in the way indicated, the fact that an organism possessed property *J* would not make it seriously wrong to kill it.¹³

As Tooley states

In short, the question is whether there is a property, *J*, that makes it seriously wrong to kill something *only because J* stands in a certain causal relation to a second property, *K*, which is such that anything possessing that property *ipso facto* has a right to life.¹⁴

In other words, in the context of our present discussion, if we define property *J* as the potential or the capacity to develop human wisdom, *K* as human wisdom, and *L* as the normal genetic development from zygote to adult, then we are asking whether the possession of *J* does grants the individual moral standing.¹⁵

¹³Michael Tooley, "In Defence of Abortion and Infanticide," in *What is a Person?* ed. Michael F. Goodman (The Humana Press, 1988), 101.

¹⁴*Ibid.*, 101.

¹⁵Tooley is talking here specifically of the "right to life" whereas previously we had been talking of the more general case of the individual having moral standing. However, I think this difference does not threaten the general point here, for to talk of an individual having the right to life is to talk of the minimally basic condition necessary for moral standing.

We can present three arguments that suggest that there is no such property, *J*, that grants the individual moral standing as defined in these terms. First, there is an argument from Tooley himself.¹⁶

Let *C* be any type of causal process where there is some type of occurrence, *E*, such that processes of type *C* would possess no intrinsic moral significance were it not for the fact that they result in occurrences of type *E*.

Then:

The characteristic of being an act of intervening in a process of type *C* which prevents the occurrence of an outcome of type *E* makes an action wrong to precisely the same degree as does the characteristic of being an act of ensuring that a causal process of type *C*, which it was in one's power to initiate, does not get initiated.¹⁷

Tooley refers to this as the "moral symmetry principle with respect to action". As Tooley admits, the moral symmetry principle may not appear self-evident, since, other things being equal, we may think that it is worse to fail to prevent a harmful outcome than not to initiate a beneficial one. However, according to Tooley, if the two actions are identical with regard to the motivations of the relevant parties, then the actions are not morally distinct.¹⁸ For example, if I have been plotting to kill my business partner, then not warning him that a bomb is about to go off does not seem to be morally superior to shooting him myself.

If the moral symmetry principle is accepted, then it presents the following challenge

¹⁶It is perhaps more accurate to say that this is an argument that is based on one presented by Tooley. The specific argument that Tooley presents with respect to the moral symmetry principle is rather too length to be restated here. It can be found in Goodman, *What is a Person?*, 104-111.

¹⁷*Ibid.*, 104.

¹⁸Tooley writes "What the moral symmetry principle implies is...that, *other things being equal*, it is just a wrong to fail to save someone as it to kill someone. If one wants to test this principle against one's own intuitions, one has to be careful to select pairs of situations in which all other morally relevant factors-such as motivation, and risk to the agent-are equivalent. And I have suggested that when this is done, the moral symmetry principle is by no means counterintuitive." *Ibid.*, 106.

to supporters of the potentiality argument. Let us say that there were a drug or treatment *H* such that if it were given to the zygote/fetus which lacked *J*, the effect of *H* would be to grant the entity *J*. (For example, perhaps in the not too distant future, the injection of certain genetic material may prevent the development of certain genetic abnormalities). We can then ask whether, other things being equal, it would be worse *not* to give the fetus this additional genetic material so that it may have the potential to develop *K* (the capacity for human wisdom) than to interfere with the development of a normal fetus so that it is prevented from developing *K*. According to Tooley, if the moral symmetry principle is correct, then it is not counterintuitive to claim that it would be no worse to do the latter than the former. That is to say, we would be obliged to aid the fetus who presently lacks the necessary genetic material to develop *K*.

The conclusion that Tooley draws from this argument is that if the moral symmetry principle is correct, then since a supporter of the potentiality argument is not prepared to accept that we have a positive duty to benefit those entities that lack the relevant potential capacity, so he or she cannot consistently claim that it would be wrong to interfere with those entities that have the potential to develop such a capacity.¹⁹ However, we can also draw a broader conclusion that directly relates to the present discussion. For if the moral symmetry principle is correct, then we would have an equal obligation to assist those entities that lack the capacity to develop *J* as we would to not interfere with those entities that potentially possess *J*. But if this is the case, then the potential possession of *J* does not point to a clear moral difference between persons and nonpersons. For if the above analysis is correct, then one has an obligation to the entity, whether or not the entity potentially possesses property *J*. That is to say, if *J* is a necessary and sufficient condition of personhood, then our

¹⁹*Ibid.*, 107-108.

obligations to nonpersons may be equivalent to our obligations to persons.

Second, there is an objection presented by Feinberg that challenges the logic of the potentiality argument.²⁰

It is a logical error, some have charged, to deduce *actual* rights from merely *potential* (but not yet actual) qualification for those rights. What follows from potential qualification, it is said, is potential, not actual, rights; what entails actual rights is actual, not potential, qualification²¹.

Thus in our present discussion, the capacity for developing human wisdom grants the entity the capacity for developing moral standing, it does not grant the entity actual moral standing. To put this a different way, if one claims that the capacity for human wisdom is necessary and sufficient for moral standing, then this would seem to imply that those entities that do not possess this capacity do not have moral standing. But it is then claimed that, in addition to this first condition, the potential possession of human wisdom is sufficient for moral standing. The implication to be drawn is that the first condition is no longer necessary, for it is no longer necessary that one actually possess the capacity for human wisdom.

It is difficult to understand why the potential possession of morally relevant properties is equivalent to their actual possession, for if this were the case, why should their actual possession matter at all? If Prince Charles can raise taxes, a militia and pardon offenders as a potential King of England, then we do not need to complicate matters further by describing this authority in terms of the potential possession of rights that must be held actually. It would be simpler to describe these rights as the actual rights of potential Monarchs.

The problem with this approach, however, is that this would mean that in the present

²⁰Joel Feinberg, "The Potentiality Argument" in *Matters of Life and Death: New Introductory Essays in Moral Philosophy*, ed. Tom Regan, (New York: Random House, 1980), reprinted in *Moral Issues*, ed. Jan Narveson (Toronto: Oxford University Press, 1983).

²¹Narveson, *Moral Issues*, 235.

case, moral standing would have to be granted on the basis of the fact that the entity has the relevant genetic information and is a member of the species *homo sapiens*. But neither of these is morally relevant -- why should being a member of a particular species or having a particular genetic code be morally relevant? Thus if Noonan is obliged to give up the potentiality argument then there seems insufficient reason to regard these actual properties as morally important, for in and by themselves they have no moral relevance. (This is not to say that Noonan should abandon the potentiality argument, though this claim is made further on).

Third, we can criticize the potentiality argument on the grounds that it is "biased towards the future". According to Noonan, the entity has moral standing on the grounds that it has the relevant genetic material, and in the normal course of events will develop the capacities for human wisdom. But it is also the case that in the normal course of events, the entity will later lose this capacity for human wisdom, through the onset of old age, senility or death. Is a supporter of the potentiality argument prepared to say that an individual who loses the capacity for human wisdom is no longer a person? If the answer is that such an individual would not be a person, then we would need an explanation why the *potential* or *future* possession of certain properties grants moral standing whereas the *past* possession of such properties does not. Why should the temporal direction of the actual possession of these properties only work one way? Additionally, why pick out just one part of the "normal course of events" as relevant to moral standing, since in the normal course of events we are all going to lack the properties, actual or potential, that are necessary for moral standing? On the other hand, if the answer is that the past possession of these properties is sufficient for moral standing, then there will be very few individuals who do not meet the criteria for personhood, for it is now the case the potential, present and past possession of the relevant

properties is sufficient for personhood. This does not mean that such a view is unintelligible - one can, of course, claim that *all* members of the human species are persons. However, it does undermine the efficacy of any moral conclusion based on potentiality. For it is now the case that from cradle to grave, a human being is a person, and has the same moral status as any other human being.

There is a further concern about the importance that is placed on the "normal course of events." We should take notice of the fact that a number of entities will never have the capacity to develop human wisdom, though this will not be known until sometime after conception. For example, chromosomal analysis of the fetus indicates trisomy 21, suggesting Down's Syndrome, or lack of fetal brain development produces anencephaly. Furthermore, there are conditions such as Alzheimer's dementia that will destroy the capacity for human wisdom after it has once been possessed. If Noonan wants to maintain that the capacity for human wisdom is a necessary condition of humanity, and (as far as our present purpose is concerned) what distinguishes persons from nonpersons, then he cannot successfully maintain that the moment of conception is morally important, unless he takes it for granted that all human beings possess this capacity from the simple fact that they have the genetic information "which is the biological carrier of the possibility of human wisdom." This cannot be the case, for some will not have this genetic information at all, while others will lose this capacity for human wisdom at a later time. This is to say that the capacity for human wisdom is not a necessary condition of humanity.

Furthermore, even if we understand the concept of "human wisdom" in terms of rationality, reason and deliberation we still need an additional argument to explain why the possession of these properties is morally relevant. One possible argument would be that only those capable of reason and deliberation can be the possessors of rights, and hence have

moral standing. But this would be to grant Noonan an analysis that not only is not contained in his work, but is also at odds with his view of humanity.

The conclusion to be drawn here is that whatever the properties necessary and sufficient for personhood, these properties must be actual rather than potential properties.

Tooley

We turn now to the account of personhood presented by Tooley. Like Noonan, Tooley's discussion of the concept of person appears in the context of a broader discussion of the moral permissibility of abortion. However, as with our discussion of Noonan's account of personhood, we are not concerned with this issue here. First, we are interested in Tooley's concept of person in the abstract, second, we are interested in its application to questions of neurological dementia and deficit.

Tooley begins his discussion of the concept of person by asking

What relatively permanent, non-potential properties, possibly in conjunction with other, less permanent features of an entity, make it intrinsically wrong to destroy an entity, and do so independently of its intrinsic value?²²

Thus the yardstick by which we determine whether a given property is a property of personhood, is whether this property (in conjunction with other less permanent features of the entity) makes it intrinsically wrong to destroy the entity. "Person" is defined as "an entity that possesses at least one of those enduring, non-potential properties." This is what Tooley means when he claims that our conditions of personhood "should be guided by moral considerations."²³

²²Michael Tooley, *Abortion and Infanticide*, (Oxford: Clarendon Press, 1983), 87.

²³At the beginning of this chapter we discussed the possibility of describing the concept of person in purely descriptive terms, and our conclusion then was that such a task would be difficult, since it is difficult to distinguish the descriptive and the prescriptive notions. This

Our task is to determine what are these enduring, non-potential properties that constitute personhood and that make it intrinsically wrong to destroy the entity.

To begin with, the matter is relatively straightforward, since Tooley presents us with the following argument

- (1) The concept of a right is such that an individual cannot have a right at time *t* to continued existence unless the individual is such that it can be in its interest at time *t* that it continue to exist.²⁴
- (2) The continued existence of a given subject of consciousness cannot be in that individual's interest at time *t* unless *either* that individual has a desire, at time *t*, to continue to exist as a subject of consciousness, *or* that individual can have desires at other times.
- (3) An individual cannot have a desire to continue to exist as a subject of consciousness unless it possesses the concept of a continuing self or mental substance.
- (4) An individual existing at one time cannot have desires at other times unless there is at least one time at which it possesses the concept of a continuing self or mental substance.

can seen from Tooley contention above that the purely descriptive concept of person is guided by moral considerations.

²⁴The right that Tooley is concerned with here is the right to continued existence (more generally referred to as the "right to life," Tooley's preference for the former phrase is that it permits non organic entities to be persons). As discussed in the previous chapter, the right that is the focus of this work is the right to refuse or not to receive life-sustaining treatment. It is important to ask how are these two rights are related to the right to continued existence, for if there were little relation then we should be skeptical of using any conclusions drawn from our discussion of Tooley's account to draw subsequent conclusions regarding the present work. However, I think that there is a clear relation between the right to continued existence and the above rights on Tooley's account. Since it is the case that only those entities that possess the concept of a continuing self possess the right to continued existence, a right to refuse or not to receive life-sustaining treatment is itself dependant on the possession of the concept of a continuing self, since some one who asserts or claims this right is thereby claiming that she does not want to continue to exist in that fashion. This is to say that the class of entities that possess the right to continued existence will include the class of entities that have the right to refuse or not to receive life-sustaining treatment. The remaining question is to ask what else the entity must possess in order to possess the latter right in addition to the former. We will return to this point later on in the chapter.

Therefore:

(5) An individual cannot have a right to continued existence unless there is at least one time at which it possesses the concept of a continuing self or mental substance.²⁵

The conclusion to be drawn is that a necessary (and perhaps sufficient) condition of the concept of person is that the entity possess at some time the concept of a continuing self of mental substance. This condition can be regarded as one of the enduring, non-potential properties that make it intrinsically wrong to destroy the entity.²⁶

An immediate conclusion follows from this strategy of formulating the conditions of personhood with respect to the right to continued existence. By adopting this strategy Tooley would seem to have thereby undermined the claims of non-persons to moral standing. For if the right to continued existence is necessary and sufficient in order for it to be intrinsically wrong to destroy an entity, and person is defined in terms of the right to continue to exist, then a nonperson does not have a right to continued existence. That is to say, nonpersons have no intrinsic moral standing. In other words, other things being equal, we should not be concerned for the interests of nonpersons.

This conclusion rests on Tooley's success in showing that there is at least one enduring non-potential property that makes it intrinsically wrong to destroy the entity in question, and that we can with reasonable success identify which entities possess this property and which do not.

According to Tooley the above argument involves three central claims. To facilitate discussion I have assigned letters to these claims.

²⁵Ibid., 121.

²⁶It may be the case that this is the *only* enduring non-potential property that grants the entity the right to continued existence. For whatever reason, Tooley does not mention any other properties apart from this one.

(A) First, there is the contention that it follows from a sound analysis of rights that there is a conceptual connection between specific rights and the capacities for specific interests. (B) Second, there is the claim that interests presuppose desires. (C) Third, there is the claim that the unification of mental states existing at different times, including desires, into a single, continuing subject of consciousness, requires that the individual possesses, at some time, the concept of a continuing self or mental substance.²⁷

It is the third of these claims that contains the core of Tooley's concept of person, and we will discuss this claim in detail. However, we need to begin by examining the first two claims.

The argument that Tooley presents in support of the first premise, and in defence of (A), is the argument presented by Feinberg with regard to the connection between rights and interests that we discussed in the previous chapter (pp **19-25). To recall, Feinberg claims that

...the sort of beings who *can* have rights are precisely those who have (or can have) interests. I have come to this tentative conclusion for two reasons: (1) because a right holder must be capable of being represented and it is impossible to represent a being that has no interests, and (2) because a right holder must be capable of being a beneficiary in his own person, and a being without interests is a being that is incapable of being harmed or benefitted, having no good or "sake" of its own.²⁸

This is to talk of the general case -- that a right holder must be the possessor of interests. If the entity has no interests at all, that is to say, that he or she is incapable of being benefitted or harmed by any action or event, then to say that such an entity might have a right to something is questionable. For if the entity is incapable of being benefitted or harmed what is the substance of a claim that such an entity has a right -- why would the entity be able to make a claim for *x* if *x* could never be beneficial to the entity?

For his part, Tooley is concerned for the particular case; whether the possession of a particular right is dependant on the possession of a particular interest. If Feinberg's

²⁷Ibid., 122.

²⁸Joel Feinberg, "The Rights of Animals and Unborn Generations" in *Philosophy and Environmental Crisis* ed. William Blackstone (Athens: University of Georgia Press, 1974).

account of the connection between rights and interests is correct in the general case, then the following more specific principle, the *particular-interests principle* is also correct.

It is a conceptual truth that an entity cannot have a particular right R unless it is at least capable of having some interest I which is furthered by its having right R.²⁹

Tooley argues that the plausibility of this principle can be seen if we compare the killing of a newborn kitten with the act of torturing one for five minutes. The second act seems indefensible on the grounds that the kitten has an interest in not being tortured, an interest which is derived from its capacity to feel pain -- the kitten can be clearly harmed by such action. But it seems less clear that the first act is indefensible, for it is doubtful that the kitten has an interest in continued existence. To say that an entity has "an interest in continued existence" means that the entity possesses the concept of a continuing subject of experiences, and that the entity is aware of itself as a subject of experiences. No matter how much we like kittens, no kitten possesses the above concept, and thus on this analysis, it would be wrong to claim that any kitten has an interest in continued existence. T h i s analysis is by no means self-evident and depends for its plausibility on the establishment of the claim that the right to life is dependant on the possession of an interest in continued existence. This, of course, is the basis for Tooley's concept of person, and we will discuss this claim in detail shortly.³⁰

The second claim behind Tooley's main argument is that interests presuppose desires

(B). This is the second premise of the above argument

²⁹Tooley, *Abortion and Infanticide*, 99.

³⁰It is worth pointing out that if in order to have the right to continued existence the entity must possess the concept of a subject of experiences, then few children possess the right to continued existence before the age of two or three. Tooley fully endorses this conclusion, since his aim is to argue for the moral permissibility of both abortion and infanticide.

(2) The continued existence of a given subject of consciousness cannot be in that individual's interest at time *t* unless *either* that individual has a desire, at time *t*, to continue to exist as a subject of consciousness, *or* that individual can have desires at other times.

Thus the entity cannot have a right to continued existence unless it has an interest in continued existence, and this interest presupposes that the entity is capable of having a desire to continue to exist. What is the basis for this claim? To understand this we need to be clear as to what is involved in having a desire.

According to Tooley, there are two important points about desires that are relevant to the connection between desires and interests. First, desires can only be attributed to things that have experiences, that are conscious. If the entity is incapable of having experiences then the ascription of beliefs and desires to the entity is, at best, metaphorical³¹. For example, it would be false to claim that a patient in a PVS desires that life-sustaining treatment be ceased, for such individuals are simply not capable of having desires.

Second, desires involve *propositional attitudes*. Propositional attitudes are defined by Paul Churchland in the following way

Let us now direct our attention to certain of those mental states themselves -- to thoughts, and beliefs, and fears, for example -- for each such state also has a 'meaning', a propositional 'content'. One has

the *thought* that [children are marvelous],

the *belief* that [humans have great potential] and

the *fear* that [civilization will suffer another Dark Age].

Such states are called *propositional attitudes*, because each expresses a distinct 'attitude' toward a specific proposition. In the technical vocabulary of philosophers, such states are said to display *intentionality*, in that they 'intend' something or 'point

³¹There are, of course, those who say that the ascription of beliefs and desires to any individual is, at best, metaphorical. This point is discussed by Paul Churchland in "Eliminative Materialism and the Propositional Attitudes," *The Journal of Philosophy* LXXVII, 2 (February 1981): 67-90.

to' something beyond themselves: they 'intend', or point to, children, humans, and civilization.³²

According to Tooley, since desires involve propositional attitudes

...an individual cannot have a desire unless he understands the proposition that he desires is true. Understanding a proposition, in turn, presupposes possession of the concepts involved in it. To have desires, then, one must possess the related concepts.³³

This is an important claim, for according to Tooley, beliefs, thoughts, fears (etc.) can only be had by individuals capable of relatively complex cognitive states. Since beliefs and desires involve propositional attitudes, the individual cannot have a desire for, or a belief in *x* unless the individual can understand the truth-conditions of the proposition involved. For example, if Mary has a desire for ice-cream then she wishes that the proposition "Mary is eating ice-cream" were true.

What are the related concepts involved in the desire to continue to exist? In order for somebody to have this desire he or she must understand what it would be for him or her to continue to exist. This involves having the concept of oneself as a continuing subject of experiences, and being aware of oneself as a subject of experiences. If Mary has a desire for continued existence then (let's say she wants to live to see her grandchildren go to university) this means she recognizes herself as a discrete individual and the subject of experiences, and wants to continue as the same subject of experiences in the future - she wants Mary to continue to exist.

However, although it is helpful to talk in terms of propositional attitudes in order to understand the intentionality of beliefs, thoughts (etc.,) and thereby clarify what someone

³²Paul M. Churchland, *Matter and Consciousness*, rev. ed. (Cambridge: The MIT Press, A Bradford Book, 1988), 63.

³³Tooley, *Abortion and Infanticide*, 104.

means when he or she has a desire for something, it does not seem that helpful in explaining just what it means to have a desire. Think, for example, of a patient, John, who is suffering from end-stage liver cancer and who desires pain medication. Have we come to a clearer understanding of this desire by reformulating it as "John desires that the proposition "John is receiving pain medication" were true"? I think it would be true to say that John could not care one way or the other about the truth of this or any other proposition. And the reason would be that although it might be of some use to talk in terms of propositional attitudes if our purpose is to understand the semantics of "John desires that x ," such a propositional analysis gives us a rather odd picture of desires. One might claim that desires are rather more "basic" than as described here. A person might have a desire for something and it not be the case (as far as a description of the actual desire is concerned) that this desire can be analyzed in relation to a propositional attitude. For example, we can imagine the desires of children to build sandcastles, or whatever. To attempt to understand these desires in terms of the truth or falsity of propositions seems to miss the point.

A possible implication of the view that desires involve propositional attitudes, and can be had only by individuals capable of relatively complex cognitive states, is that Tooley may be mistaken in claiming above that the kitten has an interest in not being tortured. For if interests presuppose desires, and desires involve propositional attitudes, it would be controversial to claim that the kitten is capable of having the sufficiently complex cognitive states that Tooley claims are necessary for the having of desires. In other words, if Tooley wants to claim that the kitten has an interest in not being tortured, then he must grant it the desire not to be tortured and the appropriate propositional attitude, and it is clear that he does not want to do this.

The general motivating point here is that an individual can have a desire for

something and it not be the case that he or she "understands the related concepts". Recall the case of Joseph Sakiewicz mentioned in the previous chapter. In that case the question was whether to perform chemotherapy on a profoundly retarded sixty-seven year old man with the I.Q. of 10 suffering from leukaemia. In such a case, Tooley's analysis would seem relevant if the appropriate desire is the desire that chemotherapy be performed. For without doubt, Joseph Sakiewicz could not understand the related concepts involved in such a treatment regimen. But what if the appropriate desire is a desire for relief of pain - what are the "related concepts" involved in such a desire? One answer is that such a desire does involve certain concepts, for example the concept of relief of pain. In this case it would be plausible to claim that even Joseph Sakiewicz could desire the relief from pain (though not in propositional form). He may simply wish for the pain to go away even though lacks the ability to describe or even conceptualize this desire. However, this seems to be a considerably weaker account of what is involved in a desire than that originally presented. Another answer is to claim that the related concepts involved in the desire for the relief from pain involve more complex propositional attitudes. However, this would seem to imply that Joseph Sakiweciz cannot desire the cessation of pain, a view that seems controversial.

The conclusion to be drawn here is that Tooley's account of the connection between interests and desires is problematic. First, it seems that Tooley is demanding too sophisticated a level of cognitive or even intellectual capacity in order for the entity to possess interests. The above examples suggest that an individual can have a desire for, and an interest in the relief of pain even though the individual may not be able to possess the relevant concepts. Furthermore, by Tooley's own admission, we can ascribe interests to entities for which it would be questionable to ascribe a high level of cognitive capacity. Thus Tooley appears prepared to grant the kitten an interest in not being tortured (though not the interest in

continued existence) on the grounds that it has an interest in not suffering pain. But Tooley can only consistently make this claim if he is either prepared to grant the kitten the appropriate desire, with the accompanying propositional attitude, or prepared to deny that the kitten has an interest in not suffering pain.

We have now looked at the first two claims that form the basis for Tooley's argument. Our discussion of interests and desires has shown why Tooley thinks that in order for the entity to have a desire to continue to exist presupposes that the entity possesses the concept of a continuing subject of experiences. Tooley summarizes these points this way.

Consider now the desire to continue to exist. In view of the first point about desires, anything that has this desire will have to be subject of experiences. And in view of the second, it will have to understand a certain proposition. Part of what is involved in that proposition is the concept of something's continuing to exist. This idea, in turn, involves the concept of a temporal ordering of events, together with the idea of a thing existing at one time being identical with a thing existing at another, later time. In order to have a desire that anything continue to exist, then, one must possess these concepts. Here, however, we are concerned with a desire that a subject of experiences continue to exist, and continue to exist *as* a subject of experiences. One must therefore have the concept of a continuing subject of experiences. Finally, the desire is not a desire that just any subject of experiences continue to exist: it is that *this* subject of experiences continue to exist. And it would seem that, in order to have this desire, one must be in some sense aware of the subject of experiences in question. *To sum up, then, to have a desire for one's continued existence presupposes that one possesses the concept of a continuing subject of experiences, and that one is aware of oneself as a subject of experiences* (emphasis added).³⁴

We need to be cautious here for there is a difference between the above conclusion and the second premise of Tooley's main argument that we are discussing. On the one hand, in the above quotation Tooley is claiming that "a desire for one's continued existence presupposes that one possesses the concept of a continuing self or mental substance." On the other hand, in the second premise he is claiming that the continued existence of the individual cannot be in that individual's interest

³⁴Ibid., 104-5.

...unless *either* that individual has a desire, at time *t*, to continue to exist as a subject of consciousness, *or* that the individual can have desires at other times.

The problem is that whereas Tooley has presented some defence of the first claim, little argument has been presented in support of the claim that interests presuppose desires. Up till now, the argument has consisted of the claims that desires involve the capacity for consciousness, and should be understood in terms of propositional attitudes. In fact, the claim that interests presuppose desires seems implausible; I have an interest in keeping fit and healthy but (at present) I have no desire to do so.

The difficulty here is that the parts of the above disjunction (the second premise) are rather different. Compare the following claims:

(i) *X* cannot be in an individual's interest at time *t* unless the individual desires *x* at time *t*,

and

(ii) *X* cannot be in an individual's interest at time *t* unless the individual is capable of desiring *x* (at any time).

Our discussion seems to have shown that (i) cannot be the case, for something can be in my interest at the present moment even though I do not have the appropriate desire. Nor is it clear (to me) why Tooley wants to defend (i), for the validity of Tooley's main argument does not appear to depend on the truth of (i). On the other hand, (ii) seems much more plausible -- it seems clear that it would be difficult to maintain that *x* could be in an individual's interest if the individual was incapable of desiring *x* at any time. The reason being that if the individual were incapable of desiring *x* this would mean that, on Tooley's analysis, since the ability to have desires is linked to the capacity for consciousness, an individual who was incapable of desiring at any time must lack the appropriate level of consciousness and could not understand the related concepts.

We can put this point a different way. If we recall the interest in continued existence, we can ask why this particular interest should presuppose a desire to continue to exist. The basis for this claim, and perhaps the reason why Tooley thinks that interests presuppose desires, is that if one *never* had the desire to continue to exist (perhaps because the individual never wanted to go on living) then it would be difficult to claim that the individual had an interest in continued existence.³⁵

The conclusion to be drawn here is that Tooley may have to amend the second premise above to include only the second part of the disjunction. This point does not seriously challenge Tooley's main argument, for his conclusion that an individual cannot have a desire for continued existence unless he or she possesses the concept of a continuing self does seem to depend on whether or not interests presuppose desires. Furthermore, the main argument still seems to work if we amend the second premise as suggested.

We have now reached the final of the three central claims that form the basis for Tooley's argument for the concept of person. This is claim (C) that

...the unification of mental states existing at different times, including desires, into a single, continuing subject of consciousness, requires that the individual possesses, at some time, the concept of a continuing self or mental substance.³⁶

Once we have discussed this claim we will be in a position to determine whether Tooley is correct in his conclusion that

An individual cannot have a right to continued existence unless there is at least one time at which it possesses the concept of a continuing self or mental substance.³⁷

³⁵A relevant case here would be that of Elizabeth Bouvia. *Bouvia v. Los Angeles County*, 195 Cal. App. 3d 1075, 241 Cal. Rptr. 239 (Ct. App. 1987) and *Bouvia v. Super. Ct. of Los Angeles County*, 179 Cal. App. 3d 1127, 225 Cal. Rptr. 297 (Ct. App. 1986).

³⁶Tooley, *Abortion and Infanticide*, 122.

³⁷*Ibid.*, 121.

This is to say that since Tooley regards the class of persons as defined by all those entities that have a right to continued existence, a necessary and perhaps sufficient condition of personhood is that the individual possesses the concept of a continuing self or mental substance.³⁸

We now turn to a discussion of (C). The question at hand is twofold. First, there is the question of why it is important that mental states, including desires, that occur at different times be attributed to a single, continuing subject of consciousness. Tooley's answer is that it is only by so organizing mental states that they become morally relevant. Second, there is the question of why the attribution of temporally diverse mental states to a single, continuing subject of consciousness requires that the individual possesses the concept of a continuing self or mental substance.

To begin the discussion of (C), Tooley asks how the term "person" should be defined. He considers four "of the most plausible" answers to that question:

- (1) A person is a subject of non-momentary interests;
- (2) A person is an entity that possesses rationality;
- (3) A person is an entity that is capable of action;
- (4) A person is an entity that possesses self-consciousness.³⁹

Tooley regards these potential definitions as candidates for necessary or sufficient conditions of the concept of person. What do we mean when we talk of the concept of person as

³⁸To be precise, Tooley's conclusion is somewhat more cautious. He writes "Again, we are assuming, tentatively, that the class of persons either coincides with, or is a subclass of, the class of things possessing a right to life. Or, more precisely, we are assuming that the class of persons either coincides with, or is a subclass of, the class of things with a *right to continued existence*," *Abortion and Infanticide*, 102. It is an interesting question why Tooley allows for the possibility of there being entities that may possess the right to continued existence yet not be persons, since this would undermine the strategy of defining the concept of person in terms of the possession of enduring non-potential properties, and which grant the entity the right to continued existence.

³⁹*Ibid.*, 123.

the "subject of non-momentary interests." According to Tooley

This is, very roughly, the concept of an entity that is capable of having a variety of desires, both at a given time, and at other times, all of which are interrelated, or "unified" in certain ways.⁴⁰

Another way that we could describe a subject of non-momentary interests is in terms of

the unification of mental states existing at different times, including desires, into a single, continuing subject of consciousness

which is the first part of (C).

We can ask what conditions the entity must satisfy in order to be the subject of non-momentary desires. According to Tooley, in order to answer this question we have to first answer the following questions: First, what is the relevant concept of a desire in this context? Second, what are the appropriate interrelations among desires?

In answer to the first question, Tooley claims that the relevant concept of desire is that a desire is not simply a state that underlies and serves as an explanation of behavior. Rather, the relevant concept of desire must be the concept of a state "that is either potentially represented in consciousness, or directed at some state of consciousness."

Tooley's defends this claim in the following way. Let us describe a desire in behaviorist terms as a state that underlies and serves as an explanation of behavior. According to such a description it would be plausible to talk of a heat-seeking missile as having a "desire" to locate its target, or to talk of a dog having a "desire" to gambol in the park. Is such a description adequate? The problem with such an account of desire is that if this is all that we mean by a desire, then it is difficult to see why frustrating the missile's "desire" is any worse than frustrating John's desire to sit by the fire. Generally, we would think that John's desires as being more important than those of the missile or the dog. The

⁴⁰Tbid., 123.

reason for this is that John's desires are tied in some part to consciousness. What distinguishes John's desire to sit by the fire from the "desire" of the heat-seeking missile is that there is rather more involved in John's desire than simply being an explanation of his behavior.

A person who wants to be near something warm is disposed not only to act in certain ways, but to have certain thoughts, such as the thought that he would like to be near something warm.⁴¹

On the other hand, even though we can say that the heat-seeking missile "desires" to find the warmest spot, we cannot claim that it has any thoughts about the matter. We should reject, therefore, the claim that desires are simply states that underlie and serve as explanations of behavior.

To put the matter another way, Tooley is claiming that the justification of ascribing desires to an entity cannot be simply in terms of the instrumental value of doing so. Rather, beliefs, thoughts and any other intentional terms should only be ascribed to those entities that possess consciousness.⁴² In this way we have reason to regard the frustration of John's desire to help the poor and needy as being worse than the frustration of the "desire" of the heat-seeking missile to locate its target.

Thus we are correct in claiming that it would be worse to frustrate John's desire than the "desire" of the missile since the difference between desires and "desires" is that the former are not assignable to subjects of consciousness.

The argument, in short, is this. If "actions", and the "desires" that underlie them, are to be assignable to one subject of consciousness, rather than to another, or none at

⁴¹Ibid., 124.

⁴²Thus Tooley's account is at odds with the account presented by Dennett. For Dennett, a sufficient reason for ascribing intentional states to an entity is if it helps us to predict the entity's behavior. See, for example, *The Intentional Stance*, (Cambridge: The MIT Press, A Bradford Book, 1987).

all, one of two things must be the case. Either the object at which the "action" and the underlying "desire" are directed must itself be a present state of consciousness, or the "action" or the underlying "desire" must be at least potentially represented by states of consciousness belonging to the subject in question. The difference between desires and "desires" is that the former are assignable to subjects of consciousness while the latter are not, and it is this, I am suggesting, that makes desires morally significant in a way that "desires" are not.⁴³

Thus for Tooley it is not sufficient that the entity in question have the capacity for consciousness. What is needed in addition is that the entity can form a "representation" of the object of the desire or of the state of consciousness. In order for an entity to have desires, the entity needs to have the capacity for some potentially conscious 'representation' of the object desired.⁴⁴

If what is liked or disliked is itself a present state of consciousness, it can be viewed as 'representing' itself. If what is liked or disliked is an object presently being perceived, the perception can be taken as 'representing' the object. If what is liked or disliked is neither a present state of consciousness nor something presently perceived, then something else, which does involve consciousness, will be needed to 'represent' the object desired. Thoughts or dispositions to have thoughts, will be needed in this case, if the state is to be a morally significant one.⁴⁵

The conclusion that Tooley draws here is that

An entity that has desires must, therefore, also be an entity that has experiences. It must be a subject of consciousness. One can, therefore, assign all the desires existing

⁴³Tooley, *Abortion and Infanticide*, 125.

⁴⁴In a similar vein, one could claim that all conscious experience involves "second-order mental states"; that if I am aware of the tree in front of me then I am aware that I am aware. Peter Carruthers draws a moral conclusion from this point. He argues: i) It is only wrong to cause harm to those entities that are capable of second order mental states. ii) The appropriate way to describe the experiences of those entities that are incapable of second-order mental states, (for example, animals), is as *non-conscious experience*. Therefore iii), since the experiences of animals lack phenomenological content, we should not be concerned to avoid causing animals "pain and suffering". Carruthers argument is open to the same type of objection as is raised against Tooley's; that an animal may desire food or yelp with pain even though it cannot think to itself that it is in pain. Peter Carruthers, "Brute Experience," *Journal of Philosophy* 86 (1989): 258-269.

⁴⁵Tooley, *Abortion and Infanticide*, 125.

at different times to different sets, based upon whether or not they belong to the same subject of consciousness.⁴⁶

It is not clear that Tooley can draw this conclusion. The above discussion of desires has shown that if we want to have a notion of desire that permits us to regard the frustration of some desires as worse than the frustration of others, we need to differentiate desires from "desires", and claim that "an entity that has desires must, therefore, also be an entity that has experiences". But from this it does not follow that an entity capable of desires is a subject of consciousness, for it depends on how we understand "subject of consciousness." If by this term we simply mean an entity that has experiences then dogs and cats can be subjects of consciousness, which is perhaps too permissive given Tooley's general purpose. On the other hand, if by this term we mean an entity that possesses the concept of a continuing self or mental substance, this would exclude cats and dogs, but would not follow from the previous discussion of desires and "desires" above. In the discussion that follows we see how Tooley attempts to link a subject of consciousness with the possession of the concept of a continuing self, but we should be wary of concluding that all and only those entities who possess the concept of a continuing self can be subjects of consciousness. For it is not clear that the ensuing discussion of desire, and of the necessity of the subject being capable of having a "representation" of the object desired warrants this conclusion.

We turn now to examine what are the appropriate interrelations amongst desires, the second point about desires. This question has moral significance for Tooley. For to ask what are the *appropriate* interrelations amongst desires is to be concerned with the question of whether there is any interrelation that is more morally significant than any other interrelation. The conclusion that Tooley intends to draw is that the most morally significant interrelation

⁴⁶Ibid., 126.

of desires is if they are ascribed to one continuing subject of consciousness.

As Tooley points out, the question of the moral significance of the allocation of desires is a complicated one. Stated simply, the question being discussed is "Can it make a moral difference whether or not two desires existing at a given time belong to the same subject of consciousness? That is to say, does this relation among desires enter into the formulation of basic moral principles?" As Tooley suggests, different ethical theories will provide different answers to this question:

On one version of utilitarianism, it is claimed that the right action is always the one that will result in the greatest satisfaction of desires. If this view is right, then no assignment of desires to different sets can be morally significant in itself. For the division of desires into different sets in no way affects the sum total of satisfaction associated with any given action.⁴⁷

Contrariwise, on a deontological theory concerned with justice, responsibility and desert, it will be important that a particular person's desires are satisfied, since our obligations are to persons rather than to maximize the good irrespective of its distribution amongst persons.⁴⁸

For the sake of argument, let's follow Tooley and assume that the greater moral distribution of desires in moral terms is if they are distributed according to persons rather than simply in terms of the good.

Let us assume, then, that there are some relations among desires that serve to unify them into morally significant groups. Since, as we have seen, desires must belong to subjects of consciousness, the unification of desires into morally significant groups will also result in the construction of some sort of enduring subjects of mental states out of the momentary subjects of consciousness to which particular desires belong. I shall refer to the individuals that result from such a morally significant unification of desires into groups as subjects of non-momentary interests.⁴⁹

⁴⁷*Ibid.*, 126.

⁴⁸Parfit discusses this matter in "Later Selves and Moral Principles," in *Philosophy and Personal Relations*, ed. Alan Montefiore, (Montreal, Canada: McGill-Queen's University Press, 1973), 137-169.

⁴⁹Tooley, *Abortion and Infanticide*, 127.

There are a number of ways that we can attempt to unify different mental states into morally significant groups, so as to ascribe them to one continuing subject of consciousness.

We begin with the following suggestion.

The first suggestion is that desires existing at different times are to be assigned to distinct, morally significant groups on the basis of whether they belong to different mental substances...this comes down to saying that desires are to be viewed as belonging to a single, morally significant group if and only if they belong to mental states that stand in certain relations of causal dependance.⁵⁰

The important question here is, what conditions must be met in order to ascribe temporally diverse desires to one continuing subject? The first suggestion is that the relevant necessary and sufficient condition is that the desires belong to the same mental substance.⁵¹ But it could be argued that being the same mental substance is neither necessary nor sufficient for being a continuing subject of consciousness. Let us imagine two temporally distinct desires along the following lines: At $t1$ Mary has the desire that a particular treatment be performed. At $t2$ Mary has the desire that the same treatment be ceased. So described, it seems that these desires are morally significant. That is to say, other things being equal, the desires belong to one continuing subject of consciousness. However, let us imagine that between these two times, Mary loses all of her memories and other psychological make-up. In this case, we would be reluctant to talk of a "continuing subject of consciousness" from $t1$ through $t2$. That is to say that continuity of mental substance is neither necessary nor sufficient for a continuing subject of consciousness.

A second approach is to talk in terms of a continuity of psychological features that are causally connected. Thus Mary's desire for treatment to be ceased at $t2$ is in some way

⁵⁰*Ibid.*, 129.

⁵¹This claim is discussed and subsequently rejected by Locke. *Essay Concerning Human Understanding*, Bk. II, Ch. 3, Sect. 9.

caused by her desire that the treatment be initiated at *t1* -- the second desire is caused by the first. But as Tooley points out, the added condition of causal dependance is neither necessary nor sufficient for a continuing subject of consciousness. He asks us to imagine that Billy Graham is reprogrammed as Bertrand Russell and then presents the following questions.

The first question is whether the satisfaction of desires associated with the individual who exists after the reprogramming is in any way in Billy Graham's interest. Most people would hold, I believe, that it is not. The second question, then, is whether the desires before and after reprogramming belong to the same mental substance....If they do, one will have a reason for rejecting the contention that desires are interrelated in morally significant ways if they belong to the same mental substance.⁵²

Thus it is neither necessary nor sufficient for being a continuing subject of consciousness that there is continuity of mental substance or that desires at different times are causally related.

As Tooley says,

Therefore it cannot be the relation of belonging to the same mental substance which serves to tie desires together in morally significant way we are considering here.⁵³

If neither of the above approaches are satisfactory, we need to determine what needs to be added to the above analysis in order that we can ascribe different beliefs to the same continuing subject of consciousness. Tooley's answer is that there must be a sufficient degree (whatever that might be) of *psychological continuity*.

What seems to matter is that later stages of a mental substance 'flow' from earlier ones in a way that involves more than mere causal dependance of later stages upon earlier ones. There must also be some relation between the contents of the beliefs, desires, and attitudes at different times. The changes that the individual must, apparently, be ones that would be expected in an individual with the experiences and psychological attitudes in question.⁵⁴

However, although such continuity is necessary to tie desires together in a morally significant

⁵²Tooley, *Abortion and Infanticide*, 130.

⁵³*Ibid.*, 130.

⁵⁴*Ibid.*, 132.

way, it is not sufficient. What has to be added is that

...there must also be *recognition of the continuity* by the enduring mental substance in question. Later mental states must be seen both as causally dependant upon, and as 'flowing from', earlier mental states.⁵⁵

Tooley concludes

If this is right, we have one way in which desires existing at different times can be interrelated in a morally significant way: they can be desires belonging to an enduring mental substance that exhibits a certain psychological continuity in the alteration of its properties over time, and where the mental substance is aware, via memory, of both the causal dependance and the psychological continuity linking earlier and later stages.⁵⁶

It would be helpful to summarize the above discussion. To recall, we are attempting to determine the plausibility of the claim (C) that

...the unification of mental states existing at different times, including desires, into a single, continuing subject of consciousness, requires that the individual possesses, at some time, the concept of a continuing self or mental substance.

We began by presenting a definition of person as "a subject of non-momentary interests" and asked what conditions the entity must satisfy in order to be the subject of non-momentary desires. This involved the discussion of the following points: First, what is the relevant concept of a desire in this context? Second, what are the appropriate interrelations among desires? We concluded that in order to have desires the entity must possess consciousness and be capable of thinking about that desire. We then looked at the question of what are the appropriate interrelations among desires. The ensuing discussion led to the conclusion that desires are interrelated in an appropriate way (which is to say a morally significant way) if the desires can be attributed to a single, continuing subject of consciousness. This conclusion led, in turn, to a discussion of what constitutes a single,

⁵⁵Ibid., 132.

⁵⁶Ibid., 132.

continuing subject of consciousness. The conclusion is that there must be

...a certain psychological continuity in the alteration of its properties over time, and where the mental substance is aware, via memory, of both the causal dependance and the psychological continuity linking earlier and later stages.⁵⁷

Stated more fully, Tooley claims that

First, sameness of mental substance plus psychological continuity is a necessary, but not a sufficient, condition, for desires existing at different times to be viewed as part of a single, unified non-momentary interest. Second, there are at least three other conditions that, taken together with sameness of mental substance plus psychological continuity are sufficient:

- (1) The desires are linked together by memory beliefs concerning the existence of earlier desires associated with the same mental substance;
- (2) Earlier desires are accompanied by beliefs that there will be later desires belonging to the same mental substance;
- (3) Earlier desires are accompanied by desires for the existence of later desires belonging to the same mental substance.⁵⁸

This provides us with an analysis of what Tooley means by a "subject of non-momentary interests," and since we began our discussion of (C) by presenting a definition of person as a subject of non-momentary interests, we now have an analysis of a plausible candidate for Tooley's concept of person.

We are now in a position to ask whether Tooley has convinced us of the plausibility of (C); that being a subject of non-momentary interests requires that the individual possesses, at some time, the concept of a continuing self or mental substance.

If (C) were true, then there would be no case where an entity could be a subject of non-momentary interests yet not possess at some time the concept of a continuing self or mental substance. This to say that the concept of a continuing self or mental substance is a

⁵⁷Ibid., 132.

⁵⁸Ibid., 134.

necessary condition of being the subject of non-momentary interests.

What we are asking here is whether mental states existing at different times can be unified into a single subject of consciousness if the entity never has, or can have, the concept of a continuing self or mental substance. To answer to this question we must recall what Tooley claimed about the connection between desires and subjects of consciousness, and the interrelation of desires based on a notion of psychological continuity.

Tooley claimed that in order to have desires the entity must be capable of having experiences, since it was only by having this requirement that we could distinguish desires from "desires." He concluded that to have desires the entity must be a subject of consciousness. His argument for this was that desires involve some form of "representation" and thus we should ascribe desires only to those entities capable of having representations of desires. Two objections were raised against this argument. First, we questioned whether in order to have a desire for x the entity must be capable of having a representation of x . Second, an objection was raised against this argument in that the term "subject of consciousness" is vague; it could mean either any entity that has consciousness, or only those entities that are "subjects" that have a concept of self. Tooley needs to support the second of these alternatives on the grounds that only by doing so can he restrict the possession of desires to those entities that possess the concept of continuing self. However, if Tooley opts for this alternative the account is question-begging. For he is attempting to argue that only those possessing a concept of self are capable of conscious experience. Thus Tooley must be claiming simply that any entity that is conscious can have desires.

Tooley's conclusion with regard to the interrelation of desires was that in order for diverse desires (or any other mental state) to be unified, some amount of psychological continuity is required. To recall, Tooley contends that continuity of mental substance and

causal dependance of mental states is not sufficient for mental states to be unified. What is required in addition is that "later mental states must be seen both as causally dependant upon, and as "flowing from", earlier mental states," that is to say there must be "some relation between the contents of the beliefs, desires, and attitudes at different times." Thus

...one way in which desires existing at different times can be interrelated...: they can be desires belonging to an enduring mental substance that exhibits a certain psychological continuity in the alteration of its properties over time, and the mental substance is aware, via memory, of both causal dependance and the psychological continuity linking earlier and later stages.⁵⁹

To return to our main question, we are asking whether such an interrelation of mental states requires the concept of self. That is to say, we need to ask whether there are other ways in which desires existing at different times can be interrelated that do not require the concept of self. Tooley thinks that there is another way. He asks,

...what is one to say about the case of an enduring mental substance that has neither any memory of the past, nor any desire about the future, but that exhibits psychological continuity, and that has a belief that it will continue to exist and will display traits that will flow from its present ones. Is such a mental substance to be viewed as a subject of non-momentary interests?⁶⁰

Tooley's reply to this question is that we should ascribe non-momentary interests to such an entity, yet he finds it "very hard to see what sort of argument could be offered in support of this view." In other words, there will be cases where an individual may lack the appropriate desires and memories yet where it would, nevertheless, be legitimate to regard such an individual as a subject of non-momentary interests. As Tooley admits, the problem is in determining just how desires are united.

The whole issue of the conditions under which desires existing at different times are to be regarded as unified in a morally significant way, and hence as constituting an

⁵⁹Ibid., 132.

⁶⁰Ibid., 133.

enduring interest, seems to me deeply puzzling.⁶¹

The problem here is twofold. First, Tooley wants to defend the second premise of his argument.

The continued existence of a given subject of consciousness cannot be in that individual's interest at time *t* unless *either* that individual has a desire, at time *t*, to continue to exist as a subject of consciousness, *or* that individual can have desires at other times.

However, such a claim seems too strong, for as Tooley suggests above, we can imagine an entity who, though lacking any desires for the future and any memory of the past, nevertheless exhibits sufficient psychological continuity between the earlier and later times such that it would be appropriate to regard the entity as the subject of non-momentary interests. In other words, the entity may have interests though lacking desires, for as was pointed out before, having an interest in *x* does not entail a desire for *x*.

Second, different mental states at different moments *may* be unified without the need for a memory connection. This unification is made up of a series of overlapping beliefs about future states, and a consistent set of beliefs and attitudes. For example, imagine the case of Mrs. W. Let us assume for the sake of argument that for some reason Mrs. W has no memory of the past. Despite this she has beliefs about future states, she believes that she will continue to exist and will "display traits that flow from her present ones" -- her views, attitudes are consistent. Tooley is inclined to regard Mrs. W as a subject of non-momentary interests, since there is both causal dependance and psychological continuity between the different mental states.⁶²

⁶¹Ibid., 133.

⁶²It might be objected that any notion of psychological continuity must involve memory, that continuity of memory is a necessary condition of personal identity. This point is discussed in the next chapter.

This may be a case where an individual could be the subject of non-momentary interests yet lacking memory and desires, and therefore this does present a challenge to Tooley's account. The question remains whether in this case the subject of non-momentary interests also lacks a concept of a continuing self. A satisfactory answer to this question necessitates an investigation into the role that memory plays in the concept of a continuing self. We look at this question in detail in the following chapter, and so we cannot fully evaluate Tooley's position here until then. But we are in a position to say that we can certainly imagine cases where the memory of the individual is sufficiently disrupted that it is unclear if not questionable that the individual possesses the concept of a continuing self. For example, let us say that Mrs. W suffers from Alzheimer's dementia to the degree that she lacks the conception of herself as a "continuing subject of experiences." She is unable to recall past desires or beliefs, nor can she recall past events. However she does have a desire to continue to exist -- she wants to continue doing the things that she likes doing (for example, she likes sitting outside in the sun) and has a consistent set of attitudes and values. I do not know if such a description is clinically possible, but if it is, it would seem plausible to describe Mrs. W as a "subject of non-momentary interests." In any case, we have imagined an example that suggests that a robust "sense of self" might not be necessary for different mental states to be unified.

The conclusion to be drawn here is that a number of aspects of Tooley's account remain unclear. First, the issue of whether interests presuppose desires remains unresolved, although a number of arguments have been presented to suggest that desires are sufficient but not necessary for interests. Second, it remains unclear whether being a subject of non-momentary interests requires a concept of self. I have tried to show that we can imagine cases where we can have the former without the latter.

We now in a position to return to the conclusion of Tooley's main argument. This is

An individual cannot have a right to continued existence unless there is at least one time at which it possesses the concept of a continuing self or mental substance.

There are a number of points to be made here. First, our conclusion regarding (C) suggests that this conclusion is open to serious question. For if in order to be a subject of non-momentary interests the individual requires the concept of a continuing self, and being the subject of non-momentary interests is sufficient but not necessary for personhood, then the concept of a continuing self is not necessary for personhood. Which is to say that one might possess the right to continued existence without the possession of the concept of a continuing self. In other words, to prove his case, Tooley needs to show (i) that the concept of a continuing self is necessary for being a subject of non-momentary interests, and (ii) being a subject of non-momentary interests is necessary for personhood. However, although one can plausibly claim that he has shown the first of these claims, Tooley has shown that being the subject of non-momentary interests is sufficient for personhood.

Second, recall that Tooley put forward a possible definition of person as a subject of non-momentary interests. It would be reasonable to think that Tooley would want to maintain that being the subject of non-momentary interests is both a necessary and sufficient condition of being a person. For if person is defined in terms of the right to continued existence, and an individual can only possess this right if he or she is the subject of non-momentary interests, possession of such interests would seem necessary and sufficient for personhood. However, Tooley does not think this. His conclusion is that having a non-momentary interest is sufficient to make something a person, and it is unclear as to whether it is also necessary.

It is important to ask why Tooley is reluctant to claim that being the subject of non-

momentary interest is necessary for personhood. For this suggests the possibility that an entity might meet the conditions of personhood though not be the subject of non-momentary interests, and this would seem to undermine much of the previous extended discussion.

Third, the requirement that there is at least one time at which the individual possesses the concept of a continuing self or mental substance leads to other difficulties. If this one-time possession of the concept of self is sufficient for the individual to have the right to continued existence, then the following seems to be an objection:

There are two patients: Mrs. W (who we have met before) and B, a seven-year-old child. Mrs. W suffers from Alzheimer's dementia and no longer possesses the concept of a continuing self, though we can plausibly claim that she is the subject of non-momentary interests. B suffers from Down's Syndrome, and though he has the cognitive ability to understand the concept of a continuing self, he has yet to do so. He, too can be described as a subject of non-momentary interests. According to Tooley, whereas Mrs. W has the right to continued existence, B does not. The reason being that Mrs. W possessed on at least one occasion the concept of a continuing self. But this is to prefer past capacity (or exercise of a capacity) to future capacity, which would appear morally arbitrary.⁶³

In other words, it is unclear why it is necessary that the individual have possessed the concept of a continuing self on one occasion (and it need be no more than *one* occasion) in order to have a right to continued existence.⁶⁴

To conclude our discussion of Tooley, I want to return to the three remaining putative definitions of person. To recall, these are

- (2) A person is an entity that possesses rationality;
- (3) A person is an entity that is capable of action;
- (4) A person is an entity that possesses self-consciousness.⁶⁵

⁶³This objection is, of course, the same one that was raised against Noonan in the previous section.

⁶⁴Tooley argues that exercise of capacity and not mere capacity is necessary. See *Abortion and Infanticide*, 146-157.

⁶⁵*Ibid.*, 123.

According to Tooley, rationality is neither necessary nor sufficient to make something a person. As with the first attempted definition of personhood, Tooley's discussion of rationality is predicated on the understanding that only those entities that possess consciousness can be rational.

I shall assume that the term [rational] is being used in such a way that something cannot be said to be rational unless it possesses consciousness -- and indeed, consciousness of a certain sort: to be rational, a being must be capable of enjoying thought-episodes -- where a thought-episode is, roughly, a structured state of consciousness involving a representation of some proposition.⁶⁶

Can we say, therefore, that to be rational is to be capable of enjoying thought-episodes? Tooley thinks not on the grounds that an entity that had no thoughts "other than ones directly caused by its immediate perceptual environment would not count as rational." An entity that is simply aware of its sensations is not rational.

In order to gain a fuller notion of rationality, we need to add something to the notion of the capacity for thought-episodes. Tooley presents two possible additional criteria: First, rationality is connected to various reasoning capacities; Second, rationality involves "the capacity to have thoughts about the past and future, together with the capacity to have thoughts which involve generalizations"; Third, "To be rational is to be capable of acting on the basis of reasons, arrived at by a process of deliberation in which competing considerations are weighed against one another."⁶⁷ According to Tooley, to be rational in this third sense is equivalent to being an agent. Since the third possible definition of person is in terms of agency, we will postpone discussion of this third sense until then.

Is the capacity for thought-episodes along with reasoning capacity necessary and sufficient for personhood? Tooley thinks not. He asks us to imagine that the mental life of

⁶⁶Ibid., 135.

⁶⁷Ibid., 135.

cows is as follows.

They do have thoughts, some of which express memory-beliefs to the effect that the individual having the present thought is identical with the individual that was eating the grass in the pasture a bit earlier, and others which express thoughts about the future -- such as a belief that this individual will be out in the pasture again tomorrow. Suppose further that cows have desires not only about their present states, but about some future states...If the mental life of a cow were really like this, I suggest that it would be difficult not to regard its destruction as intrinsically wrong.⁶⁸

Tooley's conclusion is that he does not see why the cows might not enjoy a mental life of this sort, "while lacking completely any capacity for deductive reasoning, let alone for inductive reasoning or scientific theorizing." If this is correct, then the capacity for reasoning is not necessary for personhood.

Nor, it is claimed, is it sufficient. For we can imagine "a great conscious theorem-prover" who has extraordinary deductive powers and takes pleasure in elegant proofs, but nevertheless, lacks the desires and emotions that are necessary for personhood. This is to say, even if we conjoin reasoning capacity with consciousness, this is still not sufficient for personhood if desires and interests are not present.

And the matter is not solved if *rudimentary desires* are added. Tooley asks, "What is one to say about a conscious theorem-prover who can experience pleasure and pain?" We are asked to compare the destruction of the theorem-prover with the destruction of our ruminating cows.

How does the action of destroying such a cow, and replacing it with another, compare with the action of destroying such a theorem-prover mental substance, and replacing it by another? In the case of the cow, such destruction may be contrary to some desire it has, and it will certainly mean the end of a particular series of states of consciousness that are linked, not merely causally, but psychologically as well. The situation is very different in the case of the conscious theorem-prover that has rudimentary desires, but no conscious memories of its past states of consciousness, and neither any beliefs nor any desires about its future states of consciousness. For the destruction of such an entity -- provided it were done painlessly, would not violate any

⁶⁸Ibid., 136.

desire that it had.⁶⁹

The conclusion here follows from our previous discussion about subjects of non-momentary interests. The conscious theorem-prover is simply a "conscious reasoning machine," and without "some series of psychologically linked states of consciousness" the destruction of such an entity is not morally problematic.

The second possible definition of rationality is that it is "the capacity to have thoughts about the past and future, together with the capacity to have thoughts which involve generalizations." As Tooley argues, it is difficult to see why the capacity to have generalizations should be necessary for personhood. If we were prepared to regard the destruction of the cows as morally wrong, the fact that they are unable to form generalizations does not seem to alter this conclusion. Thus this second definition of rationality does not serve as a sufficient condition of personhood. If we recall our conscious theorem-prover, we have not made his destruction and replacement any more morally problematic by granting him the ability to make generalizations, for he still lack the relevant psychological continuity.

On the other hand, the capacity to have thoughts about the past and future does seem to be part of our notion of a subject of non-momentary interests, since the entity must possess the notion of a continuing subject, and a level of psychological continuity.⁷⁰

Tooley thus rejects rationality as either a necessary and sufficient condition of personhood. An objection to his account is that his notion of rationality is somewhat too specific; rationality is understood as being synonymous for "capacity to reason." But this is not the normal use of the term. It is more readily understood in the more general sense of

⁶⁹Ibid., 137.

⁷⁰Here Tooley is assuming that the notion of a subject of non-momentary interests is unproblematic. Recall, though, that Tooley admitted that he is unable to fully determine what constitutes a subject of non-momentary interests.

deliberation and understanding, and the acting on a set of relatively consistent set of goals and values. However, Tooley uses this notion of rationality as part of his definition of "agent." This leads us to the third of Tooley's definitions of person, that it is being an agent that makes something a person.

The definition of "agent" that Tooley discusses is this

An agent is a continuing mental substance that is capable of acting on the basis of reasons, arrived at by a process of deliberation in which competing considerations are weighed against one another.⁷¹

It is Tooley's view that being an agent is neither sufficient nor necessary for being a person. It is not sufficient on the grounds that, according to the definition above, an entity could be an agent though not being a subject of non-momentary interests. It is not necessary, since the notion of agency is not part of the notion of a subject of non-momentary interests.

My own inclination is to say that if something can recall some of its past states, can envisage a future for itself, and have desires about that, and if its personality traits do not alter in too drastic a fashion over short periods of time, then it is intrinsically wrong to destroy it, and it is irrelevant whether some of its desires have arisen through a process in which conflicting considerations are weighed against each other.⁷²

The fourth of Tooley's definitions of personhood is self-consciousness. It is his contention that self-consciousness is not a necessary condition of personhood, and it is unclear whether it is sufficient.

To understand Tooley's conclusion here we have to return to the previous discussion of a subject of non-momentary interests. In that discussion we described a case "where an individual has no memory of the past, yet believes that he is a continuing mental substance, and has various hopes about his future states." (This was the case to which Tooley thought

⁷¹Ibid., 139.

⁷²Ibid., 142.

it unclear whether such an individual could be regarded as a subject of non-momentary interests). In the normal case, where desires are unified by memory beliefs, the individual is aware that these desires belong to the same continuing self. In this case, self-consciousness is present. But in the case where memory of the past is lacking that matter is more difficult.

Such a person is aware of a continuing mental substance, but he is not aware of it *qua* continuing mental substance. He believes, and hopes, that is what he is, but he does not know that it is the case. So if "self-consciousness" is interpreted in a way that implies knowledge of a continuing self, it seems that there could be something that had non-momentary interests, but that lacked self-consciousness.⁷³

The claim here is that an individual could have a series of beliefs and desires yet not possess the concept of self-consciousness. In other words, the individual could be a subject of non-momentary interests yet lack self-consciousness. This claim is difficult, for Tooley does claim that the individual possess the concept of a continuing self.

Given the ways in which desires can apparently be unified into non-momentary interests, it will be essential for something to possess the concept of a continuing self. It will also, apparently, be necessary for the individual to exercise that concept in certain ways with regard to itself. But this exercise need not involve cognitive states: certain conative states will do as well.⁷⁴

I understand Tooley's point to be this: For the individual to be the subject of non-momentary interests it is necessary that he or she possess the concept of a continuing self. This is to say, the individual recognizes that she has present and future desires and has an interest in the future. Despite this, the individual does not *know* that she is a continuing mental substance, though she might hope or believe that she is. Furthermore, as Tooley says, this difference between knowledge and belief does not amount to a moral difference.

For consider the case of the individual just mentioned, with no memory of the past, but with beliefs and hopes about the future. Would it make a morally significant difference if that individual, rather than believing that he was a continuing self, merely

⁷³Ibid., 144.

⁷⁴Ibid., 145.

hoped that he was?⁷⁵

Thus if being the subject of non-momentary interests is sufficient for personhood, it is not necessary that the individual be self-conscious. That is to say, such an individual could possess the right to life though she is not self-conscious.⁷⁶

Lastly, we look at the question of whether self-consciousness is sufficient for personhood. The question here rests with the intelligibility of the notion of a continuing mental substance that is self-conscious and possesses the concept of self, yet lacks any desires or even the capacity for desires. Tooley's conclusion is that such an entity would possess self-consciousness but it would not be a subject of interests "either continuing or momentary." The reason for this is that, as we know, Tooley does not think one can have interests without desires.

We have now reached the end of our discussion of Tooley's concept of person and it is time to take stock of where we are. The conclusion of Tooley's main argument is that

An individual cannot have a right to continued existence unless there is at least one time at which it possesses the concept of a continuing self or mental substance.

This possession of this concept is one of the enduring non-potential properties that are necessary and sufficient for personhood. We saw, however, that there are problems in determining when an individual possesses the concept of a continuing self or mental substance. Two points were made: First, it is not clear that the individual must possess the concept of a continuing self in order to possess the right to continued existence. For

⁷⁵Ibid., 144.

⁷⁶An interesting and related question is whether these "beliefs and hopes about the future" are dependant on the possession of language. Tooley's concept of person does not mention specifically the capacity for language, and so I think it would be fair to conclude that he does not regard such a capacity as a necessary condition of personhood. Dennett, on the other hand, does make such a claim. See Rorty, *The Identities of Persons*, 186-191.

although possessing the concept of a continuing self may be necessary for being the subject of non-momentary interests, it is questionable whether being the subject of non-momentary interests is necessary for personhood. Second, Tooley in fact declares it to be "puzzling" whether being the subject of non-momentary is a necessary condition of personhood. The conclusion to be drawn here is that Tooley has proved that the concept of a continuing self or mental substance is a necessary condition of personhood, and hence of the right to continued existence.

We then looked at four possible definitions of personhood to determine whether any of these definitions included the concepts of a continuing self. To recall, these four definitions are:

- (1) A person is a subject of non-momentary interests;
- (2) A person is an entity that possesses rationality;
- (3) A person is an entity that is capable of action;
- (4) A person is an entity that possesses self-consciousness.

We discussed in considerable detail the first of these claims, and concluded that although being the subject of non-momentary interests is sufficient for personhood, if it is accompanied by psychological continuity, it is not necessary. The problem here is twofold. First, there is a problem with desires, which Tooley regards as presuppositions of interests. One can imagine cases where an entity can have beliefs about the future and exhibit psychological continuity, yet lack the relevant desires. We must therefore conclude that the entity cannot have interests. Tooley is reluctant to draw this conclusion, and so leaves the question of whether being the subject of non-momentary interests is necessary for personhood unresolved. Second, there is a problem with the possession of the concept of a continuing self. As with the case involving desires above, it seems that an entity could be the subject of non-momentary interests yet not possess the concept of a continuing self.

Our discussion of the remaining definitions of personhood resulted in the conclusion

that not one of these three conditions is necessary for personhood, since they bear no relation to the concept of a continuing self. Additionally, agency is sufficient for personhood, since something cannot be an agent unless it is the subject of non-momentary interests. Tooley further concludes that rationality is not sufficient and it is unclear whether self-consciousness is.

The conclusion of Tooley's concept of person is that the most likely candidate for personhood is being the subject of non-momentary interests. But this notion has to have the additional feature of psychological continuity added to it in order to qualify as a sufficient condition of personhood. We can claim that being a subject of non-momentary interests in this full sense is an enduring non-potential property that is a necessary and sufficient condition of personhood, and thereby grants the entity the right to continued existence.

In the two remaining chapters we will return to Tooley's definition of person, and this notion of a subject of non-momentary interests. For as his account suggests, it would seem that by including the notion of psychological continuity as a condition of personhood, we are thereby suggesting that personal identity is a condition, or weaker, a part of personhood. This suggestion is made more explicit in the following and final account of Gillett.

Gillett

We will now turn to a relatively brief discussion of the account of personhood put forward by Gillett.⁷⁷ The importance of Gillett's account here is that it makes explicit certain features of Tooley's concept of person, and also serves as an introduction to issues

⁷⁷"Why Let People Die?" *Journal of Medical Ethics* 12 (1986) and "Reasoning about Persons," in *Persons and Personality*, ed. Arthur Peacocke and Grant Gillett (Oxford: Blackwell, 1987). See also "Consciousness, the Brain and What Matters," *Bioethics*, vol. 4, no. 3, (July 1990): 181-198.

to be discussed in the following chapters.

According to Gillett, there are three "marks" or features that capture the important features of the concept of person. "They are...*personal identity, quality of life, and responsibility for self.*" Gillett says of these features

They are not formal defining attributes which pick out just and only the members of a given class and therefore they may be ascribed by extension from the paradigm cases, to certain individuals by reason of their similarity to us or their involvement with us.⁷⁸

According to Gillett, the following considerations suggest the importance of personal identity in the concept of person.

- (i) We treat a person as a being with continuity over time and, because he is now the same being as he will be at some future time, we ascribe to him an interest in his future and in what he will be like at that future time.
- (ii) We are concerned for the experiences of the individual involved...a concern which lies at the bottom of much of our empathic morality.
- (iii) A person acts and takes a role in human life and relationships in which our 'reactive attitudes' are shaped and applied.
- (iv) Each person has an identity...We consider that each person thinks of himself as being a certain distinct individual, the same person in the many experiences which comprise his life.⁷⁹

We can put the matter in much stronger terms than Gillett is prepared to, and claim that what is being suggested here is that personal identity is a necessary and perhaps sufficient condition of personhood. (This is an important claim. Recall that in Chapter One we saw that Buchanan and Brock argued that the moral authority of advance directives was not challenged by the "personal identity question." In particular, the person who executed the directive is not the same person as the person to whom it is later applied, since the degree of neurological deficit that renders the patient incompetent, and brings the directive into effect, not only destroys the conditions of personal identity but also of personhood ("..the

⁷⁸ Gillett, "Why Let People Die?" 84.

⁷⁹ *Ibid.*, 84.

living being who remains is *not a person at all* and hence, *a fortiori*, not a different person.")

Thus in Chapter One we were discussing the claim that personhood is a necessary condition of personal identity, whereas in this section the claim being discussed is that personal identity is a necessary (and perhaps sufficient) condition of personhood. The question to be asked, therefore, is whether these two terms are coextensive. We discuss this question in Chapter Four).

The view that Gillett holds is that

The mental life of a person can be understood only with the aid of a substantive concept of a person as one who has an identity and an interest in that identity.⁸⁰

(A full account of this view will have to wait until the following chapter, for Gillett's main motivation for supporting such a view is that he rejects the "Bundle Theory" view put forward by Parfit that personal identity can be described "in impersonal terms.")⁸¹ Gillett's reasons for regarding personal identity as a feature of personhood is that the concept of person contains the concepts of diachronicity and self-consciousness. First, with regard to the notion of diachronicity the Gillett's concept of person suggests an entity that has a continued existence over time, composed of psychological and physical features. As Gillett says above, we treat a person "as a being with continuity over time." Thus our interactions with persons are predicated on these considerations; that this entity has existed and will continued to exist. This is to say that persons are *enduring*.

Secondly, persons are self-conscious.

We consider that each person thinks of himself as being a certain distinct individual, the same person in the many different experiences which comprise his life... We expect a person to be aware of his own attitudes as being his own and to demonstrate and

⁸⁰Gillett, "Reasoning about Persons," 87.

⁸¹See Part Three of *Reasons and Persons*, (Oxford: Oxford University Press, 1984).

define for himself a character.⁸²

This is to say, in other words, that in our interactions with other persons it is presumed that the person is aware of his or her identity and of the features that comprise this identity. An implication of this view is that persons exhibit intentionality and are capable of second-order mental states. We understand the behavior of persons in terms of the person's beliefs, desires, intentions (etc.,) and presume that the person is aware of these beliefs and desires.

The appreciation of one's experiences requires a widespread articulation of the contents of those experiences by different conceptual capacities. These include, *inter alia*, the capacities that formulate motivations into a system of reasons.⁸³

The second feature of Gillett's concept of person is quality of life. Persons necessarily interact with the world and are aware of their experiences, and it is by such interactions that their lives have quality (whether it be good or bad). Thus if an entity has suffered some injury (or whatever) the result of which is that the capacity for experiencing the world is permanently lost, then such an entity loses claim to moral standing. Such a person

...will no longer manifest, in his body, the proper functioning of a sentient human being and that is a deficiency which undermines his ethical status as a person...⁸⁴

We can put the matter another way by saying that to be a person, the individual is capable of at least minimal interaction with his or her environment. Thus like Tooley, Gillett regards consciousness as a minimal condition of personhood. According to Gillett, "the nature of consciousness and its moral significance are related." If the entity permanently lacks the capacity for consciousness then "nothing matters" to such an entity.

In order for something to matter to an individual, he must have or be able to form an attitude or preference about that thing -- it must figure in his thought in some

⁸²Gillett, "Why Let People Die?" 84.

⁸³Gillett, "Consciousness, the Brain and What Matters," 192.

⁸⁴Gillett, "Why Let People Die?" 85.

salient or significant way.⁸⁵

For Gillett, the particular application of this point is with regard to the continuation of life-sustaining treatment

When we consider which types of organism have the ethical status that we accord to human beings it soon emerges that the presence of our capacity for consciousness holds a central role in our thinking. Where an individual has not got, or has irrevocably lost, all capacity for conscious experience due to brain damage, we cannot see any hope in keeping that patient alive.⁸⁶

The third feature of Gillett's concept of person is responsibility for self. *Contra* Tooley, Gillett thinks that agency is an important feature of the concept of person. To regard a person as an "agent" is to say that a person is acting for his or her own reasons, and "we respect these reasons when we treat persons as autonomous."

We ground this respect in the fact that he has thoughts about his own life based on who he is and what he values which are uniquely determined by his own first person perspective. Each person thinks about himself in a very special way and on this basis forms intentions and performs actions.⁸⁷

Gillett's notion of agency here is not simply that of "autonomy," if by autonomy we mean "self-governing." For Gillett's notion of agency is linked closely with the earlier remarks about intentionality and self-consciousness; to be an agent the person is aware of his or her attitudes and values and acts as a result of these. In this regard, Gillett would not regard various debilitating or disabling conditions such as motor neuron diseases as involving a loss of autonomy. On the other hand, patients suffering from advanced Alzheimer's dementia, and paradigmatically, PVS, could not be regarded as agents.

From the evidence of Gillett's account, I think that is judicious to claim the following.

⁸⁵Ibid., 196.

⁸⁶Ibid., 181.

⁸⁷Gillett, "Why Let People Die?" 84.

First, the capacity for consciousness is a necessary condition of personhood. Gillett claims that

...consciousness paradigmatically or essentially is a matter of the capacity for directed and flexible sensitivity toward the environment.⁸⁸

An individual who permanently lacks this "sensitivity toward the environment" is not an individual to whom anything "matters" and so presents no claim for moral standing.

Second, personal identity, quality of life, and responsibility for self are sufficient conditions of personhood. As we have seen, the second and third conditions are closely tied to notions of intentionality and self-consciousness. A problem with such conditions (perhaps the reason why Gillett refers to them as "features" rather than as "conditions") is that the question remains to what degree of level such features must be present in order for us to regard the individual as a person. For example, if we consider the normal course of mental deterioration associated with Alzheimer's dementia, at the early stage of the disease it would seem clear that the patient is a person, since he or she is still possesses the features of personal identity, quality of life and responsibility for self. But as the patient's condition deteriorates further the features will no longer be present to the same degree, and there will come a point when it is clear that they are not possessed at all.

Our problem here is where to draw the line that would divide persons from nonpersons. This is, of course, not a problem that is exclusive to Gillett's account, and we should recall that Gillett accepts that these features of personhood cannot be strictly described. However, if the impact of the distinction is that persons have moral standing and nonpersons do not, then we need to be reasonably clear as when the appropriate criteria are met.

Furthermore, the matter of personal identity remains. This feature is, I think, the

⁸⁸Gillett, "Consciousness, the Brain and What Matters," 183.

most interesting, and the one that is most open to objection. First, this would seem to be a case of *obscurum per obscuris*, for as we shall see in the following chapter, the concept of personal identity is by no means an easily definable concept. We have some inkling of the problems in front of us by considering cases of dementia and neurological deficit and the impairment of memory and other psychological features that normally results. A particular troubling question in such a case is "How much psychological continuity is sufficient for personal identity?"

Second, we might question the moral relevance of personal identity. If the concept of person is a moral concept (it certainly functions as one on all three accounts that we have looked at so far) and personal identity is an important feature of this concept, then personal identity has moral import. An objection to this would be to recall the case of Mrs. W that we introduced in our discussion of Tooley's account. Mrs. W has no memories of the past and it is controversial whether we can talk in terms of her "psychological continuity." If we conclude that there is insufficient continuity to maintain that she has a "personal identity" does this mean that our obligations to treat her are diminished?

Personhood and Incompetence

We must now bring our three accounts together and attempt to determine if there is any agreement amongst the three as to the necessary and sufficient conditions of personhood. A further task is to determine whether, on any of the accounts, competence is necessary for personhood.

It would be easy to dismiss Noonan's account on the grounds that whatever the properties that constitute personhood are, these are actual and not potential properties. But such a decision would be unfair since Noonan's account contains elements that are useful in

our discussion.

The importance of Noonan's account is it raises the question of whether the actual possession of the relevant properties is necessary for personhood, or whether past or future possession of these properties is also relevant. Noonan's position, to recall is twofold: the possession of the relevant genetic material, and the capacity to develop human wisdom in the normal course of events. Our discussion of Noonan led to the conclusion that only the actual possession of the relevant properties is morally relevant. Such a conclusion is important and contradicts the views of those such as Engelhardt who distinguish morally between past and future persons.⁸⁹ The particular application of this point is that if the incompetent patient is a person, he or she is one by virtue of the fact that he or she retains certain properties, rather than through the future or past possession of these properties. And if Tooley and Gillett are correct in their contention that some degree of personal identity is necessary for personhood, then, according to this analysis, we would have to be able to claim that this is the same person who executed the directive, rather than this *was* the person who issued it.

The purpose of discussing Tooley's account of personhood, and discussing it in such detail, is that Tooley presents an account that attempts to show why only persons can have moral standing. By defining persons in terms of the right to continued existence, that is to say, the fundamental requirement of moral standing, Tooley thereby clearly delineates persons from nonpersons in moral terms.

However, as we saw in our discussion of Tooley, it is not at all clear that the account is successful. For Tooley is unable to state in an unqualified way what are the necessary conditions of personhood. It is suggested that being a subject of non-momentary interests and

⁸⁹H. Tristram Engelhardt, *The Foundations of Bioethics*, (New York: Oxford University Press, 1986).

having psychological continuity is sufficient for being a person. However, since Tooley cannot determine just what counts as a subject of non-momentary interests we cannot say whether this is a necessary condition or not. Additionally, there are unresolved problems with the requirement that the individual possess "at one time" the concept of a continuing self. This could be interpreted as a somewhat liberal requirement since an individual possessing this concept once and then never possessing it again would count as a person, even though he or she could later become incompetent. To recall, we presented a case to suggest that on Tooley's account, an incompetent patient could still be described as a subject of continuing interests with the requisite psychological continuity, despite the loss of memory. Perhaps it is Tooley's intention to include such individuals under the concept of person.

According to Gillett, the "marks" of features of the concept of person are *personal identity*, *quality of life*, and *responsibility for self*. (Gillett has misgivings about being able to come up with an exhaustive set of necessary and sufficient conditions of personhood, but with that *caveat*, these three features are certainly important enough that their absence would disqualify the entity from being a person). These terms are complex ones, in the sense that we can understand them as being composed of more basic notions: Personal Identity can be understood in terms of psychological and physical continuity; quality of life can be understood in terms of the capacity for consciousness and the capacity to interact with the environment; responsibility for self in terms of autonomy and self-determination.⁹⁰

As we saw, Gillett argues on a number of occasions that we should regard those individuals who lack all of these properties as morally unimportant. Thus the PVS patient

⁹⁰The use of "complex" in this sense is from Parfit. We will discuss Parfit's distinction between "Simple" and "Complex" terms in the next chapter. See Derek Parfit, "Later Selves and Moral Principles," in *Philosophy and Personal Relations*, ed. Alan Montefiore, (Montreal, Canada: McGill-Queen's University Press, 1973), 138.

is not a person, and we should feel no obligation (other things being equal) to continue life-sustaining treatment for such a patient. This is to say that Gillett regards these features as necessary for personhood. Does he also regard them as sufficient? This is a more difficult question to answer in that the complexity of the above terms also suggests that they are matters of degree. If it is clear that the PVS patient is not a person, we then have the difficult task of deciding at what point in the gradual increase of these capacities the patient becomes a person. This is to suggest that the difference between persons and nonpersons is a matter of degree on Gillett's account.

The determination of the sufficient conditions for personhood is difficult for just about any account of personhood, and I think that it is made even more complicated by the interrelation of incompetence and advance directives. *Prima facie*, Gillett's account would not rule out an incompetent patient as a person; we cannot find in Gillett's account the claim that competency is necessary for personhood, and an incompetent person may still possess the three features put forward in some capacity. The complication with advance directives is that, as we saw in the previous chapter, this method of decision-making bears directly on the notions of autonomy and self-determination. An advance directive can be seen as a way by which the patient's self-determination can be extended to cover later states of incompetence.

Tooley and Gillett agree on two points: First, that a necessary condition of personhood is consciousness. Both accounts describe conditions of personhood that necessitate that the entity has the capacity for consciousness. Second, that some notion of psychological continuity is necessary. Whereas Tooley prefers this term, Gillett is more explicit and argues that personal identity is a condition of personhood. The matter of the necessity of some degree of personal identity remains unresolved, and will remain so until we have completed the following chapters.

There is considerable disagreement on the matters of self-consciousness and agency. The disagreement can, perhaps, be explained in part by a difference in strategy. Tooley is defining the concept of person in terms of the right to continued existence, whereas Gillett is defining the term with no such specific right in mind. This is, I think, important if we regard the right to continued existence as the minimal right granted to persons and denied nonpersons. On Gillett's account, the notion of personhood described is based, in part, on a series of values that we think are important. For example, Gillett describes a person as one "who acts and takes a role in human life and relationships in which our 'reactive attitudes' are shaped and applied." To put the matter differently, Gillett's concept of person incorporates a "social sense" of person.

The disagreement over self-consciousness might also be dissipated by recalling that Tooley attempts to include some notion of self-consciousness in his concept of a subject of non-momentary interests and the concept of a continuing self or mental substance. The difference may come down to whether we agree with Tooley that we can imagine an individual who though being a subject of non-momentary interests may still be unaware of the fact that he or she is a continuing and the same subject of consciousness. Recall that Tooley does claim that in order to have the right to continued existence the individual must have the relevant desire, and thus must possess the concept of a continuing self. This should include the individual's awareness of him or herself. As Tooley says,

To sum up, then, to have a desire for one's continued existence presupposes that one possesses the concept of a continuing subject of experiences, and that one is aware of oneself as a subject of experiences.

The moral to be drawn here is that self-consciousness does play a part in Tooley's account, the problem is that it is unclear just what part it plays. I think it would be fair to say that the class of entities described by Tooley's concept of person will be the same as the

class of self-conscious entities, excluding rare and exceptional cases.

There appears to be a clear disagreement between Tooley and Gillett with regard to the matter of agency. We could attempt to highlight the disagreement by distinguishing between "active" and "passive" notions of person. On an active sense of person, the person is "reactive" and responds to his or her environment. In the fullest sense of active, persons are moral agents; a view put forward by Kant, Dennett, Rawls, Engelhardt, and Frankfurt amongst a host of others.⁹¹ As a moral agent, it is therefore required that the individual possess the notions of responsibility for self, autonomy, and be capable of action.

On a "passive" notion of person, the individual is not meant "to do much at all." For the properties of personhood describe *static* properties of the individual. By this I mean that one does not have to *do* anything to be self-conscious, a subject of non-momentary interests. It is for this reason, I believe, that Tooley rules out rationality as a necessary condition of personhood. Other supporters of such a view, in one way or another, are Warren, Noonan, and Thompson.⁹²

The difference between these two senses of person may in part be motivated by the

⁹¹Kant, *Groundwork of the Metaphysic of Morals* in *The Moral Law*, ed. H.J. Paton, (London: Hutchinson University Library, 1948). Daniel C. Dennett, "Conditions of Personhood," in *The Identities of Persons*, ed. Amelie O. Rorty, (Los Angeles: University of California Press, 1976), 175-197. John Rawls, *A Theory of Justice*, (Cambridge: Harvard University Press, 1971). H. Tristram Engelhardt, *The Foundations of Bioethics*, (New York: Oxford University Press, 1986). Harry G. Frankfurt, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* LXVII, 1, (1971): 15-20, reprinted in *What is a Person?* ed. Michael F. Goodman, (The Humana Press, 1988), 127-145.

⁹²Mary Ann Warren, "On the Moral and Legal Status of Abortion," *Monist* vol. 57, no. 1, (1973): 43-61, reprinted with "Postscript on Infanticide" in *Bioethics*, ed. Rem B. Edwards and Glenn C. Graber, (New York: Harcourt Brace Jovanovich, 1988), John T. Noonan, "An Almost Absolute Value in History" in *The Morality of Abortion: Legal and Historical Perspectives*, (Cambridge: Harvard University Press, 1970), reprinted in Rem B. Edwards and Graber, *Bioethics*, Judith Jarvis Thompson, "A Defense of Abortion," *Philosophy and Public Affairs*, vol. 1, no. 1, (1971): 47-66.

type of entities that one wishes to enfranchise, given one's prior views as to the type of entity one thinks ought to be granted moral standing.⁹³ It may also be a difference in the understanding or how the moral and non-moral senses of personhood are connected.

The conclusion of this chapter is that consciousness is a necessary condition of personhood, this much is clear. It is probably not, however a sufficient condition since an entity may be conscious though lack any interests or desires in its continued existence. It is also clear that competence plays little part in the concepts of person that have been discussed. As is clear from Tooley's account, rationality is regarded as neither a necessary nor a sufficient condition of personhood, and except in so far as incompetence would prevent the individual from not being able to conceive of herself as a continuing subject of interests, competence is not a condition of personhood. Furthermore, this is the case even on Gillett's "social sense" of person, for the incompetent patient may still "act and take a role in human life."

The search for additional necessary, and for sufficient conditions has been less successful. The most interesting, and perhaps the most promising added condition is a notion of personal identity. To recall Gillett talks of a person in the following terms.

We treat a person as a being with continuity over time and, because he is now the same being as he will be at some future time, we ascribe to him an interest in his future and in what he will be like at that future time.

Each person has an identity...We consider that each person thinks of himself as being a certain distinct individual, the same person in the many experiences which comprise his life.⁹⁴

and

⁹³This point is made by Ruth Macklin, "Personhood in the Bioethics Literature," *Millbank Memorial Fund Quarterly/Health and Society*, vol. 61 no. 1 (Winter 1983): 35-58.

⁹⁴Gillett, "Why Let People Die?" 84.

The mental life of a person can be understood only with the aid of a substantive concept of a person as one who has an identity and an interest in that identity.⁹⁵

As was previously suggested, Gillett's remarks here seem to amount to claiming that personal identity is a necessary and sufficient condition of personhood. However, on the evidence of Gillett's own account, we should be wary of claiming that an individual who for one reason or another no longer "has an identity and an interest in that identity" is therefore not a person. The minimal condition of personhood that Gillett provides is that the individual must be capable of reacting to her environment in some fashion. It is on the basis of such a claim that he excludes the PVS patient, for such an individual has permanently lost the capacity for consciousness and so "nothing matters" to such an individual.

However, it would be a mistake to regard the individual's interest in her own identity as constitutive of personhood. In preferring to talk of "marks of the concept" rather than "conditions of personhood" Gillett is presumably deliberately attempting to avoid having to state what must be the case in order for the individual to qualify as a person (apart from the capacity for consciousness). Thus we can understand Gillett as claiming that it is the capacity for consciousness that is necessary for the individual to have moral standing, and necessary for the features of personhood as described. What remains unclear is whether *quality of life*, *responsibility for self* and *personal identity* suggest that Gillett regards these conditions to be sufficient but not necessary for personhood, or that the concept of person can be applied with different meanings in different contexts. Thus although the three features above are "marks" of the concept of person, that is to say, identify important features of the concept of person, in the context of determining the rights and responsibilities of the person we may have to use more rudimentary considerations such as the capacity for consciousness.

⁹⁵Gillett, "Reasoning about Persons," 87.

In Tooley's case, the requirement is that the individual possess the concept of a continuing subject. To talk in these terms is not to suggest that what is required is a robust notion of personal identity. For Tooley accepts (perhaps reluctantly) that the demented patient may be aware of her past and have thoughts about the future though lacking a notion of a "subject of non-momentary interests" and recognition of herself as a distinct individual. In this regard, we can say that such an individual has an interest in her continued existence though lacking the notion mentioned above.

However, the notion of a "subject of non-momentary interests" is an important one. For even if we are prepared to deny that the incompetent patient is a person, we may still maintain that such a patient should be granted some degree of moral worth on the grounds that she has past and future interests. As we saw in our discussion of Tooley, it is difficult to determine under what conditions it is appropriate to claim that the individual is a "subject of non-momentary interests", and to what extent the individual must conceive of herself in this way. Nevertheless, Tooley's account does incorporate this consideration that part of what gives you or I moral standing is that we have past and future interests, and that the concept of person is connected in some way to the notion of an entity that has a continuous existence over time. A difficult and important question is just what we mean by this notion of continuous existence over time. We now turn to this question and the matter of personal identity.

CHAPTER THREE

PERSONS AND PERSONAL IDENTITY

This chapter deals with the issue of personal identity. We will be attempting to answer the following questions.

What makes a person at two different times one and the same person? What is necessarily involved in the continued existence of each person over time?¹

An answer to these questions will provide a means to resolve the "identity question" with regard to the subjective approach. If we can get a clear understanding of what is involved in the continued existence of each person over time, then we should be able to determine whether incompetence, or what level of incompetence, destroys the conditions of identity.

The Identity Question

Before we begin our discussion of personal identity it is helpful to describe a case that is germane to our discussion. To return to our starting-point in Chapter One, we are interested in those cases that challenge the moral authority of the subjective approach to decision-making for incompetent patients.

Mr. H, a former concert flautist, is sixty-three years old and suffers from cerebral arteriosclerosis. As a result of this condition Mr. has become increasingly agitated and depressed, and already shows symptoms of increasing dementia. He has recently been admitted to University Hospital suffering from pneumonia, and has been intubated. Mr. H has a Living Will which he executed seven years ago. He requested that he not be kept alive by artificial means if he should become incompetent. However, Mr. H's family have requested that the Living Will be ignored on the grounds that since the time of its execution, Mr. H has changed his mind. As evidence in support of this contention, the family point to Mr. H's apparent enjoyment in being read to, that from time to time he seems to take pleasure in being with his family, and his now

¹Derek Parfit, *Reasons and Persons*, (Oxford: Oxford University Press, 1984), 202.

obvious dislike of music.

What do cases like that of Mr. H have to do the issue of personal identity, and how do they challenge the moral authority of advance directives and substituted judgment?² One critic, Rebecca Dresser, challenges the "moral authority" of advance directives on the grounds that the incompetent patient is not the same person as the one who formerly expressed the treatment preferences.³

This criticism, which I am referring to as the "identity question" has two aspects. First, the person's psychological make-up may have changed during the interval between the time the directive was executed and the time when it would be used in decision-making. For example, if there is gap of a number of years between the time the directive was executed and when it is used in decision-making we might question whether it accurately describes the patient's treatment interests.

Second, the state of affairs (broadly speaking, some form of neurological damage or deficit) that has rendered the person incompetent might on its own cause sufficient disruption in the person's psychological make-up to challenge the view that the incompetent patient is the person who executed the directive. For example, we might think that a patient such as Mr. H, who suffers from a progressive deterioration of cognitive ability, shares so few, if any,

²The term "substituted judgment" is used to refer to those decisions that are made on behalf of the patient by another person "donning the mantle" of the patient and attempting to decide as the patient would have done.

³Rebecca Dresser has been a consistent critic of advance directives. See "Advance Directives, Self-Determination, and Personal Identity," in *Advance Directives in Medicine*, ed. Chris C. Hackler, Ray Moseley, and Dorothy E. Vawter (New York: Praeger, 1989), 155-171, and "Autonomy Revisited: The Limits of Anticipatory Choices," in *Dementia and Aging: Ethics, Values and Policy Choices*, ed. Robert H. Binstock, Stephen G. Post, and Peter J. Whitehouse (Baltimore: The Johns Hopkins University Press, 1992). Similar criticisms are made by John Robertson in "Second Thoughts of Living Wills," *Hastings Center Report*, vol. 21 no. 6 (November-December 1991): 6-10.

memories, beliefs and values with the prior competent person who issued the directive, that it would be fair to conclude that the conditions of identity are no longer met.

Rebecca Dresser presents her challenge to advance directives in the following way.

Underlying the advance directive concept is the assumption that the person executing the instruction or proxy directive is the same person whose medical care will be governed by the directive...Yet suppose an individual undergoes a significant change in circumstances during the time between executing the directive and becoming incapacitated by illness or injury. Might certain alterations create the possibility that a new individual has emerged?⁴

Dresser also presents a slightly weaker objection

Mere acknowledgment that significant changes can occur over time in an individual's psychological features undermines the position that decisions embodied in advance directives should always control an incompetent patient's care.⁵

According to the first and stronger of these two claims, it would be plausible to maintain that the person who executed the directive is a *different* person than the one to whom it is later applied. On the one hand, this would seem to be a difficult claim to defend, for the success of this claim depends on showing not only that one set of psychological features has been destroyed, but that they have been replaced by a second of features sufficiently complete to constitute a whole person. On the other hand, we could simply be claiming that since the relevant continuities do not unite the earlier and later persons the conditions for personal identity have not been met, and so we must regard the later person as a different person by default.

According to the second and weaker claim, "significant changes" could occur in the person's psychological features, (either as a result of the passage of time or for clinical reasons) to the effect that the incompetent person no longer possesses the same psychological

⁴Dresser, "Advance Directives, Self-Determination, and Personal Identity," 157-8.

⁵*Ibid.*, 158-9.

characteristics (for example, memories, beliefs and values) as the person who executed the directive. This is not to say that the later incompetent person is a *different* person, but that it would be mistaken to permit past interests to override present ones on the grounds that the past interests no longer apply.

In order to assess the force of Dresser's objections we have to be able to answer the question raised by Buchanan and Brock, "How much psychological continuity is necessary for the preservation of personal identity?" If it should turn out that very little psychological continuity is necessary, then a plausible case could be made that the incompetent and competent person are the same person. Conversely, if it should turn out that a high degree of continuity is necessary, then we would conclude that the conditions of personal identity would not have been met in cases such as that of Mr. H.

A further task is to compare the identity question in the case of neurological deficit or dementia with the more traditional challenges to personal identity. Traditionally, the so-called problem cases have been exotic -- the person's brain is transplanted or perhaps the person's psychological features are somehow "taped" and then implanted into another person's brain that had the previous "occupant's" psychological features erased.⁶ Additionally, philosophers interested in personal identity have shown considerable interest in the possibilities created by recent science-fiction. For example, Parfit discusses the identity questions raised by the possibilities of "teletransportation" as witnessed in *Star Trek*. This is all well and good, but such cases depend for their success on a not immoderate "suspension of disbelief." More importantly, the conclusions drawn from such examples is mitigated by their fictional nature.

⁶For example, see Bernard Williams' article "The Self and the Future," in *Problems of the Self*, (Cambridge: Cambridge University Press, 1973).

This is not to claim that such "thought-experiments" are without merit. By imagining various unlikely but nevertheless possible scenarios, we might be able to clarify our thoughts about what we think is important to the matter of personal identity. The problem is, however, that such scenarios depend for their success on their *similarity* with the more mundane world that we live in, and it is not always clear that there is sufficient similarity. For example, we can imagine being transported *Star Trek*-fashion to some planet far, far away. Once we "appear" on this planet, Jupiter, we go about our business with the same set of coherent memories that we have here on earth. Perhaps we will get used to such a mode of transport, and regard it as travel rather than replication, but that is not my point. Rather, the question that concerns me is whether the conclusion we draw here (transportation or replication) thereby illuminates or clarifies our understanding of "real-world" cases such as whether the incompetent Mr. H is the person as the person who executed the directive. My own intuitions are that the more exotic cases do not get us very far. In other words we may not be any the wiser whatever the conclusion we draw from these thought-experiments.

A further point that relates to this issue (and one that is central to this work) is the connection between personhood and personal identity. Let us think about the moral sense of person, and, for the sake of argument, understand "person" as applying to those individual that are the bearers of rights and responsibilities. This moral sense of person is determined by our moral "system" and values. Think, for a moment, what it would be like if we could move about the galaxies in *Star Trek*-fashion. Perhaps it would be possible to duplicate oneself, to go back in time, to have one mind and fifty bodies or fifty minds and one body. All of these cases would challenge our notion of personhood, and subsequently our notions of rights and responsibilities (for example the rule of *Habeas Corpus* could no longer be so effectively applied). This is to say, in short, that such thought-experiments involve, perhaps,

a greater shift in our conceptions of personhood and personal identity than is first apparent.

This is not to say that the more exotic cases are the only cases discussed by philosophers. A number of recent books have focused on the problem of identity associated with certain neurological disorders, and with conditions such as *Multiple Personality Disorder* (MPD).⁷ However, in a number of ways, the cases discussed in these works are dissimilar to the cases discussed here. The identity question raised by the person suffering from MPD is not like that raised by the person suffering from Alzheimer's dementia or cerebral arteriosclerosis. In these latter cases we do not have two distinct and coherent personalities; we are not dealing with the question of whether two persons can inhabit one mind.

For, in the first instance, cases of neurological dementia and deficit raise the sort of question suggested by Buchanan and Brock, "How much psychological continuity is necessary for the preservation of personal identity?". This is the appropriate way to begin but two further points need to be mentioned. First, the disruption or impairment of psychological features has a clear physical cause. For example, the symptoms associated with Alzheimer's dementia -- memory loss, problems in performing mental functions, and eventually loss of motor function -- have a physiological basis.⁸ Thus we can also be concerned with the question of "How much physical continuity is necessary for the preservation of personal identity?" for the loss or disruption of psychological features is clearly tied to changes in the

⁷For example, Oliver Sacks, *The Man Who Mistook His Wife For A Hat And Other Clinical Tales*, (Simon & Schuster, Inc., and New York: Harper Collins Publishers, Inc., 1970), Kathleen V. Wilkes, *Real People*, (Oxford: Clarendon Press, 1988), James Baillie, *Problems in Personal Identity*, (New York: Paragon House, 1993), and Jon Elster, *The Multiple Self*, ed. (Cambridge: Cambridge University Press, 1986).

⁸See Jacob H. Fox, "Alzheimer's Disease and Other Dementias," in *Neurology for the Non-Neurologist*, 2d ed., ed. William J. Weiner, and Christopher G. Goetz, (Philadelphia: J.B. Lippincourt Company, 1989), 198-206.

physical features.⁹

Personal Identity

We turn now to the examination of a number of rival views of personal identity. Our main purpose here is to discover whether there are necessary and sufficient conditions for the continued existence of the person over time. To this end, it is useful to distinguish between the question of what *constitutes* personal identity from the question of what *evidence* there is for personal identity. As Richard Swinburne states the matter, "there are two philosophical questions about personal identity."

The first is: what are the logically necessary and sufficient conditions for a person P2 at a time t_2 being the same person as a person P1 at an earlier time t_1 , or, loosely, what does it mean to say that P2 is the same person as P1?

The second is: what evidence of observation and experience can we have that a person P2 at t_2 is the same person as a person P1 at t_1 ¹⁰

We are concerned here with answering the first of these questions, for an investigation into the necessary and sufficient conditions of personal identity is an investigation into the question of what constitutes personal identity.

In the following discussion of rival views of personal identity, the initial division is between *Reductionist* and *Non-reductionist* views of personal identity. An account is termed "reductionist" if it is claimed that questions about the identity of the person "reduce" to questions about the physical and psychological continuities. In this way, questions about the

⁹So far as I know, we would not gain much by asking this question in cases of MPD, for the causes of MPD are generally thought to be social rather than physiological -- the creation of different personalities is understood as a way by which the patient can accommodate past traumatic episodes.

¹⁰Sydney Shoemaker and Richard Swinburne, *Personal Identity*, (Oxford: Blackwell, 1984), 3.

identity of persons parallel questions about the identity of any other entity (and this is regarded as one of the strengths of this account by its supporters) the identity of cats, mats and people is composed of physical and/or psychological continuities. Alternatively, an account is termed "non-reductionist" it is claimed that personal identity does not consist in these "further facts" -- we cannot reduce the question of personal identity to a question about these more particular facts.

The reductionist concludes that the existence of the person *just is* the existence of brain and body, and of a series of interrelated mental and physical events. And although we say that persons "exist," persons are not "separately existing entities" over and above the brain and body, and physical and mental events. To use Wilkes term, the term "person" is "parasitic" on these other continuities.

A further claim made by some reductionists is that since personal identity consists in psychological and physical continuity, and that these continuities can be held in degrees, one can talk about there being "strong" or "weak" continuity between *A* at *t1* and *B* at *t2*.

Non-Reductionism

Swinburne and the Simple View

Our discussions of non-reductionism will focus on the account presented by Richard Swinburne. Swinburne's rejection of reductionism, and his support for the Simple View is based in large part on his contention that reductionism fails to distinguish these two questions.

The first is: what are the logically necessary and sufficient conditions for a person *P2* at a time *t2* being the same person as a person *P1* at an earlier time *t1*, or, loosely, what does it mean to say that *P2* is the same person as *P1*?

The second is: what evidence of observation and experience can we have that a

person P2 at t_2 is the same person as a person P1 at t_1 ¹¹

Swinburne terms reductionist theories *Empiricist* theories, since they focus exclusively on "fallible" empirical evidence -- physical and psychological continuity, and do not provide a plausible account of what constitutes personal identity.

Swinburne supports the Simple View

As I shall understand it, the simple view is simply the view that the truth about identity is not analyzable in terms of the fallible empirical evidence for it of brain and memory continuity....¹²

The position that Swinburne supports is that of a substance dualist; that a person is composed of a material and an immaterial part. He further contends that there is no logical impossibility in the person continuing to exist in a disembodied state, and thus sameness of body, or even any body at all, is not essential for personal identity.

We may say that there is a stuff of another kind, immaterial stuff, and that persons are made of both normal bodily matter and of this immaterial stuff but that it is the continuity of the latter which provides that continuity of stuff which is necessary for the identity of the person over time

This is in essence the way of expressing the simple theory which adopted by those who say that a person living on Earth consists of two parts -- a material part, the body; and an immaterial part, the soul. The soul is the essential part of a person, and it is its continuing which constitutes the continuing of the person.¹³

In discussing Swinburne's account we will look first at the arguments that he presents against reductionism, and his conclusion that psychological and physical continuity are fallible evidence of personal identity. This is to say that we will be discussing the claim that

1. P2 at time t_2 could be the same person as P1 at an earlier time t_1 , even though P2 has none of the brain matter (or other bodily matter) of P1 and had no apparent

¹¹Shoemaker and Swinburne, *Personal Identity*, 3.

¹²Shoemaker and Swinburne, *Personal Identity*, 21.

¹³*Ibid.*, 27.

memory of P1's actions and experiences.¹⁴

We will then examine the arguments that he presents in favor of the Simple View.

In particular, we will focus on the following claims:

2. The soul is the essential part of a person, and it is its continuing which constitutes the continuing of the person;
3. The continuing existence of a person over time is often something experienced (non-inferentially) by the person, and experienced as belonging to a single common subject.

We turn now to examine the first of Swinburne's claims. The physical continuity theory that Swinburne criticizes is a brain continuity theory: *A* is the same person as *B* if *A* has the same brain as *B*. Swinburne's rejection of this view is based on the following "fission" argument. The human brain is composed of two very similar hemispheres, the left and right hemispheres that are linked by the *corpus callosum*. Although the left and right side of the brain generally perform different functions, these functions are not exclusive to a particular hemisphere, and each function is capable of being performed by either hemisphere. Furthermore, in those cases where the individual suffers extensive damage to one hemisphere, the functions of that hemisphere could be accommodated to a considerable degree by the other.

Neurologists discovered that severing the *corpus callosum* (*commissurotomy*) proved (moderately) successful as a means of treating certain neurological conditions such as epilepsy. This led to the theoretically plausible "two-brain" view; that the human brain is really two brains in one, and if we were to divide the brain and transplant the hemispheres into two different bodies then we would have replicated the original person. As Swinburne says,

It might be possible one day to remove a whole hemisphere, without killing the persons. There are no logical difficulties in supposing that we could transplant one

¹⁴Ibid., 22.

of P1's hemispheres into a skull from which a brain had been removed, and the other hemisphere into another such skull, and that both of the transplants should take...Each of the resulting candidates would then be good candidates for being P1.¹⁵

If this argument (and the transplant) is successful the brain theorist would have to conclude that both later persons are identical with the original, which is impossible, given the logic of identity. Nor can the brain theorist claim that *neither* of the later persons are identical with the former person, for, as Swinburne points out, if only one (and either one) of the transplants had been performed, the brain theorist would have concluded that this later person was identical with the earlier person. We can, therefore, challenge the brain theorist as to why the presence of a "rival candidate" should be relevant, why a double success should be a failure.

There are a number of ways to respond to this argument, but we will focus on two. First, we might abandon the identity relation and talk in terms of the weaker relation of "survival". The importance of this strategy is that one can have many survivors, but one can be identical with only one later person. In this case, the brain continuity theorist could consistently maintain that both new people were survivors of the original person. A recent version of such a position has been put forward by Peter Unger.¹⁶ Unger declares that he is not "overly impressed with the term 'personal identity'" focusing instead on the conditions of survival. The main reason for preferring "survival" to "personal identity" is the very contention that survival is a weaker relation than identity. Whereas it is not intelligible to say that *x* is identical to *y* and *z*, we can claim that *x* can survive as both *y* and *z*. For unlike identity, survival is not an *equivalence* relation, that is to say, it does not have the properties

¹⁵Ibid., 15.

¹⁶Peter Unger, *Identity, Consciousness and Value*, (New York: Oxford University Press, 1990).

of *reflexivity*, *symmetry*, and *transitivity*.¹⁷ Thus in the case of fission above, Unger would endorse the conclusion that the two persons composed of a new body and one of the hemispheres of the original person are both survivors of the original person.

Thus one can respond to Swinburne's argument by accepting that each of the two transplanted hemispheres could be continuous with the original person. And as long as we are reasonably happy with this notion of survivor, then such accounts should not trouble us unduly.¹⁸

Second, we could simply deny that the "two-brain" view is possible, or that if possible, is sufficiently disanalogous to the real world as to be of negligible interest. This is most easily done by rejecting the idea of such brain transplants altogether. To use Parfit's term, such operations might be "deeply" impossible. The brain is rather more integrated than the picture of two equal and free-standing cerebral hemispheres would suggest.¹⁹ We cannot simply divide the brain bilaterally for certain functions of the brain are either not located "in one place" but rather are products of the interactions of other parts (for example, language and concept use, and perhaps consciousness itself) or simply cannot be divided without destroying that function, for example, the brain stem.²⁰ Furthermore, in cases of *commissurotomy* there is evidence that the subject is aware of the different "activities" of the two hemispheres. As Gillett says,

¹⁷This point is discussed at length in Chapter Four.

¹⁸A fuller discussion of the difference between the identity and survival relations occurs in the following chapter.

¹⁹See Grant R. Gillett, "Reasoning about Persons" in *Persons and Personality*, ed. Arthur Peacocke and Grant R. Gillett, (Oxford: Blackwell, 1987).

²⁰See, for example, Daniel C. Dennett, *Consciousness Explained*, (Boston: Little, Brown and Company, 1991).

In these experiments the subjects are not only conscious, but are often conscious of an internal conflict which they try to overcome by various tricks and, by succeeding, make some ascriptions secure. Because they do try to reintegrate the information they receive, they can properly be said to be struggling with certain cognitive difficulties from a unitary standpoint, rather than to have become two mutually independent streams of consciousness only contingently related together. In other words, the subject considers himself to be a single, conscious albeit handicapped subject.²¹

Moreover, even if cases of fission are problematic for the brain continuity theorist, these cases are sufficiently disanalogous to standard cases as to question the usefulness of any conclusions drawn. That is to say that we would be agreeing with a growing trend that regards such thought-experiments as unhelpful.²²

The psychological theory that Swinburne criticizes is the "memory and character theory."

This claims that, given the importance for our attitude towards persons of their memory claims and character, continuity in respect of these would constitute personal identity....²³

Swinburne presents a number of criticisms against this psychological view of personal identity, but we can identify the following as the two main objections. First, Swinburne sees no logical impossibility in a person surviving without any memories. Second, Swinburne presents an argument against the psychological theorist arising from the possibility of duplication.

In support of the first claim Swinburne cites the example of amnesia -- hence it is a logical possibility that a person could continue to exist without any memories. But as Swinburne himself admits, we normally only suppose a person to forget much of what she has done when there is sufficient bodily and brain continuity, for it is only when these other

²¹Grant R. Gillett, "Reasoning about Persons," 78.

²²For example, see James Baillie, *Problems in Personal Identity*, and Kathleen V. Wilkes, *Real People*.

²³Shoemaker and Swinburne, 8.

continuities are present that we can say that this is the same person despite the (perhaps wholesale) loss of memories, thus he cannot describe a case where "P2 was identical with P1, even though there was neither brain nor memory continuity between them." This does not trouble Swinburne at all, for part of his thesis is the rejection of "verificationist dogma" that claims that "the only things which are true are those of whose truth we can have evidence...."²⁴

Nevertheless, against Swinburne, there is something inadequate about a position whose only support is logical possibility and unverifiability. This is particularly so in the case of the claim that it is coherent to talk of a person surviving despite a wholesale loss of memory and "character continuity" ("a minor kind of evidence, hardly of great importance on its own..."). For such a view is at odds with our general conceptions of personal identity, and thus needs to be somewhat more robust to dissuade us of those conceptions.

The second argument that Swinburne presents is a "Duplication Argument" (also referred to as a "fission" argument). In presenting the duplication argument, Swinburne discusses Bernard Williams' case of a man named Charles who turns up in the twentieth century claiming to be Guy Fawkes.²⁵ According to Williams, the memory claims of this man can be checked out and fit the pattern of Fawkes' life, and other facts that cannot be checked present plausible explanations of unexplained facts.²⁶ Then another man appears, named Robert, who presents the same claims to be Guy Fawkes. Thus the same conclusion is drawn here as above with the physical theorist: Our original person, Guy Fawkes, cannot be

²⁴Ibid., 24.

²⁵Bernard Williams, "Personal Identity and Individuation," in *Problems of the Self*, (Cambridge: Cambridge University Press, 1973), 1-19.

²⁶Williams, "Personal Identity and Individuation," 7-8.

identical with both Charles and Robert, although if only one candidate had appeared, the psychological theorist would have regarded the earlier and later persons as identical.

However, the psychological continuity theorist can respond in the same way that the physical continuity theorist did, by abandoning identity in favor of survival, or by ruling out of court the possibility of such cases.

The conclusion that the non-reductionist would like to be able to draw from the problem of fission is that the features that reductionists claim as constitutive of identity are merely evidence for identity. In cases of fission, the non-reductionist can maintain that no matter how strong or weak the psychological or physical continuities are between the earlier and later person, these continuities are merely questionable evidence for the identity of the person. The non-reductionist can consistently maintain that in the case of fission only one of the later persons is identical with the original person. Thus Swinburne can respond to the problem of fission by saying that the problem shows why it is clear that what counts as evidence for personal identity is not what constitutes that identity. Although the evidence suggests that *B* and *C* are both *A*, we should disregard such evidence. If the evidence leads us to a logically impossible conclusion then this evidence must be suspect. Swinburne concludes that personal identity is a further fact over and above physical and psychological continuity.

However, it is not clear that the non-reductionist has shown that the reductionist position should be rejected. If the reductionist is prepared to adopt talk of survival rather than of identity, then she can avoid the logical problems presented by cases of fission.

We now turn to an examination of Swinburne's positive account, the view that personal identity is constituted by sameness of soul. The claim to be examined here is the second of the claims presented above.

2. The soul is the essential part of a person, and it is its continuing which constitutes the continuing of the person.

Swinburne's argument that the soul is the essential part of the person is based on Aristotelian notions of the identity of substances:

...that a substance at one time is the same substance as a substance at an earlier time if and only if the later substance has the same form as, and continuity of matter with the earlier substance.²⁷

However, as we have already seen, Swinburne argues that with regard to the identity of persons it is possible for two persons to be the same, even if there is no continuity between the bodily matter. This would mean that either the identity of persons is unlike the identity of any other entity, or that, in the case of persons, the appropriate continuity is continuity of immaterial substance. Swinburne adopts the latter option.

The position that Swinburne adopts is a dualist one -- that the identity of persons is composed of material (the body) and immaterial substance (the soul). However, the essential part of the continued existence of the person is the soul.

While on Earth, the soul is linked to a body (by the body being the vehicle of the person's knowledge of and action upon the physical world). But, it is logically possible, the soul can be separated from the body and exist in a disembodied state...or linked to a new body.²⁸

How can the soul become disembodied? Swinburne describes the situation as follows.

A person has a body if there is one particular chunk of matter through which he has to operate on and learn about the world. But suppose that he finds himself able to operate on and learn about the world within some small finite region, without having to use one particular chunk of matter for this purpose. He might find himself with knowledge of the position of objects in a room (perhaps by having visual sensations, perhaps not), and able to move such objects just like that, in the ways in which we know about the positions of our limbs and can move them. But the room would not be, as it were, the person's body; for we may suppose that simply by choosing to do so he can gradually shift the focus of his knowledge and control, e.g., to the next

²⁷Shoemaker and Swinburne, 26.

²⁸Ibid., 27.

room. The person would in no way be limited to operating and learning through one particular chunk of matter. Hence we may term him disembodied.²⁹

From this Swinburne concludes that it is coherent to talk of a person becoming disembodied.

The arguments against dualism are well-known and do not need rehearsing here, though perhaps the most telling objection is the inadequacy of explanation of how the body and the soul are interconnected. Here, we are concerned with the plausibility of the above description as an example of disembodied existence. The basis for Swinburne's support for the above supposed description of disembodied existence is that he regards such a situation as coherent, that is to say he can imagine it being so, and there is no logical impossibility in it being so. However, the relevant question is whether such an account is intelligible -- whether we can understand what it means to say that a person could have a disembodied existence.

We may be skeptical of such an account on a number of grounds. First, the picture described bears little or no relation to our normal understanding of what it means for a person to exist. Swinburne is asking us to imagine that we could interact with the world even though we have no physical contact with the world. But the story of our existence is the story of a physical entity, moving around a physical world; we do not experience the world in a disembodied way. Second, one should notice that to imagine such a picture, we have to rely heavily on physical and spatial metaphor; we are asked to imagine that we can move objects as we move our own body. This would seem to compromise the plausibility of this disembodied existence. Third, even if we grant Swinburne that one can *imagine* existing in a disembodied state, we cannot therefore conclude that the continued existence of a person *consists* in the continuation of this state. For one could equally imagine that the continued

²⁹Ibid., 23-24.

existence of a person consists in physical continuity. In other words, from the fact that we can imagine the world in one way does mean that is how the world is constituted. It is fair to conclude, I think, that Swinburne's account of continued existence of the person in a disembodied state is both implausible and inadequate as proof of the claim that personal identity consists in continuity of soul.

Finally, we turn to the third of the three claims we presented at the beginning of our discussion of Swinburne's account. This is the claim that

3. The continuing existence of a person over time is often something experienced (non-inferentially) by the person, and experienced as belonging to a single common subject.

This is the *subjective view* which maintains that the continuing existence of a person is something that can be directly experienced through introspection. It is important to distinguish two related but distinct claims suggested by the subjective view, both of which are probably held by the non-reductionist. First, there is the claim that the person is directly aware of herself, or the self, through introspection. Simply put, this is the Cartesian doctrine that the self is directly revealed in the act of introspection. Second, there is the claim that one is directly aware that one is the same continuing self - that I am directly aware that I am the same person at t_2 as the person at t_1 . It is the second of these claims on which the non-reductionist view of personal identity rests. The first claim is not sufficient to establish the non-reductionist view, for in order for personal identity to be established I must be able to discover through introspection that I am the same person at t_2 as existed at t_1 , and not simply that I am a separately existing entity, a Cartesian ego, or that I have a self.

Swinburne presents his position in this way.

I now however wish to make the point that the continuing existence of a person over a very short period of time is something which can often be experienced by that person. It is something of which we are often aware without our knowledge of it depending on our knowledge of anything more ultimate. In that sense the continuing

of a person is a datum of experience; and if it were not, we could have little knowledge of the world.³⁰

How is the continuing of a person a datum of experience? The argument that Swinburne presents to answer this question is based on the contention that

...our knowledge of change arising from sense-experience must include experience of change which is not grounded in a more basic experience of something else.³¹

(Swinburne presents a second, shorter argument to draw the conclusion that "certain simultaneous experiences are the experiences of a common subject." The argument here being the straightforward assertion that the person is aware non-inferentially that she is having a number of simultaneous experiences).

The following is, I think, an accurate formulation of Swinburne's main argument.³²

1. Human experience is the experience of change, and as a result of experience we come to know not just that certain events occurred but the order in which they occurred.
2. On some occasions the order of our experiences is known through inference based on the normal temporal order of events.
3. But this cannot be the case on every occasion. For first, our experience of the world could be ordered in a number of ways and so one unique order cannot be inferred, and second, our knowledge of the normal temporal order of events must itself derive from experience if it is to be justified and so amount to knowledge.
4. Thus we experience change directly (non-inferentially).
5. In order to experience change directly we are aware not just of a number of experiences but that these experiences occur in a particular order.
6. But we can only experience this order non-inferentially if these experiences overlap - if we experience these experiences "as overlapping in a stream of awareness."

³⁰Ibid., 42.

³¹Ibid., 43.

³²The argument can be found in Shoemaker and Swinburne, 42-44.

7. To say that these experiences overlap is to say that the person experiences them as being the experiences of a common subject.

8. Thus the continuing existence of the person is something which can be directly experienced.³³

Swinburne provides an example to illuminate this argument. Let us say that we are watching the progression of a train from p to q, from q to r, from r to s, and so on.³⁴ In order to experience the progression of the train, it is not sufficient to describe it as a number of successive experiences, for then (according to Swinburne) we could only infer the connection between the experiences.³⁵ Rather, we experience these experiences as "overlapping as in a stream of awareness," which is to say that a person experiences them as having a common subject.

The operative premise of Swinburne's argument is the seventh one. For the main objection to the argument has to do with the claim that we *experience* the self rather than that we infer that there is a common subject to our experiences. First, we may agree with the claim that experience of continuity or simultaneity experience necessitates a common subject (a unity of consciousness) on the grounds that there must be something that unites these experiences together. However, can we draw the conclusion that this unity is experienced as opposed to being inferred?

³³ The argument above in defence of the claim that the continued existence of a person over time is often experienced by that person, resembles Kant's discussion of the *Transcendental Unity of Apperception*. "There can be in us no items of knowledge, no connection or unity of one item of knowledge with another, without that unity on consciousness which precedes all data of intuitions, and by relation to which representation of objects is alone possible." *Critique of Pure Reason*, trans. Norman Kemp Smith, (New York: The Humanities Press, 1933), A 106-107.

³⁴ Shoemaker and Swinburne, 43.

³⁵ For an interesting and different analysis of the points being discussed see Daniel C. Dennett, *Consciousness Explained*, (Boston: Little, Brown and Company, 1991), 115-126.

We could answer the question as Hume does, and deny that a person is introspectively aware of the self that is distinct from the accompanying perceptions.

For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, or heat or cold, light or shade, love or hatred, pain or pleasure, I never catch myself at any one time without a perception, and never can observe anything but the perception.³⁶

But Swinburne anticipates such a reply.

Hume says that he fails to find the 'self', i.e., the common subject. One wonders what he supposed that the common subject would look like, and what he considered would count as its discovery. Was he looking for a common element in all his visual fields, or a background noise which never ceased. Is that the sort of thing he failed to find?³⁷

According to Swinburne, Hume is looking in the wrong place because the self is the *subject* not the *object* of his experiences, it is the self that makes the cohesion or unity of these experiences possible.

Nevertheless, although we may be skeptical of the Hume's view that there is no self but just a "bundle of perceptions," Swinburne does not present any further argument to support the view that we experience the self directly. In fact, his supposed refutation of Hume is based on the claim that if the successions of overlapping experiences were not experienced by a common subject, then "we would have no grounded knowledge of succession." And there's the rub. For is Swinburne not conflating what is directly experienced and what gives us "grounded knowledge"? Swinburne's argument above may be taken as an argument in support of the claim that in order for the person to have grounded knowledge of continuity and simultaneity there must be a common subject of experience, but, to repeat, it does provide for the conclusion that this subject of experience is itself a datum

³⁶David Hume, *A Treatise of Human Nature*, (1739), ed. L.A. Selby-Bigge, (Oxford: Clarendon Press, 1960), 252.

³⁷Shoemaker and Swinburne, 44.

of experience. In other words, Swinburne has not shown us that this is something experienced non-inferentially.

Furthermore, even if Swinburne were correct in his account of direct awareness of the self, this is not sufficient to establish the diachronic identity of the person. Let us assume that I am presently aware that I am the person writing this tome. Am I able to conclude by the same method that I am the same person t_2 who was working on this tome at t_1 ? If I can determine this fact through introspection, it must be with the help of memory -- I directly remember working on this work at that time. But memory is a part of the psychological continuity view that is rejected by the non-reductionist, for the claim that I am the same person now as at t_1 is dependant on a degree of psychological continuity, which is consistent with a reductionist view. However, it does not seem possible that the non-reductionist can claim identity between these two times without relying on memory. That is to say, the non-reductionist cannot claim that one's present awareness of oneself does not rely on other psychological features.

The point at issue is this. Swinburne wants to claim that we are directly aware of the self, and this awareness applies both to the present moment and over time. In other words, both the synchronic and diachronic identity of the person is known directly. But what is it to say that I am directly aware that I am the same person who performed some action on a previous occasion? Presumably, in order to make such a claim I must remember performing that action. If this is correct, then have we not included memory as part of what constitutes personal identity?

Two conclusions deserve to be drawn here. First, it is difficult to see how the subjective view could work in absence of psychological features. Even if one wants to maintain support for the non-reductionist account, a person is directly aware of certain

psychological features. Second, it is difficult to see how one could establish personal identity through introspection devoid of these psychological continuities.

Furthermore, the reductionist need not attempt to deny that a person is aware of her own identity in a different way that on is aware of the identity of others. Unless one adopts the positions of either idealism or radical behaviorism, a reductionist such as Shoemaker is prepared to accept that there is a distinction between first and third person accounts of personal identity.³⁸ However, the important point here is that this awareness of one's own identity does not *establish* this identity -- what makes the person who she is not the fact that she is directly aware of who she is, but the psychological features that constitute her identity.

Second, a series of introspections does not establish personal identity. We can see this if we imagine a series of times, $t1...t2...t3...t4...t5$. At each specified time the person is directly aware of herself, though she is not so between times. From this series of direct but separate introspections we do not seem to have sufficient evidence to claim that the person at $t1$ is the same person as the person at $t5$.

So far we have looked at a number of arguments that the non-reductionist presents against the rival account. It is time now to decide whether the non-reductionist position is defensible. The main objection to this position is that it would seem to have nothing to say about what it is that constitutes the identity of the person. The consequence of this is that the non-reductionist avoids rather than resolves the identity question with regard to the subjective approach to decision-making for incompetent patients.

It would seem that the case of Mr. H presented at the beginning of this chapter does present a genuine question of identity. Other relevant cases can be found where the patient is profoundly demented or suffers from multiple personality disorder. One could claim, as the

³⁸Ibid., 77-79.

non-reductionist does, that the disruption in psychological and physical continuity has no relation on the identity of the person. But this seems to ignore what is genuine problem. Perhaps this is a problem of evidence rather than a problem of the constitution of identity, but it is a problem nonetheless.

Furthermore, the non-reductionist does not seem to have a satisfactory answer to this problem. The non-reductionist's contention that the identity of the person is constituted by the presence of the soul is inadequate, since, as Locke argues, a soul lacking any "consciousness of its past Existence" (and for that matter, present existence) would seem to have little to do with personal identity.³⁹ In fact, this is the reverse of the point that was made previously. For the non-reductionist would seem to have to accept either that memory is part of what constitutes personal identity, and thereby seriously weaken the non-reductionist position, or deny that memory is so related, and thereby have the difficult task of explaining what it is about the soul that uniquely identifies that particular person.

Finally, the claim that personal identity is "simple and unanalysable" is surely unsatisfactory as an explanation of personal identity.

It is not clear that either of these non-reductionist responses provides any sort of explanation to the identity question, and it is difficult to understand their motivation unless one wants to provide an argument for the existence of the soul. If the reductionist's adoption of survival rather than identity is plausible, then this would appear to be a more helpful analysis than that of the non-reductionist.

³⁹John Locke, *An Essay Concerning Human Understanding*, ed. Peter Nidditch, (Oxford: Clarendon Press, 1975), 338-39. Hereafter references to Locke will be given by the appropriate page number in Nidditch.

Reductionism

We turn now to an examination of the reductionist view. According to this view, personal identity can be understood in terms of psychological and/or physical continuity. To put the point in terms of the Complex view; facts about personal identity can be understood in terms of these more specific facts.

We will begin by examining the psychological view. To recall our purpose here, we are attempting to answer the questions

What makes a person at two different times one and the same person? What is necessarily involved in the continued existence of each person over time?⁴⁰

According to the psychological view the question above can be answered as follows.

A at *t*₁ is the same person as *B* at *t*₂ iff *A* is psychologically continuous with *B*.

Our main task, therefore is to determine what is meant by "psychologically continuous with," and to accomplish this, we will examine two versions of the psychological continuity view, those presented by Locke and Parfit, as well as the view presented by Shoemaker who argues for the importance of physical continuity in addition to psychological continuity.

In discussing the reductionist our central concern is to attempt to describe an account of personal identity that can avoid the criticism levelled by Swinburne, that the reductionist view is unable to distinguish between what constitutes personal identity and the evidence for this identity. This is to say, that the reductionist view is unable to come up with any (exhaustive) set of necessary and sufficient conditions of personal identity. Rather, the criteria that are put forth can be regarded only as fallible evidence of identity.

⁴⁰Derek Parfit, *Reasons and Persons*, (Oxford: Oxford University Press, 1984), 202.

Locke

We turn now to a discussion of the account of personal identity put forward by John Locke. In the *Essay*, Locke states that, in order "to find wherein personal identity consists," we must begin by considering "what person stands for." A person is

...a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing in different times and places; which it does by that consciousness, which is inseparable from thinking, as it seems to me essential to it: It being impossible for any one to perceive, without perceiving, that he does perceive.⁴¹

From this description of "person," Locke moves immediately to his definition of personal identity.

For since consciousness always accompanies thinking, and 'tis that, that makes everyone to be, what he calls *self*; and thereby distinguishes himself from all other thinking things, in this alone consists *personal identity*, i.e. the sameness of rational being: And as far as this consciousness can be extended backwards to any past action or thought, so far reaches the identity of that *person*; it is the same *self* now it was then; and 'tis by the same *self* with this present one that now reflects on it, that action was done.⁴²

It is clear, therefore, that consciousness plays a central role both in Locke's concept of person and of personal identity. As Harold Noonan says, "consciousness, in short, is the life of persons." Furthermore, the continued existence of the person over time is the continuation of the property that is necessary for the existence of the person at any one time. It is legitimate, therefore, to regard Locke's concept of personal identity as subordinate and subsequent to his concept of person. It is a necessary condition that for *B* at *t*₂ to be the same person as *A* at *t*₁, *B* must have a continuity of the features that are necessary for *A* being a person.

In claiming that sameness of person is constituted by sameness of consciousness,

⁴¹Locke, 335.

⁴²Ibid.

Locke is thereby rejecting the view that the identity of person is equivalent to either the identity of (thinking) substance or the identity of man.⁴³ "It being one thing to be the same *Substance*, another the same *Man*, and a third the same *Person*...."

First, personal identity is not constituted by sameness of substance, either material or immaterial. Locke's argument here is that the identity of substance "matters not at all." Speaking against those who claim that there is only material substance, Locke states that it is clear that personal identity cannot consist in this, since these people recognize that the identity of animals consists in "identity of life" and not identity of substance."⁴⁴ Against those who believe that personal identity consists in sameness of immaterial substance, Locke challenges them to show why "personal identity cannot be preserved in the change of immaterial substance," or why there might not be two persons occupying at different times the same immaterial substance.

Secondly, identity of person is not constituted by sameness of man, in other words, sameness of body. As Locke states the matter,

Could we suppose two distinct incommunicable consciousnesses acting the same Body, the one constantly by Day, the other by Night; and on the other side the same consciousness acting by intervals two distinct bodies: I ask in the first case, Whether the *Day* and the *Night-man* would not be two distinct Persons, as Socrates and Plato; and whether in the second case, there would not be one Person in two distinct Bodies, as much as on Man in two distinct clothings.⁴⁵

Locke concludes that personal identity consists in sameness of consciousness.

For it being the same consciousness that makes a Man be himself to himself, *personal identity* depends on that only, whether it be annexed only to one individual Substance,

⁴³Though Locke does admit to the "more probable opinion...that consciousness is annexed to, and the affection of one individual immaterial substance," Locke, 345.

⁴⁴Locke, 337.

⁴⁵Locke, 344.

or can be continued in a succession of several substances.⁴⁶

However, to understand Locke's concept of personal identity more fully, we need to be clear what he means by "consciousness"? We can identify two different senses of consciousness in the *Essay*: First, there is the notion of *reflective consciousness* which we find in Locke's description of a person as a being possessing "reason and reflexion who can consider itself as itself, the same thinking thing in different times and places." One can put this differently, following Harold Noonan, by referring to this as the "strong sense" of consciousness.⁴⁷ In this sense, consciousness is regarded as a "reflexive second-order act."

To be conscious of one's acts is to share, *qua* witness, knowledge of their occurrence with oneself, *qua* agent. And having been a witness in this way to one's own acts one can retain the knowledge of them thus gained. It is such shared knowledge had by a present self of a past self's actions which Locke thinks of as constituting personal identity.⁴⁸

Second, there is a weaker sense of consciousness, (identified by Baillie among others) that understands consciousness in terms of a more general notion of *awareness*.⁴⁹ This weaker sense can be found in Locke, it is claimed, in his mention of the consciousness that accompanies all perceptions, and describes our normal waking state. As Wilkes says, consciousness functions here in the sense of being simply "conscious of" something, i.e., "awake."⁵⁰

Which of these two senses of consciousness is most relevant to the matter of personal identity? An answer to this question can be gained by observing how Locke describes the

⁴⁶Locke, 344.

⁴⁷Harold W. Noonan, *Personal Identity*, (London: Routledge, 1991), 53.

⁴⁸Ibid., 53.

⁴⁹James Baillie, *Problems in Personal Identity*, 61.

⁵⁰Kathleen V. Wilkes, *Real People*, 174.

concept of *Self*.

Self is that conscious thinking thing, whatever Substance, made up of whether Spiritual, or Material, Simple, or Compounded, it matters not) which is sensible, or conscious of Pleasure or Pain,...and so is concern'd for it *self*, as far as that consciousness extends.⁵¹

Furthermore, the relation between person and self is described in this way.

Person, as I take it, is the name for this *self*. Where-ever a Man finds, what he calls *himself*, there I think another may say is the same *person*.⁵²

We can identify two important ways that the concept of self functions in the *Essay*.

First there is the notion of self that is consistent with, and suggests what is referred to as "privileged access" and self-knowledge -- that first-person facts about personal identity are incorrigible. As Noonan says

...Locke himself implicitly accepts the Cartesian emphasis on the first-person viewpoint as providing a privileged standpoint from which proposals about the nature of the self can be judged....⁵³

Second, the concept of self is a "forensic" term (to use Locke's term) "appropriating actions and their merit" as it delimits the notion of responsibility in a twofold way: First, since personal identity consists in sameness of consciousness, if a person is unable to recall a past action then she is not responsible for that actions. Second, the concept of self is central to our notion of "first-person self-interest." As Noonan points out,

One of the merits of his own account of personal identity, Locke thinks, is that it makes comprehensible the way in which our concern about our own pasts and futures is different from the concern we have about those of other people -- it explains why personal identity matters.⁵⁴

⁵¹Locke, 341.

⁵²Locke, 346.

⁵³Harold Noonan, *Personal Identity*, 32.

⁵⁴*Ibid.*, 32.

This is to say, what makes my pleasures and pains more important (to me) than your pleasures and pains is that they are *mine*. As Locke says,

This personality extends it self beyond present existence to what is past, only by consciousness, whereby it becomes concerned and accountable, owns and imputes to it self past actions....⁵⁵

Here we can identify the notions of responsibility and self-interest as they apply to the concept self.

Clearly, in order for Locke to be able to say that personal identity is related to notions of responsibility and self-interest, he must be using consciousness in the strong sense, as a reflexive second-order act. For if we were simply dealing with the weaker sense of consciousness, where consciousness simply means "awake" or attentive, then it seems difficult to see how this could provide sufficient grounds for self-interest.

Although Locke claims that personal identity consists in sameness of consciousness, we can better understand the substance of the account by focusing on the concept of memory. Locke, in fact, never uses the word "memory" in his statements about personal identity, but we can, nevertheless identify the concept. For as we saw above, Locke claims that

...as far as this consciousness can be extended backwards to any past action or thought, so far reaches the identity of that *person*; it is the same *self* now it was then; and 'tis by the same *self* with this present one that now reflects on it, that action was done.⁵⁶

The position that Locke seems to be endorsing here is that *A* is the same person as *B* if *A* can remember the past actions or thoughts of *B*.

We turn now to a discussion of Locke's account. As Mackie says

There is...considerable plausibility in at least the broad outline of [Locke's] positive thesis, that personal identity is somehow to be equated with, or based on, the unity

⁵⁵Locke, 346.

⁵⁶Ibid., 335.

of consciousness.⁵⁷

The plausibility here is that, *prima facie*, consciousness does seem to describe the *locus* of personal identity; I know that it is I who am perceiving the passing train because I am directly aware that I am doing so, and I know that I am the person who visited Buckingham Palace because I can directly remember doing so. Furthermore, memory does seem to play an important role in personal identity. Minimally, a person's claim that she remembers a past experience, action or event is often used as evidence that she is the person who originally participated in that action or event. This is not to claim that this evidence is not rebuttable. As we shall see, one can object to the memory criterion on the grounds that we can be mistaken in our memory claims. However, in one sense, our memories can be constitutive of our identity even if it should turn out that some of them are false. Additionally, from the first-person perspective, a person's identity is constituted by her memories; past experiences impact on present ones and determine the type of person she is.

However, despite the plausibility of Locke's account of personal identity there are a number of objections that have been raised against it. Here we will be concerned almost entirely with objections that relate to consciousness. This seems appropriate for consciousness is the basis of Locke's concept of personal identity, and it is also the most troubling component.

Our discussion will center around the problem of circularity and its relation to the unity of consciousness, and we will look at this problems with regard to both synchronic and diachronic identity. The importance of the problem of circularity is that if this objection is well-founded then Locke is mistaken in his contention that personal identity consists in sameness of consciousness. Just what it does consist in we do not know - perhaps we would

⁵⁷J.L Mackie, *Problems from Locke*, (Oxford: Clarendon Press, 1991), 176.

have to throw in our lot with the non-reductionists and conclude that personal identity is simple and unanalysable, and that the psychological aspects of a person's existence are merely *evidence* of identity rather than being what *constitutes* that identity.

The most well-known statement of the circularity objection is presented by Butler.

And one should really think it self-evident, that consciousness of personal identity presupposes, and therefore cannot constitute, personal identity, any more than knowledge, in any other case, can constitute truth, which it presupposes.⁵⁸

There is considerable debate as to what Butler is saying here, but the following would seem to be a plausible explanation.⁵⁹ It is part of our concept of consciousness that we can only be aware of *our own* experiences, but since consciousness of *x* presupposes the existence of *x*, so consciousness presupposes personal identity and cannot, therefore, be regarded as constituting personal identity.

To understand this objection, and the claim that we can only be aware of our own experiences, we will look at the notion of the unity of consciousness. For it is with the unity of consciousness that we can best understand the circularity objection.

As Shoemaker says, the term "unity of consciousness" is ambiguous;

...it can refer either to the unity (copersonality) of conscious states, or to the consciousness of this unity, i.e., the consciousness of a variety (or 'manifold') of mental states *as* unified into a single consciousness, namely one's own.⁶⁰

The first sense of "unity of consciousness" can be understood this way. When we talk about the "copersonality" of conscious states we are asking, "In virtue of what do different experiences of mental states occurring at the same time count as belonging to one and the

⁵⁸Joseph Butler, "Of Personal Identity" in *The Analogy of Religion*, (1736), reprinted in *Personal Identity*, ed. John Perry (Berkeley: University of California Press, 1975), 100.

⁵⁹See, for example the discussions in Mackie, *Problems from Locke*, 186-187 and in Harold Noonan, *Personal Identity*, 68-71.

⁶⁰Shoemaker and Swinburne, 102.

same person?"⁶¹ In other words, the problem of synchronic identity. But we can also raise this question in connection with diachronic identity. In this case we would be asking, "In virtue of what do different experiences of mental states occurring at different times count as belonging to one and the same person?" As we have seen, memory plays a central role in the copersonality of temporally distinct mental states of a person.

The second sense of "unity of consciousness" can be understood in terms of the notion of self-consciousness, that I am directly aware of these conscious mental states. As Shoemaker points out,

A plausible view, associated with Kant, is that unity of consciousness in the first sense in some way involves unity of consciousness in the second sense, i.e., involves self-consciousness.⁶²

When we speak of the "unity of consciousness" our concern is to explain how a number of different mental states are unified (whether they be perceptions occurring at the present moment or memories of past experiences) so that we are able to say that they "belong" to one person -- that these mental states are *co-personal*. We are attempting to answer the question, "In virtue of what do different experiences of mental states occurring at the same time count as belonging to one and the same person?" What we are asking here is how a number of distinct experiences are *organized* so as to be regarded as the common experiences of a single person.

Let us first look at the matter of synchronic identity -- the copersonality of present conscious experience, and how this type of identity might be vulnerable to the circularity objection.

We can see how a problem of circularity might appear when we try to answer the

⁶¹Ibid., 75.

⁶²Ibid., 102.

question as to what it means to say that these experiences are co-personal? An initial response is to claim that the copersonality of conscious experiences is established by the fact that they are the *common* experiences of a single person. But clearly this will not do, for our very concern is with just what is meant by "the common experiences of a single person," for we have yet to explain how the diverse experiences are united.

We need, therefore, to provide a more plausible explanation of the copersonality of conscious experience. Swinburne writes

At a single moment of time you feel cramp in your leg, hear the noise of my voice, and see the movement of my arms. It is among the data of your experience, something of which you are often aware without your knowledge of it depending on anything more ultimate, that these are all your experiences.⁶³

Swinburne concludes here that the self is a datum of experience and that personal identity is something that can be directly experienced.⁶⁴ The question facing Locke is whether the copersonality of these experiences can be described *without* having to draw Swinburne's non-reductionist conclusion.

There are a number of ways that we might try to explain the copersonality of such experiences. First, the co-personality of conscious mental states could be explained on the grounds that the person is directly aware that she is having a number of simultaneous experiences. What does this mean? To recall Swinburne's statement above, this means that the person is aware that she is simultaneously experiencing the cramp in her leg, the noise of my voice, and the movement of my arms. The conclusion that Swinburne himself draws is that this is possible if and only if the person is aware of the simultaneity of these different experiences, and that she is the subject of these experiences. Can Locke and other

⁶³Ibid., 45.

⁶⁴Previously we have argued that Swinburne may be mistaken in this conclusion.

reductionists accept this conclusion? It depends on what we mean by the term "person." If by this term we mean *simply* an entity that is conscious of a number of actions and experiences together, then we have not progressed onward from our original inquiry, for we still have to explain the "togetherness" of these actions and experiences. But perhaps that is our answer -- perhaps a person just is a collection of actions and experiences?

However, this response is problematic, for it fails to explain *how* these actions and experiences are connected, and fails to explain why we regard these actions and experiences of a as belonging to a unique individual. In fact, this is Hume's view that a person is 'just a train of perceptions' held together by certain relations. And I think it is clear that Locke's concept of personal identity is far more substantive than this. A person is more than simply a bundle of perceptions.⁶⁵ Rather, a person is "a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing in different times and places."

A second option that has to be rejected is to claim that personal identity is constituted not by the interrelation of conscious experiences, but by an immaterial substance, perhaps the soul, and it is this substance that unites the different conscious experiences and in which personal identity consists. But clearly this is not an option that Locke can support, given his rejection of the importance of identity of substance. And nor can this option be accepted by any respectable reductionist, for it is the nature of this option that the facts about personal identity cannot be reduced to facts about psychological features of the person.

Thus we need a better explanation of the organization of our conscious experiences in order to explain how they are the experiences of a unique individual. To do this we need to avoid the Scylla of the Bundle Theory and the Charybdis of non-

⁶⁵We discuss the Bundle Theory below.

reductionism. In other words, we must steer a course that on the one hand avoids the conclusion that personal identity is just a collection or bundle of perceptions, and therefore suffers from inadequate explanation of copersonality, and on the other hand, avoids the conclusion that the only way that conscious experiences can be organized is by something more primary than consciousness itself.

A third option, and one that the reductionist might feel ready to endorse, is to claim that the distinct conscious experiences can be united without implying the existence of a further substance. To see how this might be the case, let us return to the person's simultaneous experience of the cramp in her leg, the noise of my voice, and the movement of my arms. To describe this episode we can say: (1) that the person is aware of a pain - a tactile experience, as well as an aural and a visual experience, and (2) that the person is aware that these experiences occur at the same time. The first claim describes the person's conscious experiences that matches, more or less, the weak sense of consciousness described above, where consciousness can be understood as "awareness." But the second claim cannot be described in this fashion, for to say that the person is aware that these experiences occur at the same time implies a type of judgement, on the grounds that experiences on their own have no intrinsic temporal identification -- it is up to the person to recognize *when* the experience occurs.⁶⁶ Thus this second claim can only be described in terms of the strong sense of consciousness, where consciousness is described as a second-order act.

What is claimed here is that it is consciousness, understood in the strong sense, that

⁶⁶This needs some qualification. If we consider the sight of a child at age two months, and the experience of the same child at one year of age, we can infer that the first experience occurred before the second. So one might be tempted to claim that there is an ordering of events that dictates the order of experience. However, it should be pointed out that this is something that has to be inferred, given our knowledge of aging of children. Furthermore, such a case is disanalogous to the case where we are talking experiences are dissimilar.

unites the different conscious experiences. This is to say that the copersonality of conscious mental states is dependant on the fact that the person is self-conscious. We would need some explanation of here of what is meant by "self-conscious." Previously, Harold Noonan has described this strong sense of consciousness in this way.

To be conscious of one's acts is to share, *qua* witness, knowledge of their occurrence with oneself, *qua* agent. And having been a witness in this way to one's own acts one can retain the knowledge of them thus gained. It is such shared knowledge had by a present self of a past self's actions which Locke thinks of as constituting personal identity.⁶⁷

The question remains how it is that self-consciousness is able to unite the conscious mental states. One alternative is to claim that the person is aware of the copersonality of these experiences - that she is the common subject of these experiences. But it is difficult to explain, on Locke's account at least, how this could be the case. For what is this "self" that is aware of the copersonality if the self is merely the conscious mental states. One could claim as Noonan does above that "it is such shared knowledge had by a present self of a past self's actions" that constitutes personal identity. However, this still leaves unresolved how this knowledge is "shared," that is to say, how the different mental states are made copersonal.

An alternative conclusion is that the person infers that the experiences are occurring at the same time, rather than this being something that is directly experienced. In this way, one can perhaps avoid the non-reductionist conclusion that personal identity cannot be constituted by consciousness. But there is a further problem waiting in the wings.

The requirement that the copersonality and synchronic identity of conscious experiences depends on the person inferring their simultaneity would not be an insurmountable task if the entire spectrum of a person's mental states was always present before the mind, for then it would be appropriate to speak of a unique and discrete mental history, such that

⁶⁷Harold Noonan, *Personal Identity*, 53.

each "consciousness" refers to a specific set of mental states.

But unfortunately this is not the case. Following Carruthers, one can plausibly maintain that a person can perform a series of actions without being aware that she is performing such actions, what Carruthers terms "non-conscious experience."⁶⁸ To use Carruthers example, I might become so absorbed in the music of Schubert while washing the dishes that I am completely unaware of the dish-washing -- though I perform the task successfully. In such a case we would be reluctant to claim that I am not the same person who is both listening to Schubert and washing dishes, although I have only consciously experienced the former.

This objection does not depend on the circularity objection *per se*, but it does suggest that personal identity does not consist in consciousness alone. For we would not want to claim that I am not the same person who washes dishes and listens to Schubert. Locke, or a fellow-reductionist could claim that either the concept of "non-conscious experience" is barren, or that through memory, hypnotism, or Freudian psycho-analysis, this experience could be made conscious. Perhaps this is correct, but legitimate doubts have been raised here against the claim that all our actions or experiences are conscious.

If the Lockean is able to rule out the possibility of non-conscious experience, does this mean that the circularity objection as it relates to synchronic identity can be avoided? There are reasons to think otherwise. Recall Jane's simultaneous experience of the cramp in her leg, the noise of my voice, and the movement of my arms. Let us ignore the matter of the simultaneity of these experiences and concern ourselves with the matter of them being the experiences of a single person. In this regard, we are concerned with the question of what makes any experience the experience of a particular person.

⁶⁸Peter Carruthers, "Brute Experience," *Journal of Philosophy* 86, (1989): 258-269.

There are two responses here. First, we could claim that the experience of the cramp in her leg belongs to Jane because it is Jane who is feeling it. The problem with this response is that it accepts wholesale the conclusion of the Humean position, that personal identity is simple a collection of perceptions. For if all that is meant by saying that an experience belongs to a particular person is that it is this person (as opposed to any other) than is having the experience, then a person simply is a collection of experiences, and the identity of the person consists simply of this.

A second response is to claim that experiences "belong" to the person in that they are somehow "tagged." By this I mean that the person herself is included as part of the content of the experience. Thus when we say that Jane feels a cramp in her leg, the fact that it is *her* leg is part of the content of the experience. This is to claim that the subject of the experience is part of the content of the experience. This might appear to be a truism, since physiologically and perhaps logically, a person can only feel her own pain. But this is not what is meant here. (Nor is the point here simply the difference between our two senses of consciousness, that the person must be conscious in a reflexive way that she is in pain. For this would seem to run afoul of the type of conclusion that denies pain-sensations to those who cannot make second-order statements that they are in pain.⁶⁹ Furthermore, the weaker sense of consciousness need not be understood as a state where the person lacks any conception that she is the subject of her experiences). What is being claimed here is that conscious experience is *personal* in that it is experience about a particular person. To put the matter no more clearly; the person is both the subject and the object of the experience -- she is the one having the experience and the one who the experience is about.

⁶⁹In "Brute Experience," Peter Carruthers supports such a position.

The conclusions to be drawn here are not drawn with great confidence. However, a case has been made that we can understand the copersonality of mental states without having to accept the non-reductionist conclusion that personal identity is either a "further fact" of is simple and unanalysable. By thinking in terms of the content of these experiences and the notion of self-consciousness, we may be able to provide some way of explaining how these states are conjoined. However, questions remain, and they will do so until we look at the account of Shoemaker that is presented later on in this chapter.

Parfit and the Bundle Theory

We now turn to a relatively brief examination of the account presented by Derek Parfit. Parfit's account is similar to Hume's in its central claim that the identity of the person consists simply in a "bundle" of perceptions.

The central notions of Parfit's account are *psychological continuity and connectedness*

Psychological connectedness is the holding of particular direct connections.

Psychological continuity is the holding of overlapping chains of *strong connectedness*.⁷⁰

This terminology requires some explanation. Parfit provides the following example of what he means by "direct connections."

Let us say that, between X today and Y twenty years ago, there are *direct memory connections* if X can now remember having some of the experiences that Y had twenty years ago.⁷¹

However, even if there are no such direct memory connections between X and Y, there might

⁷⁰Parfit, *Reasons and Persons*, 206.

⁷¹Ibid., 205.

still be *continuity of memory* between X now and Y twenty years ago. In this sense, we may regard such connections as *indirect* (though Parfit does not use this term).

Furthermore, there can be degrees of connectedness.

Between X today and Y yesterday there might be several thousand direct psychological connections, or only a single connection.⁷²

As Parfit says, we cannot determine just how many direct memory connections are required for there to be enough to claim that X and Y are the same person. However, we can describe these degrees of connectedness as being either *strong* or *weak*, and claim that

...there is enough connectedness if the number of direct connections, over any day, is *at least half* the number that hold, over every day, in the lives of nearly every actual person. When there are enough direct connections, there is what I call *strong* connectedness.⁷³

The objective in discussing Parfit's concept of person is to determine whether one can find a sufficient amount of continuity in the account to accommodate a robust sense of personal identity.

Parfit supports a reductionist position, namely

...the truth of statements about personal identity just consists in the truth of certain statements about psychological and (perhaps) physical continuity. The fact of personal identity is not a further fact, apart from certain facts about these continuities.⁷⁴

However, Parfit's reductionism is stronger than that of other psychological continuity theorists in that he claims that the facts about personal identity could be described "in impersonal terms" without having to talk of the existence of persons.

⁷²Ibid., 206

⁷³Ibid.

⁷⁴Derek Parfit, "Lewis, Perry, And What Matters," in *The Identities of Persons*, ed. Amelie O. Rorty (Los Angeles: University of California Press, 1976), 100-101.

We could therefore redescribe any person's life in impersonal terms. In explaining the unity of his 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 descriptions 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.⁷⁵

The central objection to Parfit's account, and the reason why his critics find the account inadequate, is that it is questionable whether on such an account there is sufficient coherence to talk of the "unity" of a person's life, and there are doubts as to just what is the content of the "many thoughts, desires, memories, and so on."

The objection to Parfit's account that we will focus on is that he is mistaken to claim that the facts about personal identity can be described "in impersonal terms." For such an account fails to capture what is important about personal identity, namely, the continued existence of a *particular* individual over time.

Parfit claims that both personhood and personal identity can be understood as being constituted by the holding of "more particular facts." That is to say, we use the term "person" to refer to a set of psychological features that can be held to varying degrees, and, as Parfit says, a person *just is* this set of psychological features -- there is nothing else over and above them. Furthermore, personal identity can be understood in similar way -- personal identity consists in the holding of overlapping chains of "strong connectedness" and there is nothing over and above this connectedness.

Furthermore, Parfit attempts to provide a psychological account of personal identity that can avoid the circularity objection that challenged Locke. Parfit's solution to this problem was to make use of the notion of *quasi-memory*. According to this notion, I might be able to remember the experiences of other people. In this case, it would not be the case

⁷⁵Ibid., 251.

that a person's memory of a past action presupposes that they were the person who originally performed the action. We will examine a number of objections to Parfit's notion of personal identity to determine whether it can avoid the circularity objection.

We will begin by looking at the problem of the copersonality of diachronic conscious experience, and seeing how Parfit attempts to resolve this problem.

The problems with the copersonality of synchronic conscious experiences are repeated in the cases of diachronic experiences. In this latter case we are concerned with the question, "In virtue of what are different experiences at different times the experiences of a single person?" According to Locke, the claimed necessary and sufficient condition of diachronic identity is memory: *A* at *t*₁ is the same person as *B* at *t*₂ if *B* remembers the experiences of *A*.

However, the memory criterion of personal identity is problematic. An initial objection, and one that repeats the circularity objection, is that a person can remember only *her own* past experiences. There are two ways that we could state this objection. The first statement is based on the claim that the statement "*A* remembers a past action event or experience *x*" is true if and only if *A* was the person who originally participated in that action or had that experience. But this is to admit that *A* can remember *x* iff *A* is identical with the earlier person. Hence, memory presupposes personal identity and so cannot be a condition of it.

The problem here arises from the need to distinguish genuine from delusional memory claims. A reductionist would want to avoid the conclusion that from the fact that *A* seems to remember doing *x* this means that *A* did in fact do *x*. For on this basis, I would be the identical with the person who attempted to blow up the Houses of Parliament if I have an apparent memory of doing so. (For example, perhaps it is Guy Fawkes night). But the only

way to rule out such delusional or false memory claims, is to establish on the basis of some other criteria that the memory claim is false.

We can put the matter another way by claiming that the memory criterion cannot stand on its own but, as Mackie puts it, is "parasitic upon that of bodily criterion."⁷⁶ What is claimed here is that the only way to rule out apparent or delusional memories is by saying that there is genuine memory only if there is bodily continuity. The psychological criterion cannot, therefore, be a sufficient condition of personal identity.

However, Mackie (among others) thinks this to be a mistake. As he says

It is true that we commonly use bodily continuity, or the lack of it, as evidence for or against the truth of memory claims...But it is evidence only, not part of what is meant by calling something genuine memory; we have constantly to be on our guard against the use of the term "criterion" to slur over this distinction.⁷⁷

In fact, this evidence is "rebuttable": if we met someone who made sincere, and accurate memory claims of the prior experiences of someone else, then we may be tempted to conclude, as Mackie says, that the man was "really remembering, by some direct causal link" the other person's experiences.⁷⁸ In other words, we cannot rule out the possibility that one person could remember another person's experiences.

One can also respond to the objection that a person can only remember her own experiences is to introduce the concept of *quasi-memory*. This concept, originally introduced by Shoemaker but more recently advocated by Parfit, suggests that memory need not

⁷⁶For examples of this claim see Bernard Williams, *Problems of the Self*, 1-18.

⁷⁷J.L. Mackie, *Problems from Locke*, 186-187.

⁷⁸This is to say that we accept a causal theory of memory but expand the notion of cause to include "memory-transfers" and their ilk.

presuppose personal identity.⁷⁹ Parfit describes the notion of quasi-memory in this way.

- (1) I seem to remember having an experience,
- (2) *someone* did have this experience,
- (3) my apparent memory is causally dependent, in the right kind of way, on that past experience.⁸⁰

In this case, when a person claims to remember a past experience, this memory goes along with the belief that it is the same person who had the original experience. But according to Parfit, this is a separable belief.

Because we do not have quasi-memories of other people's past experiences, our apparent memories do not merely come to us in the first-person mode. They come with a belief that, unless they are delusions, they are about our own experience. But, in the case of experience-memories, this is a separable belief. If like Jane we had quasi-memories of other people's past experiences, these apparent memories would cease to be automatically combined with this belief.⁸¹

Thus according to Parfit, it is not part of the concept of memory that a person can remember only her own experiences, although this is how we do generally understand the terms. As Mackie says

I see no reason for saying that it is part of the meaning of the words "memory", "remember" and so on that a man can remember only that same man's experiences....⁸²

The second statement of the claim that a person can remember only her own past experiences is based on the contention that their content includes reference to aspects of the person that are unique to that person. This is to say that part of the content of the memory is that I remember the experiences as *mine* as the events happening to *me*, rather than as an

⁷⁹Sydney Shoemaker, "Persons and their Pasts," *American Philosophical Quarterly*, vol. 7 no. 4 (October 1968): 269-285, reprinted in *Identity, Cause, and Mind*, (Cambridge: Cambridge University Press, 1984) 19-49. In addition see Parfit, *Reasons and Persons*, 220-223.

⁸⁰Parfit, *Reasons and Persons*, 220.

⁸¹*Ibid.*, 222.

⁸²J.L. Mackie, *Problems from Locke*, 186.

"anonymous" experience.

Consider, for example, my memory of visiting Lord's Cricket Ground. According to Parfit, this memory could be quasi-remembered by another person if that person seemed to remember this experience and the memory was transferred in the "right kind of way." Furthermore, this person's quasi-memory of my experience "from the inside" can be understood as that this person experiences Lord's as I did, though without believing that he or she was the one who had the original experience.

However, what if there were aspects of this experience that in some form made it mine? For example, it was my father who accompanied me, those were my hands holding the scorecard. Part and parcel of this memory is the relation of these features to the rest of my psychological make-up. Now imagine the person quasi-remembering this experience. Will he or she know what this is a memory of, who is the person sitting on the next seat? It is arguable that this quasi-memory will make little sense, and as such, will not have the appearance of a "genuine" quasi-memory.

Thus in this case, the introduction of quasi-memory would not solve the problem of distinguishing genuine from delusional memory if we thought that it is part of the factual content of memory that it was I who had the original experience. This is to say, the introduction of quasi-memory would not meet the circularity objection according to the second statement of this objection. This is to say that each memory is identified or "tagged" as being mine.

This identification may be in one of two ways: It might be that part of the content of the memory is that I remember the experiences as *mine* as the events happening to *me*, as suggested above. Alternatively, it might be that our memories are sufficiently complex, they contain sufficient information and detail that they cannot be transferred from one person to

another. As Marya Schectman wrote in a recent article

...simply deleting the "nametag" from a memory does not make it nondelusional, and that in order to make an apparent memory truly nondelusional one will either have to presuppose the identity of the rememberer with the person who had the experience, or else remove so much of the content of the memory that it is not longer plausible to say that what is relevant to personal identity in genuine memory is preserved in quasi-memory.⁸³

To understand Schectman's argument we can use Parfit's own example of Jane and Paul in Venice. And we take as our example of a quasi-memory a relatively complex memory that Paul had in Venice, for example, having a nice meal with his wife and family opposite St Mark's.

According to Schectman there are only two possibilities as to how we can understand Jane's quasi-memory of this event. First (and we can call this "minimal" quasi-memory) Jane will have a quasi-memory of this event but will not interpret the event as Paul does. Simply, she will have the memories and that is all. She will not experience "as if she were Paul" but as a set of memories experienced "from the outside." She will have a set of memories of a Cathedral she does not recognize, and of a woman and two children she has never met before. Second (and we can call this "rich" quasi-memory) Jane experiences the memories exactly as Paul would have experienced them "with all the same personal elements and associations." So she remembers having lunch with her wife and her two children in front of a Cathedral she recognizes.

For Schectman, neither of these alternatives is useful in an account of quasi-memory. With regard to first alternative, Jane's quasi-memory of the experience will be so different from the original experience that it will not capture what is relevant to personal identity in genuine memory. Jane's quasi-memory of having lunch bears not a great resemblance to

⁸³Marya Schectman, "Personhood and Personal Identity," *Journal of Philosophy*, vol. 87 no. 2 (February 1990): 71-93.

Paul's experience of doing so. With regard to the second alternative, the quasi-memory would run up against the rest of Jane's "psyche" since it is not consistent with the rest of her psychological make-up. (She does not have a wife). In order to enable this quasi-memory to "mesh" with her own psychological make-up we would have to transfer other memories or radically alter Jane's present psychology. This is to say that in order to have a sufficiently rich quasi-memory as to capture what is important in genuine memory one must have the same psychological make-up as the person who had the original experience. As Schectman says,

If the memory must be such that I think of it as *my* family ..., then the mineness of the experience seems to be part of the content of the memory.⁸⁴

Schectman's conclusion here is at odds with Parfit's contention that the content of the quasi-memory is separable from a belief as to who had the original experience, and thus we would not have avoided the circularity objection.

A further objection to the memory criterion focuses on the contention that memory is not necessary for personal identity. This objection has to do with the fact that we do not remember most of our past experiences or actions. For example, I do not now remember visiting London to see the "Changing of the Guard" outside Buckingham Palace for the first time, yet we would be reluctant to conclude that, therefore, I am not the same person as that young child who accompanied my family at that time. This problem becomes more complicated if we include the fact that, twenty years ago, I did remember that visit. Thus on the memory criterion, I am both the same and not the same person as that young child.

There are a number of responses to this objection. As Mackie suggests, we could talk in terms of the notion that I *could remember* these past events, that they are the

....units of *potential* consciousness, that a person could in principle remember every conscious action or experience of his, even those that he cannot in fact be brought

⁸⁴Ibid., 83.

to remember by any stimulus.⁸⁵

However, it is not clear that a person could, in principle, remember every conscious action or experience.

A more successful position is to claim that all of the past (and present) events belong to the "same unit of consciousness, the same unified mental history."⁸⁶ This position has been put forward by a number of writers, including Grice,⁸⁷ Mackie and, in a different way by Parfit. To recall the central notions of Parfit's account, Parfit distinguishes between *psychological continuity and connectedness*

Psychological connectedness is the holding of particular direct connections.

Psychological continuity is the holding of overlapping chains of *strong connectedness*.⁸⁸

In this way, Parfit can claim that although I might not be able to remember past events from thirty years ago at this time, and hence there is no *direct* connection between me now and the person thirty years ago, nevertheless between these times there is an overlapping chain of connectedness between each moment and the proceeding and successive moment. In this way I am continuous with that earlier person.

But the problem here has to do with the notion of a "unified mental history." According to Parfit, with any memory, the actual experience can be separated from the belief that the person has that it was she who had the original experience. This is to say that it is impossible to experience the memory "from the inside" although making no claim as to who

⁸⁵J.L. Mackie, *Problems from Locke*, 180.

⁸⁶*Ibid.*, 180.

⁸⁷H.P. Grice, "Personal Identity," *Mind*, vol. 50 (October 1941): 330-50.

⁸⁸Parfit, *Reasons and Persons*, 206.

is the person "within" the memory. That is to say, to use Parfit's own example, a person might have a memory of visiting Venice although there is nothing about that memory that identifies the memory as the memory of a particular person. For if there was such an identification, then the memory would imply or entail the belief that the experience "belongs" to a particular person.

The same analysis applies to consciousness experience as a whole. If one regards one's experience of past actions as a subset of conscious experience in general, then *any* conscious experience can be described in "impersonal" terms. This is to say that for any conscious experience there is nothing that implies or entails that this conscious experience belongs to a particular person.

But this would seem to be inadequate as an explanation of how or why these conscious experiences are conjoined. Now Parfit accepts this conclusion -- it is his claim that

In explaining the unity of his 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.⁸⁹

But on what grounds do we talk of the "unity" of a person's life? Parfit's answer to this is that the unity of a person's life is explained by *Relation R*.

Relation R is what matters. *R* is psychological connectedness and/or psychological continuity, with the right kind of cause.

I also claim

In an account of what matters, the right kind of cause could be any cause.⁹⁰

This is to say that there is psychological connectedness and/or continuity amongst psychological states if there any type of causal relation whatsoever between these states. As

⁸⁹Ibid., 251.

⁹⁰Ibid., 262.

Parfit says, normally this causal relation is physical.

The normal causes of memory involve the continued existence of the brain. And some or all of our psychological features depend upon states or events in our brains. The continued existence of a person's brain is at least part of the normal cause of psychological continuity.⁹¹

However, Parfit accepts *any* cause, thus conscious experiences could be transferred, recorded or whatever, and as long as there was some overlapping chain of direct connection between conscious experiences, then it is legitimate to speak of the unity of a person's life.

This account of the unity of a person's life appears to be inadequate. For there appears to be insufficient explanation of the connection of the conscious experiences. If we were to describe the unity of a *Parfit-person's* life, it would be something fitting the following description: Imagine a set of conscious experiences that are sufficiently anonymous so as not to prompt the belief that they are the experiences of a particular person. Furthermore these experiences "could have come from anywhere" -- no physical causal connection is required. But what then unites the experiences? The only option would appear to be that they are the common experiences of a single individual. But again, we have failed to explain this commonality, and, to recall, according to Parfit, we can describe the unity of a person's life without having claim the existence of a person. This is to say that a person just is this collection of conscious experience, which as we have seen, it cannot be.

The problem that Parfit faces here is a problem that is general to the Bundle Theory. As Harold Noonan says, Parfit is mistaken in his understanding that psychological states that are constitutive of personal identity cannot be regarded as ontologically prior to the concept of person. As Noonan says, these psychological states

...should be thought of precisely *as* states in the sense to which Parfit objects, that is as entities which are ontologically and conceptually dependant on persons...In short,

⁹¹*Ibid.*, 208.

experiences are "adjectives of" their subjects and not independent entities in their own right.⁹²

Parfit's claims that his notion of psychological continuity permit's him to talk about the unity of a person's life although describing that life in "impersonal terms" and regarding the history of the person as simply the history of overlapping of conscious experiences. But as has been suggested above, there does not seem to be sufficient explanation as to how these experiences are unified or "overlap." Furthermore, as Harold Noonan suggests, it is persons who have psychological states -- these states do not exist around and about us, are organized, and then this organization is referred to as a person. This is not to say that we should agree wholeheartedly with Strawson in his contention that person is a *primitive concept* or *basic particular*, but it does suggest that Parfit is mistaken to claim that we can either meaningfully talk about or understand the concept of person in "impersonal terms."⁹³ That is to say Parfit is mistaken in claiming that

Though persons exist, we could give a *complete* description of reality *without* claiming that persons exist.⁹⁴

Shoemaker and Functionalism

We turn now to examine the third and final account. In examining Shoemaker's account we will concentrate on two important considerations. First, Shoemaker's description of the causal relation of psychological states that can be described in functionalist terms. Second, his contention (that we saw above) that first-personal descriptions of personal identity are not based on a notion of identification, and thus, according to Shoemaker, are not identity

⁹²Harold Noonan, *Personal Identity*, 121.

⁹³Peter F. Strawson, *Individuals*, (London: Methuen, 1959).

⁹⁴Parfit, *Reasons and Persons*, 212.

statements.

As Shoemaker states the matter, the search for a "first-person criterion" of personal identity (psychological facts that are directly accessible to that person) rests on a mistake. The mistake is to think that such first-person statements are identity statements at all.

The term "first-person" awareness refers to the special access we have to our own mental states. As Shoemaker says,

One comes to have knowledge of these states without observing, or gathering evidence about, the physical states of one's body....⁹⁵

Furthermore, we have "special access" to the facts about our own identity over time, our memory knowledge of our own past experiences is very different from our memory of the past histories of other persons. In more general terms, we can describe this first-personal awareness as the awareness of oneself as having had a particular experience. This is to say that I can remember that it was I who visited Buckingham Palace, or who broke the window (or whatever).

It might be thought that all our experiences are by nature first-personal. However, our earlier discussion of quasi-memory may suggest that one can have the experience of a past event "from the inside" and it not be the case that the person quasi-remembers that it was she who broke the window.

The important matter that we are concerned with here is with such first-personal statements are identity statements. That is to say, when I remember visiting Buckingham Palace am I making the following claims: (1) somebody visited Buckingham Palace, and (2) I was that person? Shoemaker contends that this second claim is not part of a first-personal claim, and it is this very feature of first-person memory claims that distinguishes them from

⁹⁵Shoemaker and Swinburne, *Personal Identity*, 69.

third-person claims. For whereas it is plausible to claim that I could be mistaken as to whether it was you or some other person who visited the White House, "...if I claim to remember that I did such-and-such yesterday, it is absurd to suppose that I could be mistaken in *that* way."⁹⁶

This is not to say that I could not be mistaken in my memory claims. It is possible that I could have an *apparent* memory of some past experience although I had not participated in the past event. In this case we would regard the memory as false or delusional. But Shoemaker is not concerned with this type of case. Rather, his concern is with genuine memory claims. In these cases it does not seem plausible, according to Shoemaker, to contend that the person remembers the someone broke the window and then discovers that person was himself.

Shoemaker supports this conclusion in the following way.

It is essential to remembering one's past actions and experiences 'from the inside' that one's past self, the subject of those actions and experiences, does not enter into the content of one's memory in the way other persons do.⁹⁷

It is not the case that I can "investigate" the memory and "discover" who it was who had that past experience. Rather this information is known directly without the need for any subsequent identification. This is not to claim that the person has a direct awareness of the self in the way suggested by Swinburne earlier, for nothing that Shoemaker has said here commits him to a non-reductionist view.

We saw above in our discussion of Parfit's account that we need a more robust account of the causal relation among mental states in order to have a coherent notion of psychological continuity and personal identity. The most plausible candidate is to claim that

⁹⁶Ibid., 70.

⁹⁷Ibid., 103.

physical continuity is required for personal identity. This is to claim that diverse mental states are causally related because there is continuity of the same brain.

As Parfit suggested above,

The normal causes of memory involve the continued existence of the brain. And some or all of our psychological features depend upon states or events in our brains. The continued existence of a person's brain is at least part of the normal cause of psychological continuity.⁹⁸

I think that this is a plausible description, and Parfit is mistaken to allow that *any* cause is "the right kind of cause." To this extent, I think that the psychological aspect of personal identity is dependant in part on notions of physical continuity.

However, although it may be legitimate to claim that "the normal causes of memory involve the continued existence of the brain," the continued existence of the brain does not explain the *unity* of conscious experience. For the existence of this brain, or any other brain, is neither part of the content of the experience nor is part of the concept of self.⁹⁹

To explain the unity of conscious experience we need a different account of the relevant causal relation. To this end, we will briefly describe one possibility, presented by Shoemaker, that describes this causal relation in functionalist terms. Shoemaker writes

What the functionalist view claims is that it is of the essence of a mental state to be caused in certain ways, and to produce, in conjunction with other mental states, certain effects (behavior or other mental states). But of course, it is in conjunction with other mental states *of the same person* that a mental state produces the effect it does; and its immediate effects, those the having of which is definitive of its being the mental state it is, will be states (or behavior) on the part of the *very person* who had the mental state in question. Thus there is, on the functionalist view, a very intimate connection between the question, "What is the nature of the various mental states?" and the question "How must different mental states be causally connected in

⁹⁸Ibid., 208.

⁹⁹This is not to say that the body cannot be part of the content of the experience, but that continuity of body does not determine psychological continuity.

order to be "copersonal", i.e., to belong to one and the same person.¹⁰⁰

Thus if we understand mental states themselves in functionalist terms, then the causal relation that is at the basis of functionalist accounts is closely connected to the copersonality of these states. As Shoemaker says,

It is only when the belief that it is raining and the desire to keep dry are copersonal that they tend (in conjunction with other mental states) to lead to such effects as the taking of an umbrella; if the belief is mine and the desire is yours, they will not directly produce any joint effects.¹⁰¹

Furthermore, as Shoemaker points out, this relation of synchronic mental states is itself related to notions of diachronic identity.

But one cannot formulate these conditions for synchronic unity of mental states without invoking the notion of diachronic unity. Mental states are synchronically unified in virtue of what they jointly cause or are capable of causing, and what they cause will be something later in time with which they are diachronically unified.¹⁰²

This is to say that the copersonality of synchronic mental states is connected to the continuity of the person over time. For the fact that mental states are causally connected at one time produces, and is produced by, the causal connection of mental states at other times.¹⁰³

The second point of Shoemaker's account that warrants discussion here is the claim that

It is essential to remembering one's past actions and experiences 'from the inside' that one's past self, the subject of those actions and experiences, does not enter into the

¹⁰⁰Shoemaker and Swinburne, 93-94.

¹⁰¹Ibid., 95.

¹⁰²Ibid., 94.

¹⁰³One can recognize a parallel here between the account that Shoemaker is describing, and the concepts of person defended by Tooley and Gillett in the previous chapter. All three accounts seem to regard the continued existence of the person over time as in some way a cause of the interrelation of mental states at any one time.

content of one's memory in the way other persons do.¹⁰⁴

That is to say, first-personal awareness of past experiences is not dependant on the person's identification of his or herself as the subject of that past experience. As Shoemaker says

When one is introspectively aware of one's thoughts, feelings, beliefs and desires, one is not presented to oneself as a flesh and blood person, and one does not seem to be presented to oneself as an *object* at all.¹⁰⁵

Shoemaker's supports this position by considering a point made by Wittgenstein in the *Blue Book*, "that the use of "I" in first-person experience ascriptions does not involve the recognition, or identification, of a person...."¹⁰⁶ Wittgenstein's point is that identification goes with the possibility of misidentification, and it is peculiar to say that I could be mistaken in the identification that it is *I* who am in pain.¹⁰⁷

One might be tempted to think that Shoemaker here is, in part, criticizing Schectman's argument above that Parfit is mistaken in thinking that memory can be described in an impersonal way. Certainly both Parfit and Shoemaker defend the concept of quasi-memory. But Shoemaker's point is not that memories (or any conscious experience) can be described in an "impersonal" way, but rather that when a person remembers a past action she does not identify herself as the person who had the original experience.

Normally I can identify a past experience only as one I remember. And when this is so, there cannot arise any question whether I remember the experience or any question as to the ownership of the experience. In such cases, where one's knowledge of a past experience is based solely on one's memory, there is no room for the employment of criteria of personal identity. No question of identity arises, and hence

¹⁰⁴Ibid., 103.

¹⁰⁵Shoemaker and Swinburne, 102.

¹⁰⁶Shoemaker and Swinburne, 103.

¹⁰⁷Ludwig Wittgenstein, *The Blue and Brown Books*, (Oxford: Blackwell, 1958), 66.

there is none to be settled by reference to criteria of identity.¹⁰⁸

As I understand it, Shoemaker's point here is that when a person makes a memory claim, for example, that she was the one who broke the bank at Monte Carlo, she is not making the claims that: (i) the bank at Monte Carlo was broken, and (ii) she is identical with the person who, in her memory, is the one who broke it. Rather, as Shoemaker says,

...first-person past-tense statements can be, and frequently are, memory statements, i.e., statements that simply report what the speaker remembers, and are not conclusions from what he remembers. When this is so the speaker is not, in making his statement, identifying a remembered self as himself, and therefore the question of how he makes such an identification does not arise. I do not remember that someone broke the front window yesterday and then identify that person as myself; I simply remember that I broke the front window.¹⁰⁹

The conclusion that can be drawn here is that we should not regard first-person memory claims as constitutive of personal identity for two reasons. First, according to Shoemaker at least, a person's memory of a past event implies that the person was a "witness" to the past event (to use Shoemaker's term). In other words, psychological continuity is dependant in part on physical continuity. Secondly, first-person claims are not identity claims at all, they are simply descriptions of past events in which the person does not make any claims as to the identification of the person.

¹⁰⁸Sydney Shoemaker, *Self-Knowledge and Self-Identity*, (Ithaca: Cornell University Press, 1963), 163. It might be thought that this account of memory would seem to be at odds with Shoemaker's defence of the concept of quasi-memory. For the notion of memory described here claims that a person can only remember their own past experiences, which is what the notion of quasi-memory denies. Shoemaker would deny that there is any inconsistency since he is here talking about memory and not quasi-memory. In other words, we can define memory as the recollection a person's own past experiences, and quasi-memory as the recollection of any person's past experiences.

¹⁰⁹*Ibid.*, 142.

Persons and Personal Identity

The main goal of our discussion in this chapter was to determine what are the necessary conditions of personal identity, what personal identity consists in. Our discussion has shown that consciousness plays a central role in the continued existence of the person over time. Exactly what that role is, is less clear. According to Locke, the continued existence of the person over time is explained by the unity of consciousness, yet it also appears to be the case that consciousness itself is what enables this unity. For it is the person as a self-conscious entity that makes the different mental states copersonal.

We have shown also, it is hoped, that the understanding of the life of a person as a bundle of perceptions is inadequate. Such an account can neither explain the unity of a person's life, nor plausibly explain the type of conscious experience that seems typical to the life of persons.

This then leads us to the question of what constitutes personal identity and what is evidence for that identity. Our brief discussion of Shoemaker showed that one plausible answer is that a necessary condition of the continued existence of the person over time is to describe the life of the person in functionalist terms. This is to say that there is an interrelation of conscious mental states appropriate for the production of other mental states or behavior.

There is then the further point that Shoemaker raised with regard to the type of claim that is made by first-person memory claims. To recall, Shoemaker's contention was that such claims are not identity statements and so cannot be regarded as a criterion of identity (since their truth or falsity does not depend on the identity of the person who makes the memory claim and the person who performed the original action). Rather, first-person memory claims are made with the implicit presumption that one can only remember one's own past

experiences. The conclusion to be drawn here (and the conclusion the Shoemaker does draw, so far as I understand) is that psychological continuity is in part dependant on physical continuity.

A further conclusion to be drawn is that such first-person claims should, therefore, be regarded as evidence of personal identity, though again we should be wary of concluding that the person herself uses them as evidence. Rather, they are evidence in conjunction with the additional evidence of physical continuity.

As a negative thesis, we have not been very successful in describing the continuity in psychological terms. To recall the question that confronted us in the first chapter, we are attempting to determine, "How much psychological continuity is necessary for the preservation of personal identity?". In one obvious way, this question is simply unanswerable, for as Parfit suggests, we simply cannot pick out the moment in a gradually weakening continuity of psychological experience and say that this is when personal identity is no longer preserved. However, like Parfit we can perhaps draw conclusions as to the degrees of psychological continuity, and differentiate between cases of strong and cases of weak continuity.

Furthermore, it is part of the reductionist view that such questions are, in fact, indeterminate. That is to say that there will be occasions when we simply cannot determine whether *A* at *t1* is the same person as *B* at *t2*, and this is true from both first and third person perspectives.

A more positive response to this question is to attempt to answer it in functionalist terms. The functionalist is no more able to pick out the moment when personal identity is no longer preserved, but he can answer the question in terms of the coherence of mental states over time. Thus the functionalist can answer the question as to whether *A* at *t1* is the same person as *B* at *t2* on the basis of whether the experiences of *A* at *t1* impact in the

appropriate way (a first-personal way) on the mental states and behavior of *B* at *t*₂.

This does not tell us how much psychological continuity is necessary for the preservation of personal identity, nor does it provide us with an exhaustive set of necessary and sufficient conditions of personal identity. But it is not clear to me at this stage that the latter is possible, and nor is it clear that the former is necessary. For there is a view that claims that the individual might have self-interested concerns in the continued existence of a later individual, although the conditions of personal identity are not preserved. We now turn to the fourth and final chapter and a discussion of this view.

CHAPTER FOUR

A MATTER OF SURVIVAL

In this final chapter we look at the issue of survival, what it means for a person to survive from the present until a future time. Our purpose here is to try and formulate an account of survival that can include cases of incompetence. If we are successful then we may have provided a possible solution to the identity question with the subjective approach to decision-making for incompetent patients.¹

This chapter will also serve as a means by which we can bring together the discussion of the three previous chapters. Particularly, our discussions in this chapter will refer back to earlier discussion of the concepts of person and personal identity. In so doing, we will be examining the interrelation of the concepts of person, personal identity and survival. *Prima facie*, a plausible description of this interrelation is that in order to talk about personal identity and personal survival it must be assumed that the conditions of personhood are met. In fact this is a questionable assumption, and we will be looking in detail at the question of whether you or I can survive the loss of personhood.

The above question has important implications for the identity question with advance directives. To see this we need recall from the first chapter that Buchanan and Brock defend the "moral authority" of advance directives by claiming that since the incompetent patient is not a person, *ex hypothesi* the patient cannot be the same or a different person. But such a response might be unsuccessful if it could be shown that despite the loss of personhood, other

¹To recall, the subjective approach is where the previously expressed wishes of the patient while competent are used to determine treatment decisions.

relevant identity relations were still present between the competent and incompetent individual. In such a case it would be appropriate for the competent individual to have a self-interested concern for the welfare of the later incompetent individual. Thus if we are to defend Buchanan and Brock support for the moral authority of advance directives in particular, and the subjective approach to decision-making for incompetent patients in general, we need to show that these other identity relations are not present and/or self-interested concern would be implausible.

This is not to say that we should defend Buchanan and Brock's position. As we shall see, part of our purpose here is to defend the moral authority of advance directives while at the same time retaining the concept of person.

As a way of introducing the central focus of this chapter we can consider the following claim by Locke.

In this *personal Identity* is founded all the Right and Justice or Reward and Punishment; Happiness and Misery, being that for which every one is concerned for *himself*, not mattering what becomes of any Substance, not joined to, or affected with that consciousness.²

Locke claims here that there is a close connection between personal identity and self-interest. In fact the claim can be stated in stronger terms: A necessary condition of a person, *A*, at *t1* having a self-interested concern for the welfare of a future person, *B*, at *t2* is that *A* is the same person as *B*, which is to say the conditions of personal identity have been met.

However, uncontroversial though this claim might seem, it is, in fact, the subject of considerable debate. The success of this claim rests on two further considerations being true: First, *A* at *t1* must be the same person as *B* at *t2*. Second, in order for *A* at *t1* to have a self-

²John Locke, *An Essay Concerning Human Understanding*, ed. Peter H. Nidditch (Oxford: Clarendon Press, 1988), 342.

interested concern for the welfare of a future individual, *B*, at *t*₂, *A* and *B* must both be persons.

In this chapter I shall present a number of arguments against these claims and examine a defence of the claim that an individual can have a "self-interested" concern for the interest of a future individual even though the future individual is not a person, and is not identical (as described by the conditions of personal identity) with the present individual.³

The chapter is divided up into the following sections. We begin by describing the identity and survival relations and examining the differences between the relation, and look at why a number of contemporary philosophers have favored survival over identity. Second, we investigate the conditions of survival and look at two recent accounts of survival. Third, we look at the concept of relative identity and discuss whether *person* is a sortal term. Four, we are then in a position to turn our attention to the question above, "Can you or I survive the loss of personhood?" An answer to this question will rest on whether you or I are *essentially* persons. Furthermore, an answer to this question will provide us with an answer to the identity question with regard to advance directives that is central to this work.

Identity and Survival

Throughout much of this work we have been discussing the concept of personal identity, and have been attempting to understand the identity relation as it applies to the class

³To avoid confusion, the following terminological points are made. I will reserve the term *individual* to stand for a member of the class of human nonpersons. Individual is preferred to the even more neutral *entity* on the grounds that "entity" can refer to almost anything, and my preference is for a term that can be understood as being adjacent to the concept of person, that is to say where we can pick out particular members of a class by their "individual" characteristics. Furthermore, I will use the term *self-interested* in a neutral sense, that is to say we can talk about the self-interest of individuals and not just the self-interest of persons or selves.

of persons. However, there might be a problem with such an endeavor. The problem can be understood if we consider the following remark by David Lewis.

Everything is identical to itself. Nothing is identical to anything else. There is never any problem about what makes something identical to itself; nothing can fail to be. And there is never any problem about what makes two things identical; two things never can be identical.⁴

The identity relation that Lewis describes here is the relation of *strict identity*: if x is identical to y , then any property true of x is also true of y , and vice versa. According to Lewis, if we can pick out two distinct individuals then this would imply that there is at least one property that is not shared by both x and y , and hence the condition of strict identity has not been met. The conclusion to be drawn with regard to the present endeavor is clear; since the incompetent patient does not share all of the properties possessed by the prior competent person, we cannot legitimately talk of a *strict identity* between these two individuals. As Harold Noonan says,

How, then, can it make sense to ask, for example, what makes a person existing at one time identical with a person existing at another? If the person existing at the earlier time is identical with the person existing at the later time, the question is a request for an account of what makes a thing identical with itself. While if the earlier person is distinct from the later one, the question is a request for an account of what makes *two* things identical. In either case, it is unanswerable.⁵

However, it is not clear that Lewis or Parfit (see below) is claiming what Noonan is perhaps suggesting. For a supporter of *strict identity* (such as Lewis) claims that if x is identical to y , then at property true of x at $t1$ is true of y at $t1$. By phrasing the matter in this way we allow for the possibility of change over time. Thus unless we want to credit Noonan with the position that every property of the entity must remain unchanged for the duration of the entity's existence, we must understand as meaning something else. In fact much rests

⁴David Lewis, *On the Plurality of Worlds*, (Oxford: Blackwell 1986), 192-3.

⁵Harold Noonan, *Personal Identity*, (London: Routledge, 1991), 106.

on what Noonan means above when he states, "...if the earlier person is distinct from the later one, the question is a request for an account of what makes *two* things identical." For the question at hand is whether the earlier and later persons are distinct.

There is another aspect of Lewis' claim. This is the contention that there is just one account of identity regardless of the type of entity being discussed. To put the matter another way, identity is not *sortal-relative*; there is just one relation governing boats, bats, cats and persons. Hence it is confusing or even question-begging to talk of *personal identity* since there is no unique or distinct application of the identity relation to the class of persons.

Does this mean that our discussions of personal identity have been for naught, since, on the evidence above the competent and incompetent persons cannot be identical? I do not think that this conclusion does follow. The problem here has to do with the concept of identity. Lewis is correct, of course, in asserting that the strict notion of identity dictates that two entities cannot be identical. However, an important question is whether it is this sense of identity that is being used when we use the term *personal identity*. There are a number of considerations that suggest otherwise.

First, we can make use of a distinction put forward by Parfit that we have seen previously. Parfit distinguishes between the *logic* of identity and its *nature*.⁶ According to its logic, identity is "all-or-nothing"; the relation holds completely or not at all. According to its nature, however, identity has degrees. Parfit's support for this distinction is based on another that we are now familiar with, the distinction between the Simple and Complex views of personal identity. To recall, this is the claim that

...about the fact of personhood and personal identity, there are two views. According to the first [the Simple view], these facts have a special nature. They are further facts,

⁶Derek Parfit, "Later Selves and Moral Principles," in *Philosophy and Personal Relations* ed. A. Montefiore, (Montreal, Canada: McGill-Queen's University Press, 1973), 137-169.

independent of certain more specific facts; and in every case they must either hold completely, or completely fail to hold. According to the second view [the Complex view], these facts are not of this nature. They consist in the holding of the more specific facts; and they are matters of degree.⁷

As we know, Parfit supports the Complex view and contends that personal identity consists in the holding of more specific facts relating to psychological connectedness and continuity. Since these more specific facts can be held to varying degrees, if the identity relation is to be applied to persons then it cannot be understood simply in terms of its logic. Rather, it must be accepted that in its nature, identity has degrees.

This is not to suggest, however, that Parfit's analysis of the logic of identity can be equated with the notion of strict identity. Rather, the point here is to ask what notion of identity is suited to the concept of *personal identity*. As we see below, a number of considerations lead to the plausible contention that the most useful notion of identity that can be applied to the concept of personal identity is what Parfit terms the "nature" of identity.

To return now to our concept of strict identity, we can see that this relation either reflects only the logic of identity and ignores its nature, or can be supported only if we adopt the Simple view and claim that identity is not a further fact. However, our discussion of non-reductionism in the previous chapter showed that this latter option is not viable. Hence when we talk about *personal identity* we are not using identity in this strict sense.

There is a second reason why we should be wary of the strict sense of identity, and this has to do with the formal properties of the identity relation. Strict Identity is termed an *equivalence relation* in that it comprises the following relations: *reflexivity* (*x* is identical with itself), *symmetry* (if *x* is identical to *y*, then *y* is identical to *x*) and *transitivity* (if *x* is identical to *y*, and *y* is identical to *z*, then *x* is identical to *z*). However, our discussion of psychological

⁷Parfit, "Later Selves," 138.

continuity in the previous chapter suggested that the condition of transitivity cannot be met. For example, Bertrand Russell at age 90 could remember winning the Noble Prize for Literature at age 63, and at 63 he could remember being arrested at age 44, but he could not remember being arrested when he was 90. Thus if personal identity is understood in terms of memory, then since memory is not transitive (certainly not in the case described) we must conclude that the relevant notion of identity is not strict identity.

This is not to say that supporters of the memory criteria hold to this strict notion of identity. As we saw in the last chapter in our discussion of Parfit, it is possible for one to support the memory criterion of personal identity and give up the transitivity of strict identity.

Third, it is controversial whether the notion of strict identity could be legitimately applied to *any* discussion of diachronic identity. If we think of an entity existing at two different times, t_1 and t_2 , then the statement "present at t_2 " is both true and false of this entity. Hence if we want retain the notion of strict identity then we must conclude either that the entity cannot persist through time or that on some occasions the reflexivity condition is not met.

There are a number of ways that we can avoid these problems with the strict notion of identity. The simplest, and perhaps ultimately the most successful, is to claim that in discussions of personal identity we are using the term "identity" in a "loose and more popular" sense wherein two entities can be described as identical even though not every property possessed by one is also possessed by the other. But to avoid the vagueness that this option suggests, we would need to formulate some criteria by which the entities could be described as identical even in this loose sense.

A second option is to distinguish between those properties of the entity that belong to the entity throughout its existence, and those which belong simply for part of its existence.

As Wiggins says, we need to distinguish between

...sortal concepts which present-tensedly apply to an individual *x* at every moment throughout *x*'s existence, e.g. *human being*, and those which do not, e.g. *boy* or *cabinet minister*. It is the former (let us label them, without prejudice *substance-concepts*) which give the privileged and...the most fundamental kind of answer to the question 'what is *x*?'. It is the latter (one might call them *phase-sortals*) which, if we are not careful about tenses, give a false impression that *a* can be the same *f* as *b* but not the same *g* as *b*.⁸

Wiggins remark here is directed against proponents of relative identity (that *a* can be the same *f* as *b* but not the same *g* as *b*). However, the claim can also be used to define strict identity in a way that avoids some of the problems described. For we can describe the properties of the objects as being either temporally essential or not. That is to say, we can distinguish between those properties that describe the object for its entire existence, and those that describe it only at certain times. In this fashion, *a* can be identical to *b* although does not share all the same properties with *a*.

Another option is to abandon the identity relation and to explain the continued existence of the person by some other means. One method is to follow John Perry and talk in terms of *person-stages*. In this way, the life of a person is thought to be made up of a series of interconnected stages, each with a different temporal and spatial description.⁹ Understood in this fashion, the question of personal identity can be restated as a question of describing the "unity relation" for these person-stages. The benefits of adopting such an approach are that we avoid having to accommodate the formal properties of identity, for it is no longer necessary to claim that one person-stage is identical with any other person-stage. However, the drawback to such an account is that *person-stages* are as far removed from our customary notion of personal identity as is the notion of strict identity. In the normal case

⁸Ibid., 1.

⁹John Perry, "Can the self divide?" *Journal of Philosophy* 73, 463-88.

of things, we talk of "persons" not "person-stages." Furthermore, describing the unity relation for persons is no easy matter as we saw in the previous chapter.

A second method is to talk not of identity but of *survival*. There are a number of advantages with such a strategy. First, one might claim that the notion of "survival" better fits our own self-interested concerns with regard to our future existence. When someone makes a statement such as, "I hope I live until I'm eighty-five" or "I do not want life-sustaining treatment to be performed if I should be incompetent" one can plausibly maintain that the implied sentiment behind such claims is that the person making the claims imagines herself to have survived as the later person, rather than being identical with the later person. To put the matter another way, when a person expresses a desire to live until he or she is eighty-five (to see his or her first great-grandchild) the desire can be understood as a desire for continued existence at a certain level of conscious ability (or whatever). This person is not making the claim that she hopes that all the features or properties true of her now will be true of her then. For she may wish that at this later time she is a "different" person (has developed a liking for children). This point is more pronounced in the second example, where a person might request that life-sustaining treatment not be performed at a later time. One may contend that for someone to have such a concern, he or she must assume that he or she has survived as this later person. For *prima facie* the person at this present moment would not be concerned for the welfare of the future person if he or she thought that he or she did not survive. This is not to say that the person would be correct in his or her assumption. One could claim that just as it would be rather odd for me to be concerned about my welfare after I have died, so it would be odd for me to be concerned for my welfare as a patient in a PVS.

Second, such an approach avoids the problems associated with the identity relation

that have been discussed. Survival is a looser relation than identity since transitivity is not required for survival; if *a* survives as *b*, and *b* survives as *c*, it does not, and need not follow that *a* survives as *c*. Second, to say that "*a* survives" need not imply (either grammatically or in fact) that *a* survives *as* a particular type of entity. When one talks of "personal identity" it is implied that we are talking about the identity relation as applied to persons, and that if the individual ceases to be a person then, *ex hypothesi*, she cannot be the *same* or a *different* person. (This, to recall, is the position of Buchanan and Brock.) But when we talk of "survival" it need not be assumed that in order to survive, the individual must remain a person, or remain a member of any particular kind or class. Thus talk of survival does not imply that the individual cannot survive the loss of personhood, whereas this implication is present in talk of personal identity. In this regard, the demand for an explanation of "*a* survives as *b*" becomes less necessary, for in order for *a* to survive it is not, *prima facie*, necessary that it survive *as* one particular thing or another. With regard to the concept of person, such a view would have to be rejected, however, if it could be shown that we are *essentially* persons, that is to say, if a necessary condition of my continuing to exist is that I continue to be a person. The question remains how it is that the entity survives. This is to ask what are the conditions of survival and we turn to this question later on in this chapter.

Furthermore, and as a consequence of the above, the notion of "sameness" that is part of the concept of identity is less apparent in the case of survival. As Wiggins says

...if someone tells you that $a=b$, then you should always ask them 'the same *what* as *b*?'¹⁰

But it is not clear that when one says "*a* survives from *t*₁ until *t*₂" that there is the same requirement for identification of the relevant sortal term. That is to say, when we say that

¹⁰David Wiggins, *Identity and Spatio-Temporal Continuity*, (Oxford: Basil Blackwell, 1967), 1.

a person survives from one time until a later time, we need not be claiming that this is the "same person". If it is the case that the individual can survive from one time until a later time despite the loss of personhood, then we need not describe this in terms of the person being the "same person." This is not to say that the survival relation is without its own challenges. Foremost amongst them is the need to determine what it means for the individual to survive. That is to say, what are the conditions of survival?

There are different opinions as to the relation between identity and survival. According to Parfit, the confusions that have occurred in discussions of the persistence of the person over time are due to the mistaken apprehension that identity is what matters in survival.¹¹ Furthermore, Brennan claims that while survival is a necessary condition of identity, instances of identity occur when the survival relation is at its strongest.¹² If we think in terms of personal identity then it can be claimed that in order for *a* at *t1* to be the same person as *b* at *t2* then it is necessary that *a* survives as *b*.

Notwithstanding these differences, for what amount to pragmatic reasons we will favor the view that is supported by Brennan, Parfit, and Unger. That is to say, for the reasons cited above we will assume that the survival and identity are distinct although they may overlap.

¹¹ Derek Parfit, *Reasons and Persons*, (Oxford: Oxford University Press, 1984), and "Lewis, Perry, And What Matters" in *The Identities of Persons*, ed. Amelie Rorty, (Los Angeles: University of California Press, 1976). See also David Lewis, *Survival and Identity in The Identities of Persons*, ed. Amelie Rorty.

¹² Andrew Brennan, *Conditions of Identity: A Study in Identity and Survival*, (Oxford: Clarendon Press, 1988).

The Conditions of Survival

We now turn to issue of what it means for the individual to survive from one time until a future time. That is to say, we are attempting to discover the necessary and sufficient conditions of survival.

Our initial concern is with what we mean by "survival." We can identify two different usages of the term as they appear in works that are discussed here. First, there is the notion of "survival" that is used by Peter Unger in a recent work.¹³ Although Unger states that the "central topic" of this work is personal identity, he declares that since he is "not overly impressed with the label "personal identity," our main question concerns only the conditions of our survival itself" Unger's reason for preferring "survival" over "personal identity" is that he distinguishes between "the question of what are the conditions of my survival and, on the other hand, what are the conditions for me to remain a person (though as Unger himself says, "if I am essentially a person, then these questions will receive answers that are essentially the same").¹⁴ This is consistent with what we said previously, that one can talk about an entity surviving and this not imply that the entity has to survive *as* particular kind of entity. In other words, the survival relation does not imply an identity of the appropriate sortal terms.

Second, there is the notion of survival that is used by Andrew Brennan.¹⁵ According to Brennan

...(W)e can think of survival as a kind of *default option*: it is the feature we default

¹³Peter Unger, *Identity, Consciousness and Value*, (New York: Oxford University Press, 1990).

¹⁴Ibid., 7.

¹⁵Andrew Brennan, *Conditions of Identity: A Study in Identity and Survival*, (Oxford: Clarendon Press, 1988).

to when conditions for identity are satisfied but something else rules out an identity claim, or when the matter of identity is indeterminate.¹⁶

Brennan's notion of survival is thus different from Unger's in that for Brennan "survival" is used in distinction to identity. This distinction is supported by two types of case. First, when the conditions of identity are not met. For example, where the presence of "rival candidates" in a case of fission negates the identity relation according to the transitivity condition, or because the level of psychological continuity might be insufficient for personal identity, though perhaps sufficient for survival.¹⁷ Second, when the degree or level of survival is so high that the conditions of identity are met. In these cases, the strongest cases of survival would be cases of identity.

Yet, having once found a maximal set of stages constituting a unified particular, the question of its identity is no longer a matter of degree of argument. Identity, or unity, having been established, that item - whatever it is - obeys the classical laws of identity.¹⁸

Importantly, Brennan does use the term "personal survival."

Let us designate the relations that constitute personal survival 'the survival relation'. The survival relation is thus a relation between stages where one survives--according to our conditions--in or as the other to a suitably high degree...a maximal set of stages connected by the survival relation and with not too much change overall will

¹⁶Ibid., 17.

¹⁷There is a sense in which I am being unfair to Brennan. On pages 326-332 of *Conditions of Identity* Brennan discusses the debate between Lewis (Rorty, *The Identities of Persons*,) and Parfit, (*Reasons and Persons*) as to whether the identity and survival relations are distinct. Brennan discusses two worlds, a world with fission and a world without. He argues that "in a world of fissioning particulars, there is at least a theoretical possibility that survival in my sense is constitutive of identity", whereas "in world without fission, but continuing replication, the situation is...rather different." In this second world the survival and the identity relations are distinct. Somewhat arbitrarily, I am going to declare this world a "world-without-fission" and thus I am going to take Brennan as supporting the general thesis that the survival and identity relations are distinct. Support for this declaration can be found by taking note of the type of cases of survival and personal identity that are discussed in this work.

¹⁸Andrew Brennan, *Conditions of Identity*, 238.

constitute the history of one person.¹⁹

However, as Brennan takes care to point out, "on my official view, survival is a relation between *stages*...". That is to say, as opposed to *persons*.

Notwithstanding the above distinctions, it is fair to say that the differences between Unger's and Brennan's accounts as regard the meaning of "survival" are not substantial. Both would agree on the following claims: (i) Survival is a more primitive notion than identity, which is to say that the conditions of survival can be met where the conditions of identity are not. (ii) Survival, unlike identity, can be understood as having degrees. (iii) Survival can be a relation between different kinds of entities, this is to say, in order to survive, the individual need not survive *as* a particular kind of entity.

A further similarity that warrants mention is that both Brennan and Unger share a view of the world as physical and a view of continuity that is causal. It may be a mark of any account of the survival of an object over time (as opposed to its diachronic identity) that the account takes a physical approach. Certainly, both regard the "underlying reality" of the world to be physical, and to be subject to physical laws. Furthermore, both endorse the view that in order for the individual to survive from one time until a later time, there must be continuity of brain. What is presumed here is a material view of the mind such that psychological states can be regarded as the product of brain states. That is to say, there is a clear rejection of any type of dualism that regards the mind and mental states as distinct from the brain and brain states.

It is an interesting question why supporters of the survival rather than identity relation tend to adopt a physical view. I do not think that we can claim that any survival account *has* to adopt such a view, that one cannot plausibly deny that in order for the person to survive

¹⁹*Ibid.*, 329-30.

from now until later certain physical features must persist. For one can claim as many do that a person can survive his or her own death. However, perhaps one can say that the motivation for talking in terms of survival rather than identity, is based on the recognition that the world is predominantly physical, and that objects appear to persist through time though they do not remain sufficiently identical, qualitatively, to count as numerically the same object.

Brennan

According to Brennan the most satisfactory way to understand the existence of an object over time is to think in terms of the *continuity* of that object. By thinking in terms of continuity we can accommodate our intuitions that although objects undergo change over time the same object can persist over time. What does it mean to talk of the "continuity" of the object? Brennan's answer is that we can understand the spatio-temporal continuity of any extended object as composed of the interrelation of *stages* or *temporal parts*.

Let us think of things now in terms of their *stages*. Each extended object (extended through space and time) consists so we suppose, of successive stages.²⁰

Thus Brennan supports, in part, the type of account suggested by Perry above (though Brennan is not concerned specifically with identifying the unity relation for persons).

A further consideration is that the stages of the object are in some way causally related to each other, and that it is by virtue of this causal relation that we can understand the continuity of the object.

The conception of ordinary things as continuants makes sense of one prevalent feature of our experience. This is that how a thing is one time very much depends on what happens to it at others.²¹

²⁰Ibid., 23.

²¹Ibid., 23.

We may think of a continuous object as composed of the relation of successive stages, and that the state of the object at a later time is dictated by what occurred to the object at any earlier time. According to Brennan, the following two general rules apply to the relation of the stages of any object in order for that object to be spatio-temporally continuous:

- (1) The stages of a persisting object succeed each other in a lawful and regulated manner.
- (2) If a latter stage manifests certain differences from an earlier stage, then something has affected the earlier stage, or some intervening stage, and that thing, whatever it is, is responsible for these differences.²²

If we turn now to the conditions of survival, this notion of the interrelation of stages provides us with the basis for determining what it means for one object to survive as another.

Brennan presents us with the following case.

For example, a child designs a marvelous bridge using a plastic construction set. Her brother, inspired by her work, models the bridge exactly using wooden blocks and glue. The original bridge is thereafter taken apart and the plastic components of which it was made are used for the next creation. Has the child's original bridge survived? The answer to this question is that it has survived *as* the wooden bridge.²³

This example of the toy, we are told, gives us "all we need for sketching an account of conditions for survival". The following three conditions are presented as sufficient for an object to survive from one time until a later time.

Structure Condition: the brother's copy is structurally the same as the original; that is, it has components of the same size and shape as those of the original arranged in the same relative positions;

Causal Condition: the causal account of how the brother's model was constructed makes reference to his sister's original model; in particular, his version was obtained from hers by being copied, albeit in a different material;

Matter Condition: one item survives as another when, in addition to the conditions on

²²Ibid., 24.

²³Ibid., 26.

cause and structure being satisfied, the item as which the original survives is made of material similar to the original.²⁴

As is clear, cases of survival cover a very broad range. They range on the one hand from cases where the former object barely meets the conditions to those on the other hand where the conditions are clearly met. Since it is appropriate to talk of degrees of survival, we might say that we have a clear case of survival, or a strong case of survival when it is the self-same components that are used to construct the earlier and later object. Thus in the case of the siblings' bridges above, we would have a case of strong survival -- perhaps even identity -- if one bridge was "reconstructed" out of the original materials after the original had been disassembled. Conversely, we would have a weak case of survival, and perhaps even a case where survival is indeterminate, when the structure, causal and matter conditions are met to a lesser degree.

Brennan claims that only the structure and causal conditions are *necessary* for survival. It is not necessary that in order for the object to survive from an earlier to a later time, the object must be composed of the same or similar material throughout. However, all three conditions are *sufficient* for survival.

If one item possesses lots of material similar to that possessed by another, if the second item has that material structured in the same way as it is in the first, and if the first item participated causally in bringing about this sharing of structure and matter, then the first item survives as the second.²⁵

We turn now to the question of how these conditions apply to the survival of the person. An important point to begin with is that Brennan contends that there is no one, clear conception of person. Rather, not only is "our everyday concept of the person confused," but there are also a number of alternative conceptions that are applicable in different

²⁴Ibid., 26-27.

²⁵Ibid., 27.

circumstances. Brennan claims that the confusion here is due in large part to the fact that our everyday conception of person is both physical and psychological. As we discussed at length in the last chapter, there is a substantial body of opinion that supports the view that personal identity can be understood in terms of psychological continuity. Such a view has its merits and seems most plausible if our intention is to capture the supposed first-person nature of personal identity. But such a view is less satisfactory in describing personal identity in third-person terms. Fallible though the evidence might be, our identification of another person over time is based on physical evidence.

A possible way to reduce this confusion would be to provide an account of personal survival that combines both physical and psychological considerations. This is Brennan's intention.²⁶ The solution that Brennan cautiously endorses is that the person has survived if the brain survives. Support for this view is taken from the consideration that the brain is necessary for the continued functioning of the living human being, and for the presence of psychological states.

In discussing what it means for a person to survive, our task is to determine whether Brennan's three conditions of survival can be usefully and successfully applied in discussion of the survival of the person. *Prima facie*, we might find it difficult to see how these three conditions can be applied to the psychological aspects of the life of the person, for to talk of the structural, causal and matter conditions implies a notion of physicality that might be at odds with our understanding of psychological states. Brennan's responds to this concern in two ways. First, we may understand the use of these terms as "metaphorical."

I might...point out that we need not think of the issue in terms of structure and matter, but simply in terms of developing psychological analogues of the notions of

²⁶For the sake of brevity I will use the term *personal survival*, though with the caveat that this is not meant to imply that survival implies sameness of person.

structure and matter. We may have no idea what such analogues are; but this should not inhibit us from thinking about them. Nor, incidentally, is my usage here a great departure from common sense. If we do think -- as neo-Lockeans tend to -- of the human mind as like the software that perhaps runs on the body and the brain as its hardware, then we can reflect that the software notions like *program*, *data*, *array*, *loop*, and so forth are all easily subsumed under our labels.²⁷

However, we do not have to claim that these terms can be used only in a non-literal way, for to recall our discussions of the past chapter, one problem that affects the psychological account of personal identity is to determine how the different psychological states (whether synchronic or diachronic) are united. As Brennan suggests, we may think of the structural and causal conditions as somehow providing a possible explanation (and it is possible that the matter condition could be relevant also). As was suggested in the last chapter, a functionalist account of personal identity may be helpful here, and such an account would be consistent with Brennan's three conditions.²⁸

Brennan claims that his conditions of survival are able to resolve a problem that has beset psychological views, namely, that problem of amnesia. Amnesia is thought to be a problem for such accounts in that although I may have forgotten much if not most of the events in my past life, I still want to claim that I am the same person as the person who participated in those events. In other words, amnesia would undermine the claim that memory is a necessary condition of personal identity.

Brennan present us with the following case. He asks us to consider a person, Pamela,

²⁷Andrew Brennan, *Conditions of Identity*, 302.

²⁸I admit that this is a controversial claim. It is clear that the causal condition is relevant, and that the matter condition unimportant. On any functionalist account the matter from which the object is constructed is not considered important in that two objects may be functionally identical though constructed of different matter. The structural condition is the difficult one to assess. If by "structure" we mean the relation and interconnection of the parts of the object, and this structure determines the relevant causal relations, then it may be necessary that in order for two objects to be functionally identical they must be similar in structure.

who because of a serious infection has lost "all memories of her illness and those extending back to some time before the illness as well." This is an example of what is termed *retrograde* amnesia. To understand this we can consider Pamela as existing at three times, $t1$, $t2$ and $t3$, and that the infection occurs at $t2$. Thus she will have experiences at $t2$ that are then lost. The result is that although at $t3$ and just before $t2$ she can remember the events at $t1$, and at $t3$ she can remember the events at $t1$, at $t3$ she cannot remember the events at $t2$. On a psychological account we would have to conclude, more or less, that the conditions of identity are not met in such a case, and hence we would have to conclude that Pamela is not the same person throughout.

Brennan claims he can draw a different conclusion.

Our *causal* condition requires that one stage survives as or in another stage provided the structure and matter of the later stage is causally dependant in an appropriate way, on the structure and matter of the earlier one.²⁹

This condition is met in Pamela's case. For according to Brennan, one can plausibly maintain that the infection at $t2$ caused her amnesia by affecting the structure and matter of her brain. It should not be thought here that Brennan is simply preferring a physicalist view to a psychological view, for his three conditions would still be pertinent if understood more metaphorically. For Pamela's beliefs and other psychological states apart from memory at $t2$, and her psychological make-up at later times will be causally connected to other psychological states at earlier times. For example, her belief that she was ill at $t2$ will be causally related to later beliefs and other psychological states.

In turning now to the question of what it means for the person to survive, it is important to recall that Brennan maintains that there is no *one* concept of person. Rather, there are a number of different concepts that can be applied in different contexts. Brennan

²⁹Andrew Brennan, *Conditions of Identity*, 303.

distinguishes amongst three levels of the concept of person: *low*, *intermediate*, and *high*. On the intermediate level,

The relata of the survival relation...are, on such an account, psychological states, episodes, and data conceived as having structure and being capable of entering into causal relations with other such states and episodes.³⁰

According to Brennan, an example of an intermediate concept of person would be the type of account of Parfit. As we saw in the previous chapter, Parfit claims that personal identity can be understood in terms of the psychological continuity and connectedness "with the right kind or cause." Thus we can understand Parfit as supporting an account that can be described in terms of the causal relation of psychological states.

A higher concept of person would be one in which the notion of consciousness, particularly self-consciousness, plays a more prominent role. Brennan regards Locke's account of personal identity as an example of such a higher concept. In such accounts

...with the emphasis on co-consciousness, the relata of the survival relation are *conscious*, or even *self-conscious*, episodes, which are in communication with each other thanks to memory, intention, and other high-level psychological operations.³¹

According to Brennan, a consequence of this high-level conception of the person is that it does not matter what changes might occur to the structure and matter of the body that supports the conscious states, so long as there is sufficient "communication" amongst the conscious states. That is to say, there could be psychological continuity between different persons even though there had been considerable change both in the physical make-up and the psychological make-up. To say that there could be changes in the latter might appear odd, but the claim here is that different psychological states could produce similar psychological effects. As Brennan says,

³⁰Ibid., 332.

³¹Ibid., 332.

Just as the conscious episodes could be supported by differently configured bodies, so too they might be supported by differently configured minds.³²

However, according to Brennan it is this lack of connection between the psychological and physical "levels" that raises problems for a high-level conception of person. To recall, Brennan adopts a physical view of the world -- that psychological states have physical causes, for example, dementia can be the result of cerebral arteriosclerosis. Thus a drawback to the high-level conception, and the implied irrelevance of physical factors, is that it seems to be at odds "with the nature of explanation in the life sciences."

So long as the nature of the conscious episodes experienced seem unchanged to me, so long as the behavioral repertoire of my successive stages changes only gradually and along clearly intelligible lines, then my identity need not be bound up with any sameness either of bodily structures or of psychological structures either.³³

If we turn now to the low-level conception of survival, a sufficient condition of survival is based on the notion of *continuity*.

Provided we have continuity of enough of a living human body, we have sameness of person. Thus persons are nothing other than living human bodies.³⁴

Brennan refers to this conception of person as an "administrative" one.

This conception has at least administrative convenience on its side. In the psychiatric or neurological ward, where there are patients with dreadfully disordered lives, we can still apply such a criterion of personhood.³⁵

Simply stated, this administrative conception may seem to be too low-level to incorporate any role for the psychological features that we think are important for the continuity of the

³²Ibid., 334. Another way of looking at this would be in functionalist terms. Our psychological states could be arranged or structured in different ways yet produce the same results.

³³Ibid., 334.

³⁴Ibid., 335.

³⁵Ibid., 335.

person. To rectify this omission Brennan contends that in speaking of the continuity of "enough of a living human body" we are speaking primarily of the brain. Thus we should allow

...the brain, or the brain and certain other parts of the nervous and regulatory systems of the body, to be the components of greatest significance in establishing identity.³⁶

As Brennan says, the faculties of "personality, memory, perception, concept-use" are associated with the brain, and without a functioning brain I would lack those qualities that are important for being a person. The brain may thus be regarded as playing the most crucial role in the continued existence of the person, and changes in the matter and structure of the brain cause changes in our psychological features.

Where does this leave us in determining the conditions for the survival of the person? The conclusion would appear to be that the person survives so long as the brain survives with, it is presumed, the capacity for consciousness. If Brennan regards that "administrative" conception of person to be a legitimate one, then it would appear that neither psychological continuity nor *coherence* -- by this we mean that the psychological states are unified -- is necessary for personhood. We can conclude, therefore, that incompetent patients are persons. For example, if we think of a patient who is suffering from Alzheimer's dementia, and is regarded as incompetent on the grounds that she cannot orientate herself to "person, time, and place". Nevertheless such a patient might meet Brennan's administrative conception of person.

An issue that remains is to determine what makes one conception of person more appropriate than the others. Brennan appears somewhat defeatist about this surfeit of riches.

In the present situation, we have several developments of the concept of person available. These are in competition with one another. But none is yet worthy of

³⁶*Ibid.*, 336.

commendation over its rivals.³⁷

The implication here is that the decision to adopt one conception over another is a pragmatic one, and is dependant on the particular case. Thus if our concern is to assign praise or blame where the intentions of the individual are relevant, we would find more success in adopting a higher-level conception of person. Conversely, if we are concerned with the minimal survival of the person, we should adopt a lower-level conception.

Unger

We turn now to examine the second account of personal survival, the account put forward by Peter Unger.³⁸ Unger is attempting to present a "physical approach to our survival" and is concerned with answering the question, "what is involved when a person, like you or me, survives from the present time to some future time..?" As we previously mentioned, Unger is careful to distinguish this question from the question, "What are the conditions for me to remain a person?". The importance in distinguishing these questions is that it should not be *presumed* that my status as a person determines the conditions of my survival. However, if it turns out that you or I are *essentially* persons, then this would mean that my existence as a person does determine my survival. Unger claims, in fact, that you and I are essentially persons, and thus cannot survive the loss of personhood. This point will be explained further as we go along.

We will now turn to an examination of Unger's physical approach to our survival. As we mentioned at the start of this section, both Brennan and Unger support a physical view

³⁷Ibid., 339-40.

³⁸Peter Unger, *Identity, Consciousness and Value*, (New York: Oxford University Press, 1990).

of survival. Furthermore, both share the same motivation for this view, that the world is physical and we interact with the world in a physical way. Unger's support for the physical view is based on his understanding that the psychological features that are necessary for our continued existence cannot exist independently of the body, but depend on physical objects for their realization. Particularly, they depend for their realization on the continued capacity of the brain to realize these psychological features.

A person's physical part and structures are important to her survival to her *only insofar as they continue to support, and to realize, her basic psychological capacities*.³⁹

As Unger points out, the main thesis of his physical account of our survival is based on the contention that our physical "part and structure" is necessary for the realization of the psychological features and is of *derivative* importance. By this Unger means that this physical part and structure does not have *basic* or *independent* value, but is valued in that without it, the person could not survive.⁴⁰ To claim that the continuity of the physical part and structure, in other words, the brain, is of derivative value, is to claim that it is of *indirect* and *relative* value. Physical continuity is indirect in that it

provides for what is more directly important: the future existence of a person's basic mental capacities and, more directly important still, the future existence of that person herself.⁴¹

This is to say that if there were some way for the person's psychological features to survive without any physical realization then physical continuity would be of little importance. Physical continuity is relative in that it is dependant on a view of the world as very largely physical in nature.

³⁹Ibid., 71.

⁴⁰Ibid., 113.

⁴¹Ibid., 117.

The most concise description of Unger's physical account of survival can be stated as follows.

The person X now is one and the same as the person Y at some time in the future if, and only if, (1) there is sufficiently continuous physical realization of a core psychology between the physical realizer of X's core psychology and the physical realizer of Y's core psychology, and [probably] (2) {some suitable clause for ruling out unwanted cases of branching}.⁴²

To understand Unger's view, we need to look more closely at what he means by "sufficiently continuous physical realization". We may identify two central notions in Unger's account as described here: (i) the distinction between psychological activities and psychological capacities; (ii) the distinction between core psychology and distinctive psychology, and the claim that survival involves only the realization of core psychology.

If we think of a working brain we can ask what level of functioning is necessary in order that the psychological features are realized. According to Unger, in order for the person to survive it is necessary the brain retain the capacity to realize these features, it is not necessary that the brain be "functioning properly throughout." As Unger says,

...appropriate physical *functioning* is needed only for psychological activity. Continuity of adequate physical *structure*, by contrast, will suffice for the capacity for that activity.⁴³

Unger supports this contention by claiming that people can survive long breaks in psychological activity. He imagines the case of *super-freezing and super-thawing*. A machine freezes the molecules of the person "in their tracks." Some time later the person is thawed out. For Unger, it would be fair to conclude in such a case that the person survived such a process. As long as the brain retained the capacity for the realization of the psychological features, the person could survive the super-freezing. However, it is important to note that

⁴²Ibid., 107.

⁴³Ibid., 105.

such a view depends on the brain having realized the psychological activity at one time. If the brain had never realized the activity at one time we would not be in a position to regard the super-freezing as having preserved the capacity for such activity.

We can understand Unger's notion of survival another way by saying that as long as the brain retains the appropriate structure, physical capacity is sufficient for survival. However Unger distinguishes between different "levels of physical structure."

Appropriate physical structure is enough to ensure the existence of the brain, whether living or not. Appropriate structure that is somewhat finer may be needed to ensure that this brain is composed of living tissue. But much finer structure still is needed, as a matter of fact, for the realization of the psychological capacities by that brain. At least in any of the more ordinary cases, in order for someone to survive, there must be continuity of these very fine structures, not just the grosser ones.⁴⁴

An example of an individual who falls below the required continuity of the finer structures is a permanently comatose patient (to use Unger's term). Although such a patient has a "highly structured living brain" the patient will never again think and experience, will never do anything in a conscious manner. In such a case, the patient will "have lost all of her mental capacities" and thus cannot survive in this state.

We now turn to Unger's distinction between *core* and *distinctive* psychology and the claim that in order to survive it is necessary the brain realize the person's core psychology. Unger presents this distinction in the following way.

Among the mental capacities that it comprises, my dispositional psychology includes those that I share with all other normal human beings, notably my capacity for conscious experience, my capacity to reason in a rudimentary way, and my capacity to form simple intentions. These capabilities are shared, too, with many humans who are markedly below the psychological norms. We may call this group of capacities my *core* psychology, with the understanding that this term may be highly, and perhaps irremediably, vague. Certain other aspects of my psychology I share with some normal humans but do not share with others...Still other aspects distinguish me from all other actual human beings...Let us say that my *distinctive* psychology comprises both of these other aspects of my psychology. So part of my distinctive psychology I share with

⁴⁴Ibid., 105.

some, but not all, other normal human beings, and the rest I share with none.⁴⁵

Why does Unger think that the person will survive if she retains her core psychology but loses her distinctive psychology?

As explanation, Unger provides the following example.

We may imagine that a maniacal surgeon records on tape precise information as to your brain's constitution. Then she tampers with your brain so as to remove all of your distinctive psychology, leaving you only with your core. A week later, she may use the recorded information to restructure your brain, again, thus restoring the distinctive psychology that, for the week, was missing.⁴⁶

Unger then asks us to imagine that the maniacal surgeon informs us that we have the choice of either receiving a slight but considerable amount of pain before the week, or receiving very much greater pain during the week when only core psychology is realized. On the grounds that we want to avoid as much pain as we can, we would choose to receive the earlier lesser pain. This would suggest that we believe that we would survive even without a distinctive psychology. In fact, as Unger says

...you will choose for her to impose the much lesser earlier pain *whether or not*, at the end of the week, the surgeon restores your distinctive psychology. For, as you may deeply believe, in either case, as long as your brain is continuously structured to realize core psychology and, thus, as long as it always is the brain of a person, you will exist.⁴⁷

Unger's conclusion here is that a person's distinctive psychology is "incidental" to their survival.

In the normal course of events, my brain realizes both my core psychology and my distinctive psychology. On the physical approach, it is owing to its realization of my core psychology that my brain is so important for my survival; by contrast, my brain's

⁴⁵Ibid., 68.

⁴⁶Ibid., 105.

⁴⁷Ibid., 106.

realization of my distinctive psychology is completely incidental.⁴⁸

From this two important conclusions follow. First, that those psychological features that are distinctive to me are neither necessary for my survival nor part of what I value in my survival. Second, we can infer from the above that since core psychology is necessary for personhood, you or I cannot survive the loss of personhood.

We are concerned with the plausibility of Unger's claim that my distinctive psychology is incidental to my survival. The success of Unger's view depends on the plausibility of Unger's case of the maniacal surgeon.⁴⁹ According to this case, Unger contends that the person would prefer less pain while possessing only core psychology rather than more pain when the distinctive psychology is restored. From this he concludes that this means that we consider ourselves to have survived though lacking distinctive psychology, and hence distinctive psychology cannot be necessary for survival, and nor is it necessary for self-interested concern.⁵⁰ However, with regard to this notion of self-interested concern, an important question that Unger must address is why I should have any particular concern for the future welfare of the person who possesses only core psychology. For if the maniacal surgeon has removed all of the psychological features that are distinctive to me, what is it about the resulting person that provokes a self-interested concern? The point here is that one might agree with Unger that the person has survived the loss of distinctive psychology,

⁴⁸*Ibid.*, 105.

⁴⁹Cases like this have a rich history. Bernard Williams presents a similar case in "The Self and the Future", *The Philosophical Review*, 79, 161-80, concluding that this shows that we have first-personal concerns for body continuity. Furthermore, as we shall see below, in "When does a life begin?" in Lockwood, M.J., *Moral Dilemmas in Modern Medicine* (Oxford: Oxford University Press, 1985), Michael Lockwood presents a version of the case to show that we have self-interested concerns for our own welfare despite the loss of personhood.

⁵⁰Precisely the same point is made by Bernard Williams in "The Self and the Future."

but disagree with the claim that the person would prefer less pain then than more pain later.

To understand this we need to be clearer just what is meant by "distinctive" psychology. It would appear that by this Unger means that part of our psychology that is unique to each person, in other words, particular or individual thoughts, memories, beliefs (etc.,) that only I have. This would imply, as Unger suggests, that "core" psychology would refer to all those psychological states that are shared by other persons, in the sense that different persons will have similar states. We can ask though (as we have asked before) what makes a particular psychological state *mine*? In other words, what makes it *distinctive*? On the above analysis, a plausible suggestion is that Unger would regard the content of the state as what makes it distinctive.⁵¹ Thus what makes my memory of visiting Venice is that it is me who is "in" the memory -- the memory is about me. These type of psychological features are removed by the maniacal surgeon, and what is left are memories that are impersonal, that have no particular content (perhaps these are quasi-memories). But is it simply the content of the memory (or other psychological state) that makes the memory mine? Perhaps not. Is not the more basic condition that it is I who is presently remembering? Thus what makes the perception of the starling my perception is that I am the one presently doing the perceiving, and not that in some fashion I am part of the content of the experience, or differently, that I identify myself as having that perception. For recall our remarks at the end of the last chapter with regard to the identification of the self. We mentioned there that if we introduce this function of identification then this raises the possibility that misidentification is possible, which is to say that I might perceive the starlings and then realize that it is not I who am doing the perceiving. A scenario that seems very unlikely.

⁵¹Recall the discussion of quasi-memory in the last chapter and Marya Schectman's objection to Parfit.

The point that is being argued for here is that we cannot so clearly separate distinctive from core psychology. My perception of pain is just as individual as my perception of reading Proust, and thus it does not make sense to say that one can be removed while leaving the other. Rather, what is meant when we talk of core psychology is that there is a sufficient reduction or impairment in psychological states that only minimal psychological function remains. In this case it is not clear we would agree with the conclusion Unger draws, for it is not clear that the person would have survived in this state.

However, perhaps we should endorse Unger's conclusion in part. If we regard a patient such as Mr. H who suffers from cerebral arteriosclerosis as someone possessing only core psychology, then *prima facie*, the lack of distinctive psychology would have little effect on the intensity of the pain. For Mr. H is still able to feel pain even if his psychological characteristics. Nevertheless, the success of such an account rests on the plausibility of the distinction between core and distinctive psychology, and for that matter, a greater understanding of what counts as "core psychology". For we may think that it is not the lack of distinctive psychology that has no effect on Mr. H's responsiveness to pain, but that parts of Mr. H's psychological make-up are still integrated in some form or another.⁵²

We now turn to Unger's controversial claim that "once a person always a person."

Unger presents his view in a rather cautious fashion.

...with no great confidence, I finally move to endorse a position that I first floated near the beginning of the book: In that I now am a person, from now on, at any time that I exist, I will be a person. Without great confidence, I do offer this as a necessary truth, but, even then, only if the remark is understood as meaning this: From now on,

⁵²In fact the example of Mr. H appears to support the remarks in the previous paragraphs. Mr. H still has what we might term "distinctive psychological episodes" -- memories of past events, moments of lucidity -- suggesting that the split between core and distinctive psychology is confusing at best.

whenever I will exist, I will be, at the least, a minimal person.⁵³

Unger's describes his notion of a "minimal" person as a person who "may be effected in certain very basic ways, such as being caused severe pain by being made to feel worthless." This might not be what immediately comes to mind when one talks of a "minimal" person, but that is not our concern. Rather, we are interested in why being a person is a necessary condition of survival. So far as I can determine, Unger can be understood as making the following claim. As suggested above, so long as core psychology is preserved a person can survive despite the loss of distinctive psychology. Conversely, the person will not survive if core psychology is not preserved. Thus there will no case where the person survives though lacking the psychological features necessary for personhood (whatever they are). Thus being a person is a necessary condition of survival.

Unger does not state in clear terms the necessary and sufficient conditions for personhood, and by so doing it is difficult to claim that the conditions of survival could be met though the conditions for personhood are not. However, as Ingmar Persson argues, there is a certain asymmetry in Unger's account since Unger admits that the individual existed prior to being a person. As Unger says,

As is at least arguable, there may have been a time when I was not a person, say, when I may have been a one week old fetus newly conceived. At that point, I might not have had even a capacity for quite rudimentary feeling; perhaps, I then had only the disposition to develop into a being with psychological attributes.⁵⁴

However as Persson suggests, this suggests that nonpersons can be transformed into persons, and so we may ask, why can persons not be transformed into nonpersons? Persson presents us with the following possibility.

⁵³Unger, *Identity, Consciousness and Value*, 249.

⁵⁴*Ibid.*, 6.

What if the surgeon had reduced the psychological dispositions to a level just below that constitutive of personhood and had then restored whatever of core and distinctive psychology had been removed. Would the post-operative person now be identical to the preoperative one? Well, it strikes me as plausible to hold that they are identical under *some concept*; so even if they cannot be counted as the same *persons*, because in-between them there is a conscious being that does not qualify as a person, they are nevertheless the same *conscious being or subject*. Similarly, if the surgeon had reduced the mental abilities to nil, the pre-operative and post-operative person could still be the very same living or material being.⁵⁵

Part of Persson's analysis depends on the notion of relative identity that we will turn to shortly. Let us say, for the sake of argument, that we agree with Unger that the individual described above is no longer a person. But notice that I might still be concerned for the welfare of the future nonperson so long as there had been no significant and permanent impairment of brain capacity. However, Persson's suggestion here seems plausible. For the question being asked is why is it necessary that core psychology be preserved as opposed to distinctive psychology? If we recall Unger's concepts of *super-freezing* and *super-thawing* then we may think that the person who was thawed out had permanently lost her distinctive psychology, and thus it is not clear that she has survived the process.

We have looked at two different accounts of survival. These accounts show marked similarities in proposing both physical accounts of survival and for suggesting that the person can survive even when there is only minimal cognitive function. The accounts can be distinguished with regard to the attention that is paid to psychological features. To recall Brennan's account, Brennan claims that there is no one conception of the person, but rather that there are a number of competing conceptions. For Unger, the concept of person is determined by whatever minimal level of cognitive capacity is necessary for survival, so long as the individual had previously been a person.

⁵⁵Ingmar Persson, "The Indeterminacy and Insignificance of Personal Identity," *Inquiry* 35, 271-283.

It is the mark of any such account, perhaps, that one cannot clearly state how much psychological capacity is necessary for survival, nor how much is necessary for personhood. Nor, to be fair, can an answer be provided by supporters of the identity relation. What we have seen is that on both Brennan's and Unger's account you or I cannot survive the loss of personhood. We turn now to the notion of relative identity, and look at an account that is at odds with this conclusion.

Sortal Concepts and Relative Identity

In the *Essay*, Locke makes a distinction between the same man and the same person.

...if it be possible for the same Man to have distinct incommunicable consciousness at different times, it is past doubt the same Man would at different times make different persons...⁵⁶

This seems to be an instance of a position referred to as the thesis of relative identity. According to this thesis, *a* may be the same *F* as *b* but not the same *G* as *b*. That is to say, *a* may be the same Man as *b* but not the same person. We can refer to the descriptive terms or concepts used here ("man" and "person") as *sortal* concepts. Although a critic of the notion of relative identity, Wiggins states the following as consistent with what a supporter of relative identity would claim.

There are, or may be, a whole battery of sortal concepts under which a material particular *a* may fall, and under which it may be counted, individuated, and traced through space and time, *a* may coincide with some specified material particular *b* when individuated under some of these sortal concepts and not coincide with *b*, but be wholly distinct from *b*, when individuated under others.⁵⁷

The concept of relative identity cannot be defended by someone like Wiggins who

⁵⁶John Locke, *An Essay Concerning Human Understanding*, ed. Peter H. Nidditch (Oxford: Clarendon Press, 1975), 342.

⁵⁷David Wiggins, *Identity and Spatio-Temporal Continuity*, (Oxford: Basil Blackwell, 1967), 1.

supports a notion of strict identity, that is, where any property true of *a* is also true of *b*. Above, we have suggested that it might not be a bad thing if we abandoned the concept of strict identity. Should we therefore support this notion of relative identity? There are reasons to think otherwise. The problem with the notion of relative identity so described is that it is too permissive. Consider the sortal terms: graduate student and human being. If we accept Wiggins' contention that the claim $a=b$ must specify "same what", then our conditions of identity will not pick out one and the same entity over time. For the claim $a=b$ will provide a different description if the sortal term is "graduate student" than if the term were "human being."

The difference between these two sortal terms would appear to be a temporal one. Whereas *a* is a human being for the duration of *a*'s existence, this is not true (it is hoped) of the sortal term "graduate student". If identity statements are to be meaningful, and pick out the same entity through the duration of the entity's existence, we need to apply the appropriate sortal term that meets this condition. To this end, Wiggins further distinguishes between

...sortal concepts which present-tensedly apply to an individual *x* at every moment throughout *x*'s existence, e.g. *human being*, and those which do not, e.g. *boy* or *cabinet minister*. It is the former (let us label them, without prejudice *substance-concepts*) which give the privileged and...the most fundamental kind of answer to the question 'what is *x*?'. It is the latter (one might call them *phase-sortals*) which, if we are not careful about tenses, give a false impression that *a* can be the same *f* as *b* but not the same *g* as *b*.⁵⁸

Thus in the above example, "human being" is a substance sortal and "graduate student" a phase-sortal. In this fashion, Wiggins hopes to defeat the claims of relative identity, for as long as we pick out the appropriate sortal term we can show that other conflicting identity claims only describe part of the entity's existence, and thus should not be regarded as genuine

⁵⁸Ibid., 1.

identity claims.

This leads us to the following important question, "Is *person* a substance-concept or a phase-sortal?" Unfortunately, the answer is not clear. If we use "applying to an individual *x* at every moment throughout *x*'s existence" as a necessary and sufficient condition of a substance sortal, then one could conclude that person is either a substance-concept or a phase-sortal. For on the one hand, one could claim that you and I are persons for the duration of our existence. This, of course, is a controversial claim (and one that will be discussed in detail further on in this chapter). Support for this claim can be evidenced if we recall Gillett's concept of person. For Gillett, the concept of person can be understood as composed of the following "themes": personal identity, quality of life, and responsibility for self, all of which are dependant on a notion of consciousness. If the individual was to permanently lose each of these components one could claim that the individual has ceased to exist, since this would be an individual with such a reduced level of consciousness and cognitive capacity (for example, a patient in a PVS) that it is difficult to claim that the individual continues to exist.

On the other hand, there are problems with categorizing person as a substance-concept, and this suggests that it is better described as a phase-sortal. Apart from possibly the type of account of personhood put forward by John Noonan that was discussed in Chapter Two, and that attempts to ascribe personhood to all living human beings, it is more general to regard the class of persons as composed of only some members of the class of human beings (and perhaps members from other classes). There are pragmatic reasons for this position. Recall that one function of the concept of person is to enable us to make moral judgements -- we accord persons greater intrinsic moral worth than nonpersons. However, if it should be the case that we are persons throughout our existence then,

arguably, we must conclude that either prior to, and subsequent to the possession of the properties necessary and sufficient for personhood we do not exist (although there is a living human being that does exist before and after these times) or that all humans are persons. The second option is less attractive than the first, for in this case "person" could no longer adequately function as a means of distinguishing members of this class from the class of human beings. The first option is problematic in that it re-presents two of the key questions of this chapter, namely "Am I essentially a person?" and "Can I survive the loss of personhood?" If person is a substance-concept, then we would have to answer both questions in the negative.

The conclusion to be drawn here is that we cannot conclusively determine whether "person" is more appropriately regarded as a substance-concept or a phase-sortal until we have looked at these two questions. However, we can see the relevance of the matter of relative identity to these questions. To see this, we need to recall Locke's distinction between "same man" and "same person." If we consider the case of Mr. H from the previous chapter (the patient who suffers from dementia as a result of cerebral arteriosclerosis) we can see that although it is a controversial question whether the now incompetent Mr. H is the same person as the prior competent individual, on the grounds that the former does not meet the conditions of personhood, it is uncontroversial that he is the same man. This is exactly the type of reasoning employed by those who claim that severe and permanent cognitive impairment leads to the loss of personhood.

Furthermore, there is a connection here between relative identity and advance directives. To see this we can recall Buchanan and Brock's defence of the moral authority of advance directives (and more generally, the subjective approach to decision-making for incompetent patients). Buchanan and Brock contend that the authority of advance directives

is upheld on the grounds that the incompetent patient is not a person and hence not a different person. But, if this is the case, how does the advance directive apply? Let us say that we agree with the claim that Mr. H is no longer a person, and that he has an advance directive. If we understand the function of the advance directive as a means by which a competent person can make treatment decisions to cover a later period of incompetence, then, if Mr. H is no longer a person, we can conclude that the individual's self-interested concerns, and the functioning of the advance directive, are not dependant on that individual remaining a person. However, there must be some identity between the prior competent person and the later incompetent individual, for otherwise the directive would have no authority, or differently, that particular prior competent person would have no more authority to make treatment decisions for the later incompetent individual than any other prior competent person. This is to say although the incompetent patient is not the same person, he or she is the same individual.⁵⁹

And there is a further consideration. If we grant that the prior competent person has self-interested concerns in the welfare of the later incompetent patient and the incompetent patient is not a person, this would imply that one does not have to be a person to be the subject of self-interested concerns. This runs counter to those accounts of the concept of person, for example Tooley's account, which regard being the subject of self-interested concern to be a sufficient condition of personhood. Thus if Tooley wants to defend the view that self-interested concerns are connected to the individual's status *qua* person (to recall, he

⁵⁹There is another option, and this is to claim, along with Buchanan and Brock, that the competent person has a "right of disposal" over the later incompetent patient. But this option runs the risk of appearing question-begging, for we would need to show how this right is established. In other words, why does the prior competent person have this right as opposed to any other competent person? Secondly, if we imagine cases where the patient's cognitive impairment is less severe, then we might be wary of claiming that the competent has the right of disposal of the incompetent patient.

maintains that to talk of *self-interest* presupposes that the person recognizes herself as a subject of non-momentary interests which he claims tentatively is necessary and sufficient for personhood) then he must conclude either that the incompetent patient is a person, or that advance directives do not count as evidence of the competent person's self-interested concern for the later period of incompetence.

The first of these options is problematic since there will be cases of incompetence, such as PVS, where if these were to be regarded as cases where the patient is still a person, it would be difficult to come up with a case where the incompetent patient was *not* a person. Thus if the term "person" is to be used to pick out morally relevant criteria, we must accept that it must, as a minimum, exclude those patients who have permanently and irreversibly lost the capacity for consciousness.

The second option is more difficult, for it is not clear just what a person might consider as important when she executes an advance directive to cover later periods of incompetence. To recall the case of Mr. H, can we conclude that when he issued the directive he did so according to the judgement that being in a demented state was not a condition in which *he* wanted to go on living? In other words, prior to the onset of the dementia, did he execute the directive because he considered that he would survive the onset of dementia? Presumably this is the case, for otherwise there would be no need to execute the directive.

A further point is that we would have to explain what it is about being an incompetent nonperson that warrants self-interested concern. To show this we have to show that whatever was relevant and necessary for such concern in the case of the prior competent person is still present in the case of the incompetent individual. That is to say, if the incompetent patient can be the subject of self-interested concern this implies that the prior

competent person considers himself to have survived the onset of incompetence, and that this is a state in which he no longer considers to be state worth living in, which is to say that *his* interests would be harmed by the continuation of life-sustaining treatment.

This is to grant perhaps more than we would want, for notice that it would now be the case that (i) there is an identity relation that connects the competent and incompetent, for we could now talk of the same individual, and (ii) the incompetent individual has interests that *prima facie* restrict the actions of others. For if one can have a self-interested concern for the incompetent individual, this suggests that the incompetent individual can be the possessor of interests.

If the above two conditions are met then we would have to seriously question the defence of the moral authority of advance directives supported by Buchanan and Brock. For the claim that the authority of advance directives is upheld on the grounds that incompetent patient is not a person becomes less relevant if it need not be the case that the appropriate *relatum* is "person." For it would now appear to be the case that the identity conditions can be met, and the later individual have interests, even if the later individual is not a person. Moreover, this raises the possibility that there might be cases where the identity conditions are not met, and thus there would be a different individual with interests that might, theoretically, be at odds with those of the prior competent person.

Relative Identity and the Loss of Personhood

We turn now to the question, "Can I survive the loss of personhood?" The importance of this question is apparent from the discussion above. If we can plausibly argue that the individual can survive the loss of personhood, and have self-interested concerns for the later nonperson, then this would undermine the importance of the concepts of person and

personal identity with regard to the moral authority of advance directives. To see this we need only recall Buchanan and Brock's claim that the moral authority of advance directives can be upheld on the grounds that since the incompetent patient is not a person, we need not be concerned whether he or she is a different person.

If we are to defend Buchanan and Brock's argument, and more generally, the "primacy" of the concept of person, we need to show that the individual cannot survive the loss of personhood in the way described. That is to say, that the resultant individual cannot be claimed to possess the appropriate features necessary and sufficient for identity (here understood as picking out the individual and unique features of the individual) nor necessary and sufficient for moral relevance (here understood as the features necessary for moral standing).

Lockwood

One account that endorses a notion of relative identity, and regards person as a phase-sortal is the account put forward by Michael Lockwood.⁶⁰ Lockwood argues that you or I can survive the loss of personhood. He distinguishes between the concepts *human being* and *person*.

"Person"...is not a biological concept at all. A person is a being that is conscious, in the sense of having the capacity for conscious thought and experiences, but not only that: it must have the capacity for reflective consciousness and self-consciousness. It must have, or at any rate have the ability to acquire, a concept of itself, as a being with a past and a future. Mere sentience is not enough to qualify a being as a person. But a person, in this sense, need not be human.⁶¹

Lockwood then asks the following question, "Could I cease to be a person without ceasing

⁶⁰Michael J. Lockwood, "When does a life begin?" in *Moral Dilemmas in Modern Medicine*, ed. Michael J. Lockwood (Oxford: Oxford University Press, 1985), 10.

⁶¹Ibid., 10.

to be?" Lockwood answers in the affirmative, on the grounds that personhood is not a *temporally essential* attribute: it is not an attribute that an entity must have throughout the duration of its existence. In other words, according to Lockwood, "person" is a phase-sortal.

In defence of these claims, Lockwood asks us to consider the following thought-experiment.

Suppose that you knew that you were going to suffer from some terrible disease, which would slowly extinguish your mental capacity for reflective consciousness or self-consciousness....But imagine that this disease still left the organism capable of sentience: it would still be aware, still be capable of experiencing sounds, colors, pleasure and pain, and so forth. Only higher cognitive functions would have gone. Now suppose this being were to be subjected to the most excruciating pain imaginable for some extended period. If you knew that it was going to be you that suffered the disease, and that it would be the brain you now possess, albeit pathetically reduced in cognitive capacity, whose pain centers were going to be stimulated, what would your attitude be towards the pain? Would you consider that it was going to happen to you? Would you deem it rational to fear this pain in a self-interested way? Or would you think of it as something that was going to happen to someone or something that was not, after all, you....If you think as I do, that it would be rational to fear this pain in a self-interested way, then you will be forced to conclude that being a person is not a temporally essential attribute but merely a phase sortal. I could cease to be a person and yet still exist....⁶²

If Lockwood is correct, then it would seem possible for you or I to survive the level of cognitive impairment that robbed us of the attributes of personhood. Furthermore, Lockwood is suggesting that not only would we survive, but that we would be concerned in a self-interested way for the welfare of the later nonperson, that is to say, I would be concerned for the welfare of the nonperson in a manner consistent with my concern for my present welfare.

It is important to distinguish the following claims. First, that the individual would survive the loss of personhood, and second, that the individual would be concerned in a self-interested way for the welfare of the nonperson. For one can imagine the situation where

⁶²*Ibid.*, 12-13.

it would be plausible to say that you or I survive the loss of personhood but, nevertheless, have no interest in the welfare of the future nonperson. A possible example of this would be if you were in a PVS. And the difference between this situation and the one that Lockwood describes is the level of cognitive ability of the surviving nonperson. In Lockwood's example the surviving nonperson is sentient, whereas this is not the case with PVS.

Furthermore, there is a third possibility. One could agree with Buchanan and Brock and contend that interests can survive even though the individual herself does not survive.⁶³ Thus Buchanan and Brock argue that since interests can survive death (otherwise we would not be concerned with probate) so they can survive the permanent and irreversible loss of consciousness. Thus the person might be "self-interested" in the fate of the later individual, and there might be such a thing as the survival of the person's interests even though the person herself does not regard herself as having survived the loss of consciousness (or in any way identical with the incompetent patient).

This third possibility is troublesome for it is controversial whether it makes sense to say that interests can survive even though the person (in any sense) does not. Although it might be thought that a person's financial interests survive his/her death, it is questionable whether what we are witnessing is simply a legal process for disposing of assets rather than any robust claim that interests can survive "outside of" the person to whom they belong.

If Lockwood is correct in his contention that "person" is a phased sortal then as Lockwood says,

We need to find a term for whatever it is that you and I are essentially, what we can neither become nor cease to be, without ceasing to exist. I use the term *human being*

⁶³Allen Buchanan and Dan W. Brock, *Deciding For Others: The Ethics of Surrogate Decision Making*, (Cambridge: Cambridge University Press, 1989), 165-169.

to fill this slot.⁶⁴

Why does Lockwood think that we are essentially human beings? There is a negative and a positive answer. The negative answer is that neither of the two alternatives are plausible. If the above argument is correct then it seems difficult to maintain that we are essentially persons, given that it seems that you or I could survive the loss of personhood. But neither can it be maintained that we are essentially *human organisms*, for the (our) human organism can continue to exist long after we think that you or I have ceased to exist. Thus Lockwood distinguishes between a human being and a human organism. What this distinction involves becomes clear from his positive argument.

Lockwood's positive argument rests on his understanding of what he takes to be the underlying feature of our continued existence. This, in turn, relates to the question as to when a human being comes into existence. Lockwood argues that if one could answer the question "What has to happen for a human being to have come into existence in the first place?" then one could also answer the question "What is it that makes someone at one time the very same human being as someone at another time?"

Briefly, if I know what relationship is constitutive of my identity through time, then I know that I come into existence at the moment when there first existed something that stands in that relationship to me...Our problem thus reduces to that of determining what that relationship might be.⁶⁵

To talk of "the relation (that) is constitutive of my identity through time" is to talk of the matter of "so-called personal identity," for is this not a different way of talking about the

⁶⁴Lockwood, 14.

⁶⁵Ibid., 14.

"unity relation of person?"⁶⁶ Here I am assuming that one can make parallel remarks for when there last existed something that stands in that relationship to me. In other words, an answer to the question "What is it that makes someone at one time the very same human being as someone at another time?" would also provide us with an answer to the question "What has to happen for a human being to cease to exist?"

To support his claim that you or I are essentially human beings (that is to say, we are not essentially persons) Lockwood further distinguishes between the questions "What do we mean when we speak of the identity through time of a human being?" and "What does that identity actually consist in?" These are two different questions, says Lockwood, and an answer to the first would still leave the second unresolved. The difference between the questions is the difference between the *evidence* for the identity of *a* and *b* and what that identity is *constituted* by.⁶⁷ As we discussed in the previous chapter, we might describe the diachronic identity of the person in terms of her psychological features, and this might be the customary evidence that is presented as suggesting personal identity. Alternatively, we may regard physical features as evidence of personal identity. But if we recall Swinburne's objection to reductionist accounts of personal identity, it is Swinburne's claim that such evidence is "fallible" in that it does not provide us with the necessary and sufficient conditions of personal identity, but merely rebuttable evidence for that identity.

⁶⁶Unless we are careful, we are going to have a problem with sortal terms and identity relations. It is Lockwood's contention that the relevant identity relation is *human identity*, though he speaks of the relation of "so-called personal identity". That is to say, Lockwood thinks that we can replace the relation of personal identity with that of human identity and still capture what is important about being a person -- individuality and continuity over time. Following Lockwood's usage, I will use the phrase "so-called personal identity" to refer to a relation that could be described by some as personal identity and by Lockwood as human identity. By using this term I am not therefore claiming that the conditions for personhood have been met of that person is the appropriate *relatum*.

⁶⁷Recall that this distinction was discussed in Chapter Three.

As an example of this difference between these two questions ("What do we mean when we speak of the identity through time of a human being?" and "What does that identity actually consist in?") Lockwood considers the question, "What is gold?". We could answer the question in terms of its yellowness, malleability and so forth, but according to Lockwood, this would not be a satisfactory definition of gold, since under different conditions these features of gold would not be present. A better definition would be in terms of its atomic weight, that is to say, in terms of what Lockwood calls the "underlying continuity". The same analysis would go for the question, "What are you or I?". And just as it would be a mistake to attempt to define gold in terms of its yellowness and malleability, so it would be a mistake to define human identity ("so-called personal identity") in terms of memory, continuity of attitudes and interests (etc.). These could be used as *evidence* for the conclusion that the earlier and later persons are the same person, but, since like the properties of gold, they are changing, we cannot understand the constitution of human identity in terms of these attributes. Rather, we should search for the underlying continuity of these attributes or qualities.

What is this underlying continuity? Lockwood writes

The view that I favor, and that I think is favored by the available scientific evidence, is that what underlies the discernible continuities of memory and personality is a continuity of physical organization within some part or parts of a living human brain persisting through time.⁶⁸

Thus Lockwood's answer to the question "What is it that makes someone at one time the very same human being as someone at another time?" is continuity of brain. "A human being cannot be me unless he has my brain, or at any rate some crucial part of it". Furthermore, an answer to this question would also provide us with an answer to the question

⁶⁸Lockwood, 22.

"What has to happen for a human being to have come into existence in the first place?"

Just as I shall live only as long as the relevant part of my brain remains essentially intact, so I came into existence only when the appropriate part or parts of my brain came into existence, or more precisely, reached the appropriate stage of development to sustain my identity as a human being, with the capacity for consciousness. When I came into existence is a matter of how far back the relevant neurophysiological continuity can be traced.⁶⁹

This, in turn, provides us with Lockwood's answer to why it is that we are essentially human beings: we are essentially human beings because the underlying and defining continuity is the presence (and continuity of) the brain, and in the absence (or cessation of function) of the brain, you or I would cease to exist.

If Lockwood is correct, then he has shown us that person is a phase-sortal and not a temporally essential attribute, and that self-interested concerns can survive the loss of personhood. But we need to be careful here as to just what Lockwood has shown. For we cannot presume that Lockwood's conclusion here implies that *you* or *I* have survived the loss of personhood. As we saw above, the presence of a self-interested concern for the welfare of a future entity need not imply that the necessary and sufficient conditions for personal identity ("so-called" or otherwise) are met. For it has been claimed that since interests can survive death, so they can survive the loss of personhood, and so one can speak of a person's interests or self-interested concerns surviving the loss of personhood without having to claim any identity between the earlier and later individuals.

A further point is that although we do talk of "same man", "same human being," "same person" or "same human organism" this variety of terms does not commit us to relative identity. To recall, the strict identity thesis is that $a=b$ if any property true of a at $t1$ is true of b at $t1$. We have not violated this notion of strict identity if b should have a different

⁶⁹Ibid., 23.

property at t_2 . As Wiggins says above, "...if we are not careful about tenses [we can] give a false impression that a can be the same f as b but not the same g as b ." Thus Lockwood is right to focus on the question of what you or I are *essentially*. For clearly whatever that property is, it must be held by b at t_1 and t_2 if it is an essential property of a and it is claimed that $a=b$.

The important question is to determine whether Lockwood is correct in his contention that we are not essentially persons and can, therefore, survive the loss of personhood. There are a number of points to consider here. First, there is a problem regarding the following claim:

The question of when a human being comes into existence is really the same question as that of what constitutes the identity of a human being over time: the so-called question of personal identity.⁷⁰

According to Lockwood, the only plausible candidate to be the defining and underlying property of humanhood for a at t_1 is that a have a brain supporting the capacity for consciousness at t_1 . Thus the answer to the question, "When does a human being comes into existence?" is when the brain reaches the appropriate level of neurological development. (Similarly, the human being ceases to exist when his brain function falls below this level.) Furthermore we can ask what is it to say that a at t_1 and b at t_2 are the *same* human being? According to Lockwood, this is to say that the individual has the same brain with the capacity for consciousness.

However, the important question is whether our answer to the first of these questions does provide us with an answer to the second. For let us say that we agree with the plausible suggestion that the human being came into being when the brain reached the appropriate level of neurological development. Does this claim present us with the necessary and

⁷⁰Lockwood, 14.

sufficient conditions for human identity? It is clear that at the time when the human being comes into or goes out of existence the conditions for identity are destroyed. For one cannot claim that *a* at *t1* can be the same person as *b* at *t2* if *a* ceased to exist in the interim. But this is trivially true, for it is clear that we cannot make identity statements if one of the *relata* is missing. One cannot say the car at *t1* is the same car as the car at *t3* if the car ceased to exist at *t2*.

Furthermore it should be pointed out that there is no claim here as regards the essential properties of the entity involved, for all that is being claimed is that *a* cannot be the same *F* as *b* if *a* is not an *F*. Now it might be the case that *F* is an essential property of *a*, but the truth of the statement does not depend on this fact.

Thus what Lockwood is claiming cannot be that in order *a* to be the same *F* as *b*, *a* must be an *F*. Rather, he can be regarded as claiming that whatever the conditions are that determine that boundaries of the individual's existence, so these must also determine the identity conditions. But clearly this is not true. Let us say for the sake of argument that I came into existence when my brain first starts to work, and I will cease to exist when my brain stops working. Can we conclude from this that *a* at *t1* is the same human as *b* at *t2* if the same brain is working? I do not think that this can be concluded from the fact that I would cease to exist if my brain ceased functioning.

We can see why if we recall the case of Mr. H. Imagine that in a few years neurologists could save a person from the onset of cerebral arteriosclerosis or Alzheimer's dementia by somehow (don't ask me how) transferring the person's psychological make-up to a new brain. If this were possible it is not obvious that Mr. H would cease to exist and the "Mr. H" in the new brain would simply be a replica. Nor is having the same brain with the capacity for consciousness sufficient for human ("personal") identity. If we performed a *brain*

zap and destroyed or erased *all* of the human's psychological features, then although the entity would still have an appropriately functioning brain it is not clear that we have retained anything that is customarily thought to be relevant to human ("personal") identity.

Thus it is not clear that Lockwood is making an identity claim of any type at all. What is being claimed is that you or I would cease to exist without an appropriately functioning brain. And this is plausible in just the same way as the claim that you or I would cease to exist without an appropriately functioning heart. But not only would it be mistaken to claim that I need the same heart to continue to exist (for this is false) so it would equally be mistaken to conclude that I need the same brain (there is no logical reason to rule out my continuing to exist with a different brain). Moreover, just as it would be unwarranted to claim from the fact that a functioning heart is necessary for my continued existence so I am *essentially* a heart, so it would be unwarranted from the fact that a functioning brain is necessary for my continued existence so I am *essentially* a brain. And Lockwood appears to endorse this latter claim. This claim might be true, however the argument does not warrant the conclusion.

A further question is whether Lockwood is correct in concluding that the identity of the individual should be equated with the survival of the human being. By this we mean whether I think that *I* would have survived as a human being. To recall, Lockwood's argument in support of this claim is that it would be rational for me to suppose I would feel the pain even though I had lost the relevant psychological capacities usually associated with personhood. As we are well aware, we saw just the same type of example in our discussion of Unger, and the same points that were made at that time are worth repeating.

First, Lockwood suggests that a clear division can be made between the psychological capacities necessary for personhood and those for humanhood (Unger, to recall argued that

a clear division could be made between core and distinctive psychology). But it is not clear that this is the case, and hence the plausibility of the example is questionable.

Second, and more importantly, does Lockwood show that you or I can survive the loss of personhood? It is fair to say he does, but his success here is due to a particular definition of personhood. To understand this it is worth recalling that both Brennan and Unger deny that you or I can survive the loss of personhood. The reason being that both Brennan and Unger support, in some contexts, the view that personhood is preserved when the individual possesses only reduced cognitive ability. Brennan, to recall, speaks of the "administrative" conception of the person that can be applied to those individuals whose mental life is confused and who have a reduced level of conscious ability. Unger speaks of a minimal conception of the person as one, perhaps, who possesses only core psychology (whatever that might mean). Thus if we think of Mr. H, he may meet the conditions of personhood on some of these accounts but not on others, and as such might be regarded as a case of personal or human survival or personal or human identity.

This is to say that what on one account would count as an example of human identity (for example, Lockwood's) might on another account (Unger or Brennan) count as an example of personal survival, while on another (Tooley) might count as a case where the person has not survived. Our problem here is due in considerable part to the concept of person. As Brennan suggests, this is a vague concept and one that has different meanings in different contexts. Given that we have no clear and exhaustive concept of person agreeable to all, it is difficult to find unanimity with regard to the conditions of personal identity or survival.

A further point is that there simply is not clear answer to the question whether Mr. H is identical with or has survived as the later incompetent patient. If we adopted a notion

of *strict identity* then we could conclude that Mr. H is not the same person when incompetent as he was previously. But for the reasons stated above, the notion of strict identity is not the notion of identity used in our discussions of personal identity. We must, therefore, make use of a looser sense of identity or talk instead in terms of survival. But even if we talk in terms of survival, there will be cases where we cannot say for certain whether the individual has survived. Perhaps our best strategy is the one that we have seen used here -- the use of "avoidance of future pain tests." By this method we can gain some insight as to whether we think that we would have survived in some later reduced state. However, as I have suggested, these tests present a somewhat skewed account of cognitive impairment, and as such should be treated with caution.

Our conclusion here is that although Lockwood contends that you and I can survive the loss of personhood, this should not undermine the authority of the subjective approach to decision-making for incompetents. Lockwood's conclusion does raise questions as to the relevance of the claim of Buchanan and Brock that since the incompetent patient is not a person, he or she cannot be the same or a different person. For Lockwood raises the possibility that being a person is neither necessary for the possibility of self-interested concern nor for a relevant identity relation. The consequence of this is that on certain concepts of "person" and of "identity" we can regard the incompetent patient to have survived although in a considerably reduced cognitive state, and thus, perhaps, not as a person. So long as the patient maintains (to some degree which cannot be determined) the capacity for conscious activity, we can regard the patient as having survived the incompetence.

We have thus rejected those accounts that suggest that in order to survive in a way that is relevant to self-interested concern, the individual must retain a high level of self-consciousness, if by this we mean that the individual recognizes herself as a "subject of non-

momentary interests" (to use Tooley's phrase). For it seems defensible to claim that the person can survive although lacking this self-reflexive capacity.

It should be added, however, to talk of the conditions of survival is not to talk of the conditions of identity or of value. Although I might prefer to suffer less pain while incompetent rather than more pain when my competence is restored, it should not be concluded that the incompetent state describes the conditions of personal identity. For although I might consider myself to have survived in that incompetent state, I might consider such a state to lack most of the features that uniquely describe who I am.

Furthermore, one should not conclude that from the fact that I would avoid future pain in this state, that this is a state that has value. To better assess this fact we would have to compare this state with other states -- would I prefer to suffer pain at some future time while competent or incompetent? It is Unger's contention that all cases of survival are equally cases of survival, but we have presented reasons to challenge this. I might regard survival-as-incompetent as being a case of less survival than survival-as-competent on the grounds that the former state lacks much of what is valued about continuing to exist.

Nevertheless, by talking in terms of survival rather than identity we can provide an account that defends the moral authority of advance directives and the subjective approach to decision-making while at the same time doing justice to our notions of self-interest and what we value.

We are now in a position to draw some final conclusions. Our investigations have shown that the subjective approach to decision-making for incompetent patients can be defended. Particularly, we have shown that one can present a defence against the identity question on the basis of either the notion of relative identity or that of survival. In this regard, the moral authority of advance directives can be upheld, for one can claim either that

you or I can self-interested concerns despite the loss of personhood, or claim that the concept of person derives its meaning from its instrumental value, and as such, can be applied with different meanings in different contexts.

CONCLUSION

We have now reached the end of this work and it is time to draw some final conclusions. We began this enterprise by examining in the first chapter the notions of competence and incompetence, the rights of incompetent patients, and three approaches to decision-making. It was in this chapter that we first presented the "identity question" and examined Buchanan and Brock's attempted resolution to this question by denying that the incompetent patient is a person.

Our conclusions in this chapter were various. In our discussion of rights we concluded that a necessary condition of having a right to x is that the individual can be benefitted or harmed by the presence or loss of x . We rejected the conclusion drawn by White that only persons can be the possessors of rights since only persons can claim or recognize a right, and endorsed instead the more modest approach of Feinberg who grants rights to individuals even though others may have to claim or exercise the right on their behalf. In our discussion of Buchanan and Brock's attempted resolution of the identity question we argued that the success of their resolution cannot be determined until we have examined the conditions of personhood.

In the second chapter we discussed three different accounts of the concept of person. Our discussion of Noonan's account led to the conclusion that it is problematic to contend that moral standing can be granted on the possession of potential properties. The distinction between the *moral* and *metaphysical* senses of persons helped in showing that arguments that base moral standing on potential properties cannot adequately explain why the potential and not the past possession of these properties is morally relevant.

The examination of the second of the three accounts by Michael Tooley revealed that one possible necessary and sufficient condition of personhood is being the subject of non-momentary interests. That is to say, the individual is aware that she has a past, present and future, and has interests that overlap these times. However, our discussion of Tooley's account revealed that it was difficult to determine just what we mean by a *subject* of non-momentary interests in the sense of whether the individual is aware of herself as a continuing subject of interests. For this reason, Tooley was unable to conclude that having the concept of oneself as a continuing subject of mental substance is one of the "non-potential properties" that are necessary for personhood and the right to continued existence.

Gillett's account identified three *marks* of the concept of person, *quality of life*, *responsibility for self*, and *personal identity*. Our conclusion here was that these three marks are sufficient conditions of personhood, and all rely on the more primary condition of consciousness. The identification of personal identity as a possible condition of personhood is very important, for it suggests that the interrelation of these two concepts is rather more complex than might at first be suspected. Furthermore, it raises the issue that one feature of the moral importance of persons is the identification of a specific individual. That is to say, it is by unifying interests and desires into the desires of a particular individual that these interests and desires become morally relevant.

In the third chapter we turned our attention to the question of personal identity. In this chapter we discussed Swinburne's non-reductionist account, and the reductionist accounts of Locke, Parfit and Shoemaker. Our main focus in this chapter was to attempt to explain the synchronic and diachronic unity of conscious mental states. The purpose here being to attempt to avoid the twin problems of the implausibility of non-reductionism and the inadequacy of the Bundle Theory. Swinburne's non-reductionism was rejected on the grounds

that it would be unable to distinguish the type of cases that are the concern of this work from cases where *prima facie* the conditions of personal identity are met. For since on the non-reductionist account personal identity is not constituted by psychological or physical continuity, so cases such as that of Mr. H who suffers from cerebral arteriosclerosis would present no challenge.

Parfit's Bundle Theory was rejected on the grounds that it is unable to explain both how different conscious mental states belong to one individual, or why such states should constitute the identity of such an individual. For by allowing the diverse conscious mental states to be R-related by "any kind of cause," and the use of quasi-memory, it is difficult to see how any particular individual could be uniquely identified (whether or not identity matters in survival).

We tried to describe an account of psychological continuity that avoided these other problems, and identified Shoemaker's functionalist account as the most attractive. The advantage of such account, in conjunction with other aspects of Shoemaker's philosophy of mind, is the recognition that first-person memory claims are not constitutive of personal identity. In this fashion one could claim that the diverse conscious mental states of an incompetent patient could be unified even though such a patient does not have a coherent set of memories. Furthermore, in supporting Shoemaker's account, we endorsed his conclusion that in the normal cause of events, physical continuity plays an important part in our notion of personal identity.

In the final chapter we have looked at the matter of survival. Our purpose in this chapter was to try and present an account that avoids having to determine exactly what type of entity the incompetent patient is. In this fashion we can side-step discussions of whether the incompetent patient is a person or not, for the survival relation does not demand that the

entity continue to be described by the same substance sortal concept in order to survive. In discussing the accounts of Unger and Brennan our aim was to investigate the conditions for survival, and whether you or I could survive the loss of personhood. For his part, Unger answers in the negative. However, we attempted to argue that Unger might be mistaken in this conclusion since it rests on the questionable distinction between *core* and *dispositional* psychology. According to Brennan, there is no one conception of person. Rather, "person" may be regarded as an "instrumental" term having different meanings given different contexts.

We then turned to a discussion of Michael Lockwood's claim that we are not essentially persons, and that you or I could have self-interested concerns in the future individual despite the loss of personhood. According to Lockwood, an essential feature of our continued existence is continuity of brain, and if the brain continues to function at the appropriate level (whatever that might be) then even though the higher cognitive abilities that are necessary for personhood are lost, you or I would still be concerned for the welfare of the future individual.

Despite the differences amongst Brennan, Lockwood, and Unger there is considerable unanimity. All three endorse some notion of continuing brain functioning as necessary for survival, and so support a view of survival and identity that is physicalist in nature. Furthermore the differences amongst the accounts with regard to the conditions of survival, and in answer to the question whether you or I could survive the loss of personhood, are differences due to the vagueness of the concepts of person and survival. As we have seen in this work, there is no general agreement on the conditions of personhood apart from perhaps the capacity for consciousness (though John Noonan would not agree even with this condition). Particularly in this chapter, the differences amongst Brennan, Lockwood, and Unger with regard to the conditions of survival have much to do with differences in their

respective understanding of the conditions of personhood. For example, to be a person in Brennan's "administrative" conception of person requires less neurological functioning than Lockwood's nonperson survivor. And Unger's argument for claiming that you or I are essentially persons is that the level of psychological function necessary for *core psychology* is presumably so low that the loss of this core psychology would make survival impossible.

We can put the matter another way by saying that we can describe one and the same case in a number of ways. If we recall the case of Mr. H, then this could be regarded as a case of nonpersonal survival, personal survival, nonpersonal identity or personal identity. There is no fact of the matter here. What I have tried to show in this work is that one can present a plausible defence of the view that the incompetent patient is a survivor of is identical with the prior competent individual. I have also tried to show that the concept of person is at best only moderately useful in explaining either our moral concerns or our metaphysical ones. For it does not appear necessary that in order to survive, or in order to be identical, it is necessary that you or I continue to be persons.

SELECTED BIBLIOGRAPHY

American Neurological Association. "Persistent Vegetative State: Report of the American Neurological Association Committee on Ethical Affairs." *Annals of Neurology* vol. 33 no. 4 (1993)

Andrews, Kenneth. "Recovery of Patients After Four Months or More in the Persistent Vegetative State." *British Medical Journal*. vol. 306 1597-1600.

Annas, George. "Law and the Life Sciences: In Re Quinlan: Legal Comfort for Doctors." *Hastings Center Report*. vol. 6 no. 3 (1976): 29-30.

Baillie, James. *Problems in Personal Identity*. New York: Paragon House, 1993.

Beauchamp, Tom L., and James F. Childress. *Principles of Biomedical Ethics*. 4th ed. New York: Oxford University Press, 1994.

Bechtel, W., and Adele Abrahamsen. *Connectionism and the Mind: An Introduction to Parallel Processing in Networks*. Cambridge: Basil Blackwell, 1991.

Binstock, Robert .H., Stephen G. Post, and Peter J. Whitehouse. *Dementia and Aging: Ethics, Values and Policy Choices*. Baltimore: The Johns Hopkins University Press, 1992.

Brazier, Margaret. *Medicine, Patients and the Law*. London: Pelican, 1987.

Brazier, Margaret A. and Mary Lobjoit. ed. *Protecting the Vulnerable: Autonomy and Consent in Health Care*. London: Routledge, 1991.

Brennan, Andrew. *Conditions of Identity: A Study in Identity and Survival*. Oxford: Clarendon Press, 1988.

Buchanan, Allen. "Advance Directives and the Personal Identity Problem." *Journal of Philosophy and Public Affairs* 17 (1988): 277-302.

Buchanan, Allen and Brock Dan W. *Deciding For Others: The Ethics of Surrogate Decision-Making*. Cambridge: Cambridge University Press, 1989.

Brody, Baruch. *Identity and Essence*. Princeton: Princeton University Press, 1980.

_____. *Life and Death Decision Making*. New York: Oxford University Press, 1988.

- _____. "Ethical Questions Raised by the Persistent Vegetative Patient." *Hastings Center Report*. vol. 18 no. 1 (February/March 1988): 33-37.
- Butler, Joseph. "Of Personal Identity." In *The Analogy of Religion* (1736). Reprinted in *Personal Identity*. ed. John Perry, 99-105. Berkeley: University of California Press, 1975.
- Calvin, William H. *The Cerebral Symphony: Seashore Reflections of the Structure of Consciousness*. New York: Bantam Books, 1990.
- Chisholm, Roderick. *Person and Object*. London: George Allen & Unwin, 1976.
- Carruthers, Peter. "Brute Experience." *Journal of Philosophy* 86 (1989): 258-269.
- Carruthers, Peter. *Introducing Persons: Theories and Arguments in the Philosophy of Mind*. London: Routledge, 1989.
- Cassel, Ernest. "Recognizing Suffering." *Hastings Center Report*. vol. 21 no. 3 (May/June 1991): 24-31.
- Clark, Andy. *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing*. Cambridge: The MIT Press, A Bradford Book, 1990.
- Churchland, Paul M. "Eliminative Materialism and the Propositional Attitudes." *The Journal of Philosophy* LXXVII 2 (February 1981): 67-90.
- _____. *Matter and Consciousness*. rev. ed. Cambridge: The MIT Press, A Bradford Book, 1988.
- Cranford, Ronald. "The Persistent Vegetative State: The Medical Reality (Getting the Facts Straight)." *Hastings Center Report*. vol. 18 no. 1 (February/March 1988): 27-32.
- Dennett, Daniel C. "True Believers: The Intentional Strategy and Why it Works." In *Scientific Explanations: Papers based on Herbert Spencer Lectures Given in the University of Oxford*. ed. A.F. Heath, 53-75. Oxford: Oxford University Press, 1975. Reprinted in *The Nature of Mind*. ed. David Rosenthal, 339-350. Oxford: Oxford University Press, 1991.
- _____. "Conditions of Personhood." In *The Identities of Persons*. ed. Amelie O. Rorty, 175-197. Los Angeles: University of California Press, 1976.
- _____. *The Intentional Stance*. Cambridge: The MIT Press, A Bradford Book, 1987.
- _____. "Precis of *The Intentional Stance*." *Behavioral and Brain Sciences* 11 (1988): 495-546.
- _____. *Consciousness Explained*. Boston: Little, Brown and Company, 1991.

- Doran, Kevin. *What Is A Person: The Concept and Implications for Ethics*. New York: The Edwin Mellen Press, 1989.
- Dougherty, John .H. Jr., Donald G. Rawlinson, David E. Levy, and Fred Plum. "Hypoxic-Ischemic Brain Injury and the Vegetative State: Clinical and Neuro-Pathological Correlation." *Neurology* 31 (1981): 991-97.
- Drane, James. "Competency to Give an Informed Consent: A Model for Making Clinical Assessments." *JAMA: Journal of the Americal Medical Association* vol. 252 (August 17, 1984): 925-927.
- Dresser, Rebecca S. "Advance Directives, Self-Determination, and Personal Identity." In *Advance Directives in Medicine* ed. Chris Hackler, Ray Moseley, and Dorothy Vawter, 155-171. New York: Praeger, 1989 .
- _____. "Autonomy Revisited: The Limits of Anticipatory Choices." In *Dementia and Aging: Ethics, Values and Policy Choices*, ed. Binstock, Robert H., Stephen G. Post, and Peter J. Whitehouse, 71-86. Baltimore: The Johns Hopkins University Press, 1992.
- Dworkin, Roger. "Autonomy and the Demented Self." *Millbank Quarterly* 64 (Supplement 2): 4-16.
- Edelman, Gerald M. *Neural Darwinism: The Theory of Neuronal Group Selection*. New York: Basic Books, 1987.
- _____. *The Remembered Present: A Biological Theory of Consciousness*. New York: Basic Books, 1989.
- Edwards, Rem B. ed. *Psychiatry and Ethics*. New York: Prometheus Books, 1982.
- Edwards, Rem B., and Glenn C. Graber. ed. *Bioethics*. New York: Harcourt Brace Jovanovich, 1988.
- Elster, Jon. ed. *The Multiple Self*. Cambridge: Cambridge University Press, 1986.
- Engelhardt, H.Tristram. *The Foundations of Bioethics*. New York: Oxford University Press, 1986.
- _____. "Medicine and the Concept of Person." In *Contemporary Issues in Bioethics*. ed. Tom L. Beauchamp, and Leroy Walters. Wadsworth Publishing Company, 1981. Reprinted in *What is a Person?* ed. Michael F. Goodman, 169-185. The Humama Press, 1988.
- Feinberg, Joel. "The Rights of Animals and Unborn Generations." In *Philosophy and Environmental Crisis*. ed. W. Blackstone. Athens: University of Georgia Press, 1974. Reprinted in *Ethics for Modern Life* ed. Raziel Abelson and Marie-Louise Friquegnon, 260-270. 4th ed. New York: St. Martin's Press, 1991.

- _____. "The Potentiality Argument." In *Matters of Life and Death: New Introductory Essays in Moral Philosophy*. ed. Tom Regan, T. New York: Random House, 1980. Reprinted as "A Question about Potentiality." in *Moral Issues*. ed. Jan Narveson, 234-239. Toronto: Oxford University Press, 1983.
- _____. *Harm to Self*. New York: Oxford University Press, 1986.
- Fetzer, James. *Philosophy and Cognitive Science*. New York: Paragon House, 1991.
- Flanagan, Owen. *The Science of the Mind*. Cambridge: The MIT Press A Bradford Book, 1984.
- Fox, Jacob. "Alzheimer's Disease and Other Dementias" in *Neurology for the Non-Neurologist*. ed. William J. Weiner, and Christopher G. Goetz, 198-206. 2nd. ed. Philadelphia: J.B. Lippincourt Company, 1989.
- Francis, Leslie P., "Advance Directives for Voluntary Euthanasia: A Volatile Combination?" *The Journal of Medicine and Philosophy* vol. 18 no. 3 (June 1993).
- Frankfurt, Harry. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* LXVII 1 (1971): 15-20. Reprinted in *What is a Person?* ed. Michael F. Goodman, 127-145 Clifton N.J.: The Humama Press, 1988.
- French, Peter A. "Kinds and Person." *Philosophy and Phenomenological Research* 44 2 (1983). Reprinted in *What is a Person?* ed. Michael F. Goodman, 127-145. Clifton, N.J.: The Humama Press, 1988.
- Gareth Jones D. "Brain Death and Personal Identity." *Journal of Medical Ethics* vol.15 no.4 (December 1989)
- Gillett, Grant R. "Why Let People Die?" *Journal of Medical Ethics* vol.12 (1986): 83-86.
- _____. "Reasoning about Persons." In *Persons and Personality* ed. Arthur Peacocke, and Grant R. Gillett, 75-88. Oxford: Blackwell, 1987.
- _____. "Consciousness, the Brain and What Matters." *Bioethics* vol. 4 no. 3 (July 1990): 181-199.
- Glantz, Leonard H., "The Role of Personhood in Treatment Decisions Made by Courts." *Millbank Memorial Fund Quarterly/Health and Society* vol. 61 no. 1 (Winter 1983): 76-101.
- Glover, Jonathan. *Causing Death and Saving Lives*. London: Penguin, 1977.
- _____. *I: The Philosophy and Psychology of Personal Identity*. London: The Penguin Press, 1988.

- Goodman, Michael F. *What is a Person?* ed. Clifton, N.J.: The Humama Press, 1988
- Gregory Richard L. ed. *The Oxford Companion to the Mind*. Oxford: Oxford University Press, 1987.
- Grice, H.P., "Personal Identity" *Mind* vol. 50 (1941): 330-50.
- Hackler, Chris, Ray Moseley, and Dorothy E. Vawter. ed. *Advance Directives in Medicine*. New York: Praeger, 1989.
- Halevy, Amir, and Baruch Brody. "Brain Death: Reconciling Definitions, Criteria, Tests." *Annals of Internal Medicine* vol. 19 no. 6 (1993): 519-525.
- Hirsch, Eli. *The Concept of Identity*. New York: Oxford University Press, 1982.
- Hume, David. *A Treatise of Human Nature*. ed. L.A. Selby-Bigge. Oxford: Clarendon Press, 1978.
- Hundert, Edward M. *Philosophy, Psychiatry and Neuroscience: Three Approaches to the Mind. A Synthetic Analysis of the Varieties of Human Experience*. Clarendon Press: Oxford, 1989.
- Jaynes, Julian. *The Origin of Consciousness in the Breakdown of the Bicameral Mind*. Boston: Houghton Mifflin Company, 1976.
- Jennet, W.B., and F. Plum. "The Persistent Vegetative State: A Syndrome in Search of a Name". *Lancet* 1 (1972): 734-737.
- King, Nancy M.P. *Making Sense of Advance Directives*. Boston: Kluwer Academic Publishers, 1991.
- Kolak, Daniel, and Raymond Martin. *Self and Identity: Contemporary Philosophical Issues*. New York: Macmillan Publishing Company, 1991.
- Kuhse, Helga, and Peter Singer. *Should The Baby Live?* Oxford: Oxford University Press, 1985.
- Lamb, David. *Death, Brain Death and Ethics*. Beckenham: Crook Helm Ltd, 1985.
- Lavin, Michael. "Ulysses Contracts." *Journal of Applied Philosophy* vol. 3 no. 1 (1986): 89-101.
- Lewis, David. "Survival and Identity." In *The Identities of Persons*. ed. Amelie O. Rorty, 175-197. Los Angeles: University of California Press, 1976.
- _____. *On the Plurality of Worlds*. Oxford: Blackwell, 1986.

- Locke, John. *An Essay Concerning Human Understanding*. ed. Peter H. Nidditch. Oxford: Clarendon Press, 1975.
- Lockwood, Michael J. "When does a life begin?" In *Moral Dilemmas in Modern Medicine*. ed. Michael J. Lockwood, 9-32. Oxford: Oxford University Press, 1985.
- Lycan, William G. *Consciousness*. Cambridge: The MIT Press/A Bradford Book, 1987.
- _____. ed. *Mind and Cognition*. Cambridge: Basil Blackwell, 1990.
- Mackie, John L. *Problems from Locke*. Oxford: Clarendon Press, 1976.
- Macklin, Ruth. "Personhood in the Bioethics Literature." *Millbank Memorial Fund Quarterly/Health and Society* vol. 61 no. 1 (Winter 1983): 35-58.
- McCall, Catherine. *Concepts of Person: An Analysis of Concepts of Person, Self and Human Being*. Aldershot, England: Avebury, 1990.
- McDonald, Cynthia. *Mind-Body Identity Theories*. London: Routledge, 1989.
- McGinn, Colin. *The Character of Mind*. Oxford: Oxford University Press, 1982.
- _____. *The Problem of Consciousness*. Oxford: Basil Blackwell, 1991.
- Meisel, Alan, Loren H. Roth, and Charles W. Lidz. "Toward a Model of the Legal Doctrine of Informed Consent." *The American Journal of Psychiatry* 134 (1977): 285-289. Reprinted in *Psychiatry and Ethics*. ed. Rem B. Edwards, 192-201. New York: Prometheus Books, 1982.
- Moussa, M., and T.A. Shannon. "The Search for the New Pineal Gland: Brain Life and Personhood." *Hastings Center Report* vol. 22 no. 3 (May/June 1992): 30-37.
- Nagel, Thomas. "What it is like to be a bat?" *Philosophical Review* 83 (October 1974). Reprinted in *Mortal Questions*, 165-181.
- _____. *Mortal Questions*. Cambridge: Cambridge University Press, 1979.
- _____. *The View From Nowhere*. Oxford: Oxford University Press, 1986.
- Narveson, Jan. ed. *Moral Issues*. Toronto: Oxford University Press, 1983.
- Noonan, Harold W. *Personal Identity*. London: Routledge, 1991.
- Noonan, John T. *The Morality of Abortion: Legal and Historical Perspectives*. Cambridge: Harvard University Press, 1970.
- Nuland, Sherwin B. *How We Die*. New York: Alfred A. Knopf, 1993.

- Parfit, Derek. "Later Selves and Moral Principles." In *Philosophy and Personal Relations* ed. Alan Montefiore, 137-169. Montreal, Canada: McGill-Queen's University Press, 1973.
- _____. "Lewis, Perry, And What Matters." In *The Identities of Persons*. ed. Amelie O. Rorty, 91-109. Los Angeles: University of California Press, 1976.
- _____. *Reasons and Persons*. Oxford: Oxford University Press, 1984.
- Peacocke, Arthur, and Grant R. Gillett. ed. *Persons and Personality*. Oxford: Blackwell, 1987.
- Perry, John. "The Importance of Being Identical" in *The Identities of Persons*. ed. Amelie O. Rorty, 67-91. Los Angeles: University of California Press, 1976.
- _____. *Personal Identity* Berkeley: University of California Press, 1975.
- Persson, Ingmar. "The Indeterminacy and Insignificance of Personal Identity." *Inquiry* 35, 271-283.
- Plum, Fred, and Jerome B. Posner. *The Diagnosis of Stupor and Coma*. 3rd. ed. Philadelphia: F.A. Davis Company, 1980.
- Post, Stephen G., and Peter J. Whitehouse. "Dementia and the Life-Prolonging Technologies Used: An Ethical Question." *Alzheimer's Disease and Associated Disorders* vol. 6 no. 1, Raven Press Ltd., New York. (1992): 6-10.
- Priest, Stephen. *Theories of the Mind*. New York: Viking Penguin, 1991
- Puccetti, Ronald. "The Life of a Person." In *Abortion and the Status of the Fetus* ed. W. B. Bondeson et al. D.Reidel Publishing Company, 1983. Reprinted in *What is a Person?* ed. Michael F. Goodman, 265-281. Clifton, N.J.: The Humana Press, 1988.
- Quinlan, J., J. Quinlan, and P. Batelle. *Karen Ann: The Quinlans Tell Their Story*. New York: Doubleday & Company, Inc., 1977.
- Ramsey, Paul. *The Patient as Person*. New Haven: Yale University Press, 1970.
- _____. *Ethics at the Edge of Life*. New Haven: Yale University Press, 1978.
- Rawls, John. *A Theory of Justice*. Cambridge: Harvard University Press, 1971.
- Regan, Tom. ed. *Matters of Life and Death: New Introductory Essays in Moral Philosophy*. New York: Random House, 1980.
- Reid, Thomas. "Of Memory." In *Essays on the Intellectual Powers of Man*. Reprinted in *Personal Identity*. ed. John Perry, 99-107. Berkeley: University of California Press, 1975.
- Rescher, Nicholas. *Human Interests*. Stanford: Stanford University Press, 1990.

- Rey, George. "Survival." In *The Identities of Persons* ed. Amelie O. Rorty, 41-67. Los Angeles: University of California Press, 1976.
- Robertson, John A. "Involuntary Euthanasia of Defective Newborns: A Legal Analysis." *Stanford Law Review* 27 (1975): 246-261.
- _____. "Second Thoughts on Living Wills." *Hastings Center Report* vol. 21 no. 6 (November-December 1991): 6-10.
- Rorty, Amelie O. ed. *The Identities of Persons*. Los Angeles: University of California Press, 1976.
- _____. *Mind in Action: Essays in the Philosophy of Mind*. Boston: Beacon Press, 1988.
- Rosenthal, David M. ed. *The Nature of Mind*. New York: Oxford University Press, 1991.
- Roth, Loren H., Alan Meisel, and Charles W. Lidz. "Tests of Competency to Consent to Treatment." *The American Journal of Psychiatry* 134 (1977): 279-284. Reprinted in *Psychiatry and Ethics*. ed. Rem B. Edwards, 201-211. New York: Prometheus Books, 1982.
- Sacks, Oliver. *The Man Who Mistook His Wife For A Hat And Other Clinical Tales*. New York: Simon & Schuster, Inc., and New York: Harper Collins Publishers, Inc., 1970.
- _____. *Awakenings*. New York: Harper Collins Publishers, Inc., 1973.
- _____. *A Leg To Stand On*. New York: Simon & Schuster, Inc., and New York: Harper Collins Publishers, Inc., 1984.
- _____. *Seeing Voices*. Berkeley: University of California Press and New York: Harper Collins Publishers, Inc., 1989.
- Schechtman, Marya. "Personhood and Personal Identity." *Journal of Philosophy* vol. 87 no. 2 (February 1990): 71-93.
- Searle, John R. *Minds, Brains and Science (The 1984 Reith Lectures)*. Cambridge: Harvard University Press, 1984.
- Shoemaker, Sydney. *Self-Knowledge and Self-Identity*. Ithaca: Cornell University Press, 1963.
- _____. "Persons and their Pasts" *American Philosophical Quarterly* vol. 7 no. 4 (October 1970): 269-85. Reprinted in *Identity, Cause and Mind*, 19-49.
- _____. *Identity, Cause and Mind*. Cambridge: Cambridge University Press, 1984.
- Shoemaker, Sydney, and Roger Swinburne. *Personal Identity*. Oxford: Blackwell, 1984.

- Soderberg, William. "Do Living Will Statutes Embody a Claim Right or a Mere Privilege To Refuse Life-Prolonging Treatment?" In *Advance Directives in Medicine*, ed. Chris Hackler, Ray Moseley, and Dorothy E. Vawter, 75-85. New York: Praeger, 1989.
- Stitch, Stephen. "Dennett on Intentional Systems." *Philosophical Topics* vol.12 38-62.
- Strawson, Peter F. *Individuals*. London: Methuen, 1959.
- Taylor, Charles. *Sources of the Self*. Cambridge: Harvard University Press, 1989.
- The Hastings Center. *Guidelines on the Termination of Life-Sustaining Treatment and the Care of the Dying*. Bloomington: Indiana University Press, 1987.
- Tooley, Michael. *Abortion and Infanticide*. Oxford: Clarendon Press, 1983.
- _____. "In Defence of Abortion and Infanticide." In *What is a Person?* ed. Michael F. Goodman, 83-115. Clifton N.J.: The Humana Press, 1988.
- Thomasma, David, and Glenn C. Graber. *Euthanasia: Toward an Ethical Social Policy*. New York: The Continuum Publishing Company, 1990.
- Thompson, Judith Jarvis. "A Defense of Abortion." *Philosophy and Public Affairs* vol. 1 no. 1 (1971): 47-66.
- Unger, Peter. *Identity, Consciousness and Value*. New York: Oxford University Press, 1990.
- Veatch, Robert V. "The Impending Collapse of the Whole-Brain Definition of Death." *Hastings Center Report* vol. 23 no. 4 (July/August 1993).
- Warren, Mary Ann. "On the Moral and Legal Status of Abortion." *Monist* vol. 57 no. 1 (1973): 43-61. Reprinted with "Postscript on Infanticide" In *Bioethics* ed. Rem B. Edwards, and Glenn C. Graber, 583-591. New York: Harcourt Brace Jovanovich, 1988.
- Weiner, William J., and Christopher G. Goetz. *Neurology for the Non-Neurologist*. 2nd. ed. Philadelphia: J.B. Lippincourt Company, 1989.
- Weir, Robert F. *Ethical Issues in Death and Dying*. New York: Columbia University Press, 1977.
- White, Alan R. *Rights*. Oxford: Oxford University Press, 1984.
- White, Stephen L. *The Unity of the Self*. Cambridge: The MIT Press/A Bradford Book, 1991.
- Wiggins, David. *Identity and Spatio-Temporal Continuity*. Oxford: Blackwell, 1967.
- _____. *Sameness and Substance*. Cambridge: Harvard University Press, 1980.

- _____. "The Person as Object of Science, as Subject of Experience, and as Locus of Value." In *Persons and Personality: A Contemporary Inquiry*. ed. Arthur Peacocke, and Grant R. Gillett, 56-75. Oxford: Blackwell, 1987.
- Wikler, Daniel. "Not Dead, Not Dying? Ethical Categories and Persistent Vegetative State." *Hastings Center Report* vol. 18 no. 1 (February/March 1988).
- Wilkes, Kathleen V. *Real People*. Oxford: Clarendon Press, 1988.
- Williams, Bernard. "Personal Identity and Individuation." *Proceedings of the Aristotelian Society*, 57 229-252. Reprinted in *Problems of the Self*, 1-19.
- _____. "The Self and the Future." *The Philosophical Review* 79 161-80. Reprinted in *Problems of the Self*, 46-64
- _____. *Problems of the Self*. Cambridge: Cambridge University Press, 1973.
- Wittgenstein, Ludwig. *The Blue and Brown Books*. Oxford: Blackwell, 1958.
- _____. *Philosophical Investigations* 3d. ed. Translated by Gillian E.M. Anscombe. New York: Macmillan, 1968.

VITA

Thomas Gilbert Buller was born in Surrey, England on April 3, 1958. He attended Leighton Park School, Reading and Cambridge College of Arts and Technology, Cambridge. He read Philosophy at York University, England, graduating in 1981. In 1982 he studied at the University of Waterloo, Canada, receiving a Master's Degree in Philosophy in 1984. From 1985-86 he studied Medical Ethics at King's College, London and received a Postgraduate Diploma in Medical Ethics and Law. In 1988 he began his doctoral studies at the University of Tennessee, Knoxville graduating with a Ph. D. in Philosophy in 1994.