

Claremont Colleges

Scholarship @ Claremont

---

CMC Senior Theses

CMC Student Scholarship

---

2021

## The Self and Memory: How Theories of Self Are Crucial in Aiding End-of-Life Decision Making for Dementia Patients

Eric Oregel

Eric E. Bauer

Follow this and additional works at: [https://scholarship.claremont.edu/cmc\\_theses](https://scholarship.claremont.edu/cmc_theses)

---

### Recommended Citation

Oregel, Eric and Bauer, Eric E., "The Self and Memory: How Theories of Self Are Crucial in Aiding End-of-Life Decision Making for Dementia Patients" (2021). *CMC Senior Theses*. 2735.

[https://scholarship.claremont.edu/cmc\\_theses/2735](https://scholarship.claremont.edu/cmc_theses/2735)

This Open Access Senior Thesis is brought to you by Scholarship@Claremont. It has been accepted for inclusion in this collection by an authorized administrator. For more information, please contact [scholarship@cuc.claremont.edu](mailto:scholarship@cuc.claremont.edu).

Claremont McKenna College

The Self and Memory: How Theories of Self Are Crucial in Aiding End-of-  
Life Decision Making for Dementia Patients

Submitted to  
Professor Dustin Locke

By  
Eric Oregel

Submitted For  
Senior Thesis  
Fall 2020 - Spring 2021

*Acknowledgements*

As graduation nears and the submission of this thesis is finalized, I find myself overwhelmed with gratitude for the experiences I've been afforded leading up to this moment and for the people that have made an indelible impact on my life. I feel proud and humble to have had the opportunity to attend CMC with such excellent faculty and to have benefitted from all that a liberal arts education offers. I would like to thank all of my professors that I have had the privilege of being taught by, especially my professors in economics and philosophy. Additionally, I would like to thank Professor Locke, my thesis reader, for his enduring patience and support throughout this process.

I feel blessed and so very lucky to know that my family has always believed in me and encouraged me to follow my passions. To Mom, Dad, Mark, and Peter, thank you for all that you've done and continue to do for me day in and day out. As will hopefully be apparent by the end of this thesis, who we are is the stuff of memories and experiences and you have all given me the very best.

Finally, I would like to acknowledge my grandmothers, both of whom, unfortunately endured dementia and have since passed away. Your warmth has been with me throughout the writing process and you have both inspired me to take a deep dive into the subject of who we are as persons. Thank you so much, I miss you always.

*Table of Contents*Chapter One:

<i>Introduction</i> .....	(3)
1. <i>Lockean Theory of Self</i> .....	(5)
1.1 <i>The Identity of Objects Over Time</i> .....	(5)
1.2 <i>The Identity of Persons Over Time</i> .....	(9)
2. <i>Popular Alternatives to the Lockean Theory of Personal Identity</i> .....	(11)
2.1 <i>Immaterial Soul</i> .....	(11)
2.2 <i>Materialist Alternative</i> .....	(12)
3. <i>Important Critiques of Locke's Theory</i> .....	(14)
3.1 <i>The Breakfast Problem</i> .....	(14)
3.2 <i>The Brave Officer</i> .....	(15)
3.3 <i>The Branching Problem</i> .....	(16)
4. <i>Chapter One Summary</i> .....	(16)

Chapter Two:

<i>Chapter Two Introduction</i> .....	(18)
1. <i>The Branching Problem</i> .....	(19)
1.1 <i>Brain Split</i> .....	(19)
2. <i>The Fickle Nature of Memory</i> .....	(24)
3. <i>Severe Memory Loss</i> .....	(24)
3.1 <i>Advance Euthanasia Directive</i> .....	(26)

Chapter Three:

<i>Chapter Three Introduction</i> .....	(30)
1. <i>Centering the Issue</i> .....	(30)
2. <i>Irrational Weight of the Past</i> .....	(32)
3. <i>Four Philosophers Discuss AEDs</i> .....	(35)
3.1 <i>Dworkin</i> .....	(37)
3.2 <i>Hawkins</i> .....	(37)
3.3 <i>Koppelman</i> .....	(38)
3.4 <i>Locke</i> .....	(39)
4. <i>Altered AED?</i> .....	(39)
5. <i>In Conclusion</i> .....	(41)
<i>Works Cited</i> .....	(42)

## Chapter One

### Introduction

One of the simplest, yet, most daunting questions one could ask is, “Who are you?” Of course, we are many things. One might respond with their name, occupation, racial identity, sexual orientation, or a plethora of other identities. However, behind all of those identifiers is something more foundational and it doesn’t take much digging to know intuitively that who we consider ourselves to be is an extremely profound and complex notion. For millennia, philosophers have contemplated this question; from Socrates to contemporary thinkers such as John Perry and Derek Parfit.

Likewise, popular answers to this inquiry have varied widely. Traditional Materialists might posit that the essence of ‘who we are’ resides in the body, or the physical components of our being. In accordance, any alterations or changes to the body would coincide with a change in the self. Others, such as Descartes, would posit that there is an immaterial substance, such as a soul, that embodies our being. However, in 1690 the English philosopher John Locke added another layer to the puzzle in his work *Of Identity and Diversity*. It is in this book that Locke would provide his groundbreaking account of personal identity that this paper will focus on.

Though various theories of self have been extensively explored, the reason it has persisted for so long may not seem obvious. Why should we concern ourselves with what the self is and where it resides? In part, the goal of this paper is to answer this question in presenting what I take to be the most promising theory of self. I will then offer an intriguing wrinkle concerning the notion of severe memory loss and how this might change how we think about the

self. The paper will close by closely examining Alternative Euthanasia Directives (AEDs) and their plausibility.

Before I delve into the various theories of self, it may help to lay the foundation for this paper using a practical example. As we will see later on, millions of folks around the world currently suffer from dementia and the number is slated to grow. Dementia and its specific forms, such as Alzheimer's, cause severe memory loss in their victims and thus a sense of self becomes more difficult to distinguish. This can make decisions extremely difficult, especially end of life decisions. To quell this uncertainty, Advance Euthanasia Directives (AEDs) have been developed to help dementia patients die on their own terms. Simply put, these are agreements in which people in the early stages of dementia authorize the terms of their death in advance. These agreements make it so that if someone that previously authorized an AED were to fall ill then doctors would not intervene to help save the patient, they would simply let them die. This seems straightforward but we will see that the situation becomes complicated when the patient's dementia has progressed far enough that they cannot remember authorizing said AED and wish to opt out of it altogether. Intuitively, we might think that there should be no question in this scenario and that the person in question should be free to opt out of the AED should they wish. However, this poses challenges such as respecting the wishes and autonomy of the person that authorized the initial AED. As we will see, this case warrants investigating the question of the self in order to arrive at a thorough solution to this case.

In Chapter One, I will begin with a survey and an outline of Locke's theory, beginning with how we should go about defining objects and persons and ultimately arriving at his theory of personal identity. I will then address the two most popular alternative theories, the Cartesian View and the Materialist View, and will conclude by briefly addressing a few of the most

important influential critiques of the theory. Chapter Two will seek to address the Branching Problem identified later in Chapter One as well as the ramifications of severe memory loss and how we might view persons as a result. Finally, in Chapter Three I will delve deeper into the issue of Advanced Euthanasia Directives and ultimately determine whether or not they are still plausible within our current knowledge framework.

## 1. Lockean Theory of the Self

### 1.1 The Identity of Objects Over Time

John Locke's (1632-1704) theory of self can be found in his seminal work, *Essay on Human Understanding*. It should be said that this was one of the most influential theories of the self at the time because for Locke 'the self' resides not in the body or the soul but in what he calls a 'continuity of consciousness.' Locke defines this notion as follows:

“For, since consciousness always accompanies thinking, and it is that which makes every one 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 a 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 it is by the same self with this present one that now reflects on it, that that action was done.”

Importantly, Locke identifies that sameness of body and soul are neither necessary, nor sufficient for sameness of person. As we will see later, Locke's theory posits that the crux of consciousness resides in memory. In order for there to be continuity of consciousness, said consciousness must extend into the past via memory: I am only myself insofar as I have past memories of my own thoughts and actions and they serve to inform my future self.

However, before delving into the core of his theory a foundational question must first be answered, namely, “What does it mean for a thing to be the ‘same thing’ over time?” Locke states:

“When we see anything to be in any place in any instant of time, we are sure (be it what it will) that it is that very thing, and not another, which at that same time exists in another place, how like and indistinguishable soever it may be in all other respects: and in this consists identity, when the ideas it is attributed to vary not at all from what they were at the moment wherein we consider their former existence, and to which we compare the present”(qtd. in Perry 33).

Essentially, Locke says that the criteria for identity involves having the same set of ideas and beliefs over time. If these ideas and beliefs come to differ in the future then we may be dealing with a categorically different person. Importantly, at this point the question can pertain to literally anything: a rock, chair, person, animal, etc. If we were to take a pencil, for example, what can we say about its essence that would allow us to determine whether it is the same pencil when we see it again in three days time? Or, is it different? Locke would phrase this question as: “Is this pencil *identical* (same) or *distinct* (different) from that which I saw three days ago?”

Now, in order to even begin to answer this question it is necessary to delve a level deeper and ask what it means to be identical over a period of time. Intuitively, we know that the answer to this question will likely rely on the nature of the subject in question. For example, seeing a rock through time would be much different than observing a university or person over time. This too is a bedrock principle for Locke’s theory: when we ask what it means for an object or entity to be identical over time, the answer depends on what type of thing the object is. Locke begins his answer by defining four foundational types of ‘things’: atoms, complex bodies, living organisms, and people. For Locke, it seems as though these categories are not meant to be

mutually exclusive. Rather, as we continue they will undergo further refinement. For Locke, the case of atomic identity is somewhat simplistic:

“Let us suppose an atom, i.e., a continued body under one immutable superficies, existing in a determined time and place; it is evident, that considered in any instant of its existence, it is in that instant the same with itself. For, being at that instant what it is, and nothing else, it is the same, and so must continue as long as its existence is continued; for so long it will be the same, and no other” (qtd. In Perry 35).

The case with atoms seems straightforward as Locke takes them to be ‘immutable.’ Therefore, they cannot change and one atom from time A to time B will be the same. The sameness will persist over time. After all, at least in Locke’s perception these are the simplest entities with a finite, unchanging structure. However, things become more complicated when considering a *complex body*:

“In like manner, if two or more atoms be joined together into the same mass, every one of those atoms will be the same, by the foregoing rule: and whilst they exist united together, the mass, consisting of the same atoms, must be the same mass, or the same body, let the parts be ever so differently jumbled. But if one of the atoms be taken away, or one new one added, it is no longer the same mass of the same body” (qtd. In Perry 35).

So, we now have a body of atoms and were this body to maintain the same atoms over time, then it would be unchanging from time A to time B. However, it does in fact change if atoms are added or subtracted from the whole. Now, if we continue with this line of thinking we will soon encounter a few issues. Take a plant for example: I intuitively believe that a rose in my garden is the same rose even if it sheds petals and leaves. The petals may be different but the stem and root system are essentially the same. Inherently, it will shed and acquire new atoms throughout its lifetime. This is an issue that Locke addresses:

“In the state of living creatures, their identity depends not on a mass of the same particles, but on something else. For in them the variation of great parcels of matter alters not the identity: an oak growing from a plant to a great tree, and then lopped, is still the same

oak; and a colt grown up to a horse, sometimes fat, sometimes lean is all the while the same horse...

We must therefore consider wherein an oak differs from a mass of matter, and that seems to me to be in this; that the one is only a cohesion of particles of matter any how united, the other such a disposition of them as constitutes the parts of an oak; and such organization of those parts, as is fit to receive, and distribute nourishment, so as to continue, and frame the wood, bark, and leaves, etc. of an oak, in which consists vegetable life.

The case is not so much different in brutes, but that one may hence see what makes an animal, and continues it the same" (qtd. In Perry 35).

Implicitly, Locke identifies a core of the object and while appendages of the core may alter over time, so long as the core is the same we may refer to the object as essentially the same over time. At this stage, we now have clarificatory points regarding the status of living things and the conditions under which they retain identity, namely, 'one continued body.' Insofar as parts may change or be altered, we know that a mass persists over time as long as it retains this unifying factor. Specifically, the unifying factor in question is organizational, i.e. in order for an object at T1 to be the same object at T2, it must have the same organization or the same way in which matter is structured. This will pertain to the way in which we talk about bodies as well.

Finally, we may move to humans:

"This also shews wherein the identity of the same man consists; viz. In nothing but a participation of the same continued life, by constantly fleeting particles of matter, in succession vitally united to the same organized body. He that shall place the identity of man in anything else, but like that of other animals in one fitly organized body, taken in any one instant and from thence continued, under one organization of life, in several successively fleeting particles of matter united to it, will find it hard to make an embryo, one of years, mad and sober, the same man, by any supposition, that will not make it possible for Seth, Ismael, Socrates, etc. to be the same man. For, if the identity of the soul alone makes the same man, and there be nothing in the nature of matter why the same individual spirit may not be united to different bodies, it will be possible that those men living in distant ages, and of different tempers, may have been the same man... And that way of speaking would agree yet worse with the notions of those philosophers who allow

of transmigration, and are of the opinion that the souls of men may, for their miscarriages, be detruded into the bodies of beasts, as fit habitations, with organs suited to the satisfaction of their brutal inclinations'' (qtd. In Perry 35).

Locke has now offered his theory regarding the identity of man. Locke states that if the self were to reside in an immaterial soul then we might be able to unite individual spirits to different bodies, and by such logic, these new beings would be numerically identical. This is not to be confused with qualitatively identical persons, those which share like properties. Similarly with the case regarding living things, we see that there is a condition of a 'continued life.' However, Locke is careful not to endorse a notion of an immaterial soul, similar to the Cartesian theory of self. You might think that by offering a theory of identity for humans, Locke has provided a theory of identity for persons. However, he draws a distinction between the two.

## 1.2 The Identity of Persons Over Time

Now that we have an understanding of the four different categories of things that Locke has chosen to isolate and how we might view their identities over time, we can move towards identity concerning persons. For Locke, there is a critical distinction to be made between persons and what we might call *humans* or *bodies* in the previous section:

"It is not therefore unity of substance that comprehends all sorts of identity, or will determine it in every case; but to conceive and judge of it aright, we must consider what idea the word it is applied of stands for: it being one thing to be the same substance, another the same man, and a third the same person, if person, man, and substance are three names standing for three different ideas; for such is the idea belonging to that name, such must be the identity" (qtd. In Perry 37).

Locke again draws on the notion established prior: determining what it means for a thing to remain identical over time relies heavily on what it is. He adds an element here by stating that because the man, person, and substance are all categorically different, their respective criteria for

identity must accordingly differ. So, what does it mean for a person to be identical over time?

Well, it might help to begin by providing Locke's definition of a person:

“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 only by that consciousness, which is inseparable from thinking, and as it seems to me essential to it...” (qtd. In Perry 39).

Here, Locke posits that a metacognition of sorts is a necessary condition for persons, that we must be able to think about ourselves as ourselves through reason and reflection. The mechanism that reflection constitutes is crucial to understanding how Locke defines persons. ‘Reflection’, in this context, should perhaps be thought of as a sixth sense of sorts. In *Personal Identity*, John Perry puts it quite well:

“Locke's controversial belief that reflection is best thought of as a sense - like vision or hearing - is not crucial. What is important seems less controversial: that we do have such a way of finding out about the occurrence of experiences, and that all the experiences we find out about in this way are our *own*. I can be ‘reflectively aware’ of the occurrence of my own experiences, but not of yours” (Perry 13).

This leads directly into his notion of personal identity:

“For since consciousness always accompanies thinking, and it's 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 a 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...” (qtd. In Perry 13).

Now, we know that there may exist a person A at a time X and a person B at time Y and that persons A and B may be the same person if and only if they contain the same consciousness.

However, while this seems resolute, there is still one more piece of the puzzle, namely, identifying what consciousness is in this context. To answer this question, again we turn to Locke:

“For as far as any intelligent being can repeat the idea of any past action with the same consciousness it had of it at first, and with the same consciousness it has of any present action; so far it is the same personal self” (qtd. In Perry 40).

This is the essence of Locke’s theory: for a person to be identical over time would mean that they maintain the ability to recall and act based on prior belief, action, experience, etc. and that such actions are rooted in memory. For Person B to contain the same consciousness as Person A, Person B must be able to draw on or ‘repeat’ the same elements of Person A’s consciousness. In this sense, memory becomes the crucial mechanism which facilitates the necessary recall.

**Formally:**

If, and only if, person B can remember person A’s thoughts are they identical.

2. Popular Alternatives to the Lockean Theory of Personal Identity

2.1 Immaterial Soul:

Of probable alternatives, there have historically been two which rival Locke’s memory theory. The first states that the essence of self is constituted in the sameness of the immaterial soul. This is in line with the Cartesian theory of self and posits that each of us contains an immaterial soul in which the soul is a non-physical thinking substance that is distinct from the body. In this view, it is both a necessary and sufficient condition that for Person B to be Person A, they must contain the same soul. One implication of this view is that the physical elements of a person are irrelevant to the self. In sections 12-14 of *On Identity and Diversity*, Locke argues against this idea. Here are two poignant passages that highlight his skepticism:

“And therefore those, who place thinking in an immaterial substance only... must shew why personal identity cannot be preserved in the change of immaterial substances,, as

well as animal identity is preserved in the change of material substances..." (qtd. In Perry 40).

"All those who hold pre-existence are evidently of this mind, since they allow the soul to have no remaining consciousness of what it did in that pre-existent state, either wholly separate from the body, or informing any other body; and if they should not, it is plain experience should be against them" (qtd. In Perry 43).

Inherent in this idea is the notion that the soul is a theoretically transferable entity. If my soul can occupy another animal or object, then that object will effectively become Eric. While it is important to note that Locke's theory is fundamentally opposed to this idea, he does not deny that it might be possible that a person's consciousness might reside in the soul, though it is not the soul entirely which accounts for personal identity.

## 2.2 Materialist Alternative:

Another popular alternative to Locke's hypothesis is that of the Materialist view. This view states that identical selves rely wholly on sameness of body over time. Formally, maintaining sameness of body is both a necessary and sufficient condition for Person B to be identical to Person A. The following passage presents Locke's argument against this view:

"Happiness and misery, being that, for which everyone is concerned for himself, not mattering what becomes of any substance, not joined to, or affected with that consciousness. For as it is evident in that instance I gave but now [of having one's little finger cut off], if the consciousness went along with the little finger, when it was cut off that would be the same self which was concerned for the whole body yesterday, as making part of itself, whose actions it then cannot but admit as its own now. Though if the same body should live, and immediately from the separation of the little finger have its own peculiar consciousness, whereof the little finger knew nothing, it would not at all be concerned for it, as a part of itself, or could any of its actions, or have any of them imputed to him" (qtd. In Perry 46).

Now, this is perhaps an exaggerated example. However, a more applicable example might be that of cells as they pertain to our skin composition. We all shed skin cells, such that, over the course of our respective lifetimes we each will completely shed our skin approximately every six weeks. Granted, every six weeks there is no mass exodus of skin from our bodies, it happens gradually over time. But if you were to take Person A at Time 0 weeks and Person A at Time 6 weeks, this person would likely have none of the same skin cells that they had at Time 0 weeks. In this case, the Materialist would say that they are essentially different people whereas Locke would maintain that the core of self remains unchanged.

However, that is not to say that the physical components of our being do not affect the self within Locke's framework. For example, let's say that right now, at age 21, I was to lose my left arm. This would certainly affect my being under Locke's theory, but only insofar as it would hinder my ability to participate in tasks that I would have normally been able to when I had access to my left arm. I would not cease to be Eric, but I would certainly be different from Eric that had his left arm. I would have to relearn to write with my right hand, couldn't rock climb anymore, and my everyday life would be altered relative to my life before. In this sense, so would the memories that I make and therefore, my consciousness would fundamentally be altered. It is plausible that being a person with a disability now changes my core preferences in that I might become an advocate for disability rights. So, Locke would say that, yes, losing my left arm could alter myself but not simply because I lost my arm and that being is arbitrarily tied to this conglomeration of tissue, bone, and muscle, but because in losing it I lose something that alters the way I form memories and the types of memories I am able to have. In this way, it alters my consciousness.

Arguably, one of the main benefits of Locke's view is that it moves away from a dependency on the body in identifying the self. Though it should be said that people derive a substantial part of their identity from their outward characteristics, it is hardly the keystone element. It could be argued that outwardness is, in some ways, a manifestation of a person's inward inclinations. In this sense, their outward identities are informed almost entirely by perceptions and ideas resulting from inward reflection and internalization of external stimuli. On the other hand, ability is a part of one's outward identity that is influential to the experiences and memories of said individual. The fact that my body has afforded me certain experiences is indisputably a reason it should be included as part of my identity. However, again it can be said that it only goes so far as to inform my inward ideas and perceptions.

### 3. Important Critiques of Locke's Theory

#### 3.1 The Breakfast Problem

One famous critique of Locke's theory is known as the 'Breakfast Problem' (Della Rocca, "Locke on Personal Identity"). Imagine you have a rather bland and uneventful breakfast one morning and are randomly asked by a colleague the next day what you had for breakfast yesterday. When called upon to recall what you had for breakfast the next day, you can't remember. Now, the question of what one ate for breakfast on any given morning seems comically trivial with regards to who they are as a person. However, in strict accordance with Locke's theory, this would constitute a change in person because it would violate the necessary condition with regard to memory. Locke acknowledges this fault in his theory but he doesn't really offer a satisfactory solution. Instead, he essentially chooses to 'bite the bullet'. Ultimately,

he will claim that Person A, they who consumed the breakfast, will be a different person from Person B, they who failed to recall what Person A ate. However, due to their same organization, we may know them as the same human being. If this seems like a weak response, it certainly is and one which I will touch upon later in this paper. I believe there to be an intuitive hierarchy in experience, roughly rooted in whether or not an experience contributes transformatively to who the person considers themselves to be.

### 3.2 The Brave Officer

The “Brave Officer” is another important objection to Locke’s theory, brought about by Thomas Reid. Recall that the sameness of consciousness is both necessary and sufficient for personal identity under Locke’s view. Now, the problem is as such: There are three time periods, T1, T2, and T3 (sorted chronologically) each of which correspond to a certain Person A, B, and C. Now, imagine Person A at T1 to be a 10 year-old boy that steals some candy or commits a petty act. In 15 years, this boy is now Person B at T2 who happens to be an officer in the military and wins a major battle. This is the ‘brave officer’. At point T2, Person B remembers the act that he committed as Person A at T1. From T2, fast forward to T3, 40 years later. Person C is now 65 years old and a retired general. Person C can distinctly remember being 25 and winning the major battle, however, at this point he has lost the memory from when he was 10. Now, according to Locke, Person C is identical to Person B because he can recall winning the battle and Person B is identical to Person A because Person B can remember stealing the candy. However, Person C and Person A have no memory-based connection. This essentially breaks the notion of transitivity:  $A = B, B = C, A \neq C$  and fails to obtain the necessary condition for personal identity. Now, this issue wasn’t addressed in *On Identity and Diversity* because Locke

wasn't alive to hear it but it is one that contemporary Locke scholars have been forced to reckon with. It is my belief that this is one of the major weaknesses in Locke's theory as it stands and one which will be addressed in Chapter 2 of this paper.

### 3.3 The Branching Problem

The issue of the branching problem assumes the ability to essentially take a person, Person A, and duplicate their brain twice such that there is a Person B with Person A's memory and a Person C with the exact same thoughts and memories. This effectively duplicates the consciousness of A within the minds of B and C. This renders both, Person B and C equally likely competitors for being A. One might say, "Well, couldn't and wouldn't they both simply be A?" At first glance, this might seem plausible, however, they both cannot be A because B and C are not quantifiably identical. And thus we have another violation of the principle of transitivity of identity. Then, the question remains: Who should be considered A? This points to an arbitrariness present in the theory. Many philosophers think that this example shows there to be too much continuity in the theory and that this is a failure of personal identity over time within Locke's framework.

## 4 Chapter One Summary

After having presented Locke's view, popular alternatives, and several key issues, question remains: has Locke put forth a comprehensive and acceptable theory of the 'self'? In many ways, I believe he has succeeded. I have already compared his theory to the likes of Descartes as well as the materialist view, however, in considering the examples used to explain his argument we see how profound his statement regarding memory is. It should be said that we

are only ourselves so long as we have a stock of experiences which influence our future endeavors in thought and action. As I hope to prove in subsequent chapters, whether we like it or not our 'self' is certainly fallible concerning memory. We may hold the notion of the self to be an entity that resides beyond the reach of something so corruptible as memory, however, in reality it cannot.

In the next chapter, I will look closely at two key examples that have the highest probability of pushing Locke's theory to its breaking point. One is the Brave Officer example and I will enlist the help of Derek Parfit in developing a sufficient revision to the theory so that it accounts for discontinuous memory. Secondly, I will consider the case of those with significant degenerative cognitive diseases, such as dementia, and how we are to go about considering testimony from Person A who exists pre-dementia and Person B who exists during dementia.

*Chapter Two*

In the first chapter, I provided a thorough explanation of John Locke's theory of personal identity. In doing so, a few very significant concerns were raised regarding the strength of the theory and how it might hold up practically. In this chapter, I aim to address two of those issues in an effort to see how we might either augment Locke's theory to more effectively encompass such issues, or provide insight as to how we might view these issues from a different lens. In doing so, we might realize that they do not actually pose an issue. The first issue I will consider is that of the 'branching problem' which I briefly discussed in the first chapter. This issue might seem like an arbitrary thought experiment that has no real bearing on Locke's theory. However, the ambiguity that it presents with regards to personhood is certainly a vulnerability of the model. The second issue will be referred to as the 'memory-loss issue'. Hopefully, by now it is clear that the idea of memory is the keystone of Locke's theory of personal identity. However, we know very well, and I'm sure Locke did as well, that memory is a fallible entity. We have seen smaller-scale instances of how its corruptibility threatens to undermine his theory (the brave officer example). However, folks that suffer from severe memory-loss such as amnesia or dementia pose an interesting case study for this theory. What happens when we lose a significant amount of our memory? My goal for the end of this chapter will be to have explained and resolved both of these issues sufficiently so that I may ultimately incorporate them in chapter three.

## 1 The Branching Problem

### 1.1 Brain Split

The following is an example which has been expanded upon in popular literature but first published by the philosopher David Wiggins in *Identity and Spatio-Temporal Continuity* (Wiggins 50). Imagine that there is a person, Person A, and they are an adult with a significant amount of lived experience and thus have various memories about their life. Now, imagine that we have the technology to separate the brain into two equal parts (which we do indeed have), both equal in size and cognitive capacities. We then take one half and implant it equally into two different people, Person B and Person C. When Persons B and C both wake up from surgery, they now contain the memories and beliefs of Person A. To clarify, Persons B and C now contain only the memories of Person A. The implications of this transformation are that, according to Locke's theory, Persons B and C would have met both the necessary and sufficient conditions for being Person A. Though this is a highly unusual situation, the question still remains of whether this result necessarily spells trouble for Locke's theory. Couldn't we simply say that, yes it is odd that there are now two people which are essentially Person A? Unfortunately, it isn't so simple. Though at the point of conception, both B and C are both equally good 'candidates' for being person A, we cannot say that they are both identical to A because B and C are not identical to each other. They are numerically different. Who, then, should we refer to as Person A?

This example exposes one of the larger Achilles' heels of Locke's theory, in that we are now led to believe that the notion of who Person A is becomes essentially arbitrary. At this

junction, Locke's theory takes us no further. There is no rule of thumb regarding who we should take Person A to be; we might as well flip a coin.

In having read a significant portion of the literature concerning this argument, I find it appropriate to begin with Derek Parfit's proposed resolution. Parfit begins by imagining the same scenario that we have with regards to Persons A, B, and C. Parfit then considers three possibilities: (1) that person A does not survive (2) Person A survives as one of the two people and (3) that person A survives as both. Parfit quickly does away with (1) by stating the following:

"The trouble with (1) is this. We agreed that I could survive if my brain were successfully transplanted. And people have in the past survived with half their brain destroyed. It seems to follow that I could survive if half my brain were successfully transplanted and the other half were destroyed. But if this is so, how could I *not* survive if the other half were also successfully transplanted? How could a double success be a failure?" (qtd. in Perry 201)

In disregarding the possibility the Person A (him in this case) does not survive the transplant, Parfit states that even if the 'numerical' version of A, the original physical vessel, were to die but the memories were to persist, then that would indeed qualify as survival. This would fall into line with Locke by essentially rejecting any materialist notions of the self.

The second option, that person A survives as one of the two people, seems to be a weak consideration and Parfit quickly dismisses it citing that even were we to consider it successful in the transfer of persons, because both sides of the brain were purported to be the same, they should indeed have the same ramifications insofar as the transfer of memory is concerned. Therefore, if one were a success then so too should the other.

The third case seems to warrant much more thought from Parfit. To frame his case, he imagines a modern scenario in which one's brain is cut into two pieces. In the first body, he loses the ability to see out of one eye and loses control over the corresponding side of his body. His

experiences have been separated into what he calls two separate ‘streams’. The following quote may help to further clarify,

“In describing the episode, I assumed that there were two series of thoughts, and that they were both mine. If my two hands visibly wrote out two calculations, and if I claimed to remember two corresponding series of thoughts, this is surely what we should want to say. If it is, then a person’s mental history need not be like a canal, with only one channel. It could be like a river, with islands, and with separate streams” (qtd. in Perry 203).

At this point, Parfit claims that we would have two bodies and a ‘divided mind’.

This is the moment at which Parfit diverges from the popular rhetoric and proposes a somewhat radical change, “The alternative, for which I will argue, is to give up the language of identity. We can suggest that I survive as two different people without implying that I am these people” (qtd. in Perry 203). Here, he turns the conversation away from ‘identity’ as we have traditionally considered it and towards ‘survival’. Survival of the self, that is.

Parfit continues by arguing that what we should be most concerned with is not the one-to-one relation that scholars have historically held as the ideal with regards to ‘identity’, but a way in which, “...one person can survive as two” (qtd. in Perry 206). If we think about the one-to-one relation that we usually strive for with regards to personal identity, the incentive in doing so seems obvious: if we can establish one-one relationships, relationships in which we can say that  $x$  is numerically identical to  $y$ , then we can make definitive statements about the two because we would view them as the same person. Parfit, however, thinks that this ideal might be impossible to achieve with regards to the branching problem so we must necessarily seek a new means of relationship:

“If psychological continuity took a branching form, no coherent set of judgements of identity could correspond to, and thus be used to imply, the branching form of this relation. But what we ought to do, in such a case, is take the importance which would attach to a judgement of identity and attach this importance directly to each limb of the

branching relation. So this case helps to show that judgements of personal identity do derive their importance from the fact that they imply psychological continuity. It helps to show that when we can, usefully speak of identity, this relation is our ground” (qtd. in Perry 207).

There is, however, a rhetorical snag as ‘psychological continuity’ has always implied a one-one relationship between the two entities that it bridges. Therefore, we need a new concept if we are to move forward with what Parfit suggests.

Instead, Parfit suggests the idea of ‘psychological connectedness’ in replacing ‘psychological continuity.’ The notion of psychological connectedness is less stringent and ‘weaker’ than psychological continuity, however, this is necessarily so. Again, when speaking in terms of identity, the branching problem presents a challenge that strict one-one relations are unable to account for. To clarify, psychological connectedness implies a direct connection between parts. We are not identical to our immediate past or future selves, rather, we have strong relational degrees. For example, I am only continuous with a past self insofar as I remember the thoughts and beliefs of that person. I am psychologically connected with a past self insofar as a past self of mine remembered the thoughts and actions of a previous person. For example, in the *Brave Officer* case the brave officer is not psychologically continuous with his boyhood self because he cannot remember those previous actions. However, he is psychologically connected because as a young man he could remember the pertinent details of his boyhood and he can, as an old man, remember the thoughts and actions of himself as a young man. ‘Continuity’ requires a one-to-one relation whereas ‘connectedness’ requires only an overlapping of memory.

The above passage is crucial in understanding how this can aid us in thinking about the relations between two entities with the same consciousness. Parfit concedes that psychological-connectedness, because of its flexibility, is not as ‘important’ as psychological-continuity.

However, it can help us establish relations of degrees. This can be thought of as an acceptable middle ground between psychological-continuity and no psychological connection at all.

Further, Parfit suggests that we think of ‘past selves’ and ‘future selves’ as this allows a being to persist through time without necessarily implying identity. Therefore, if a split were to occur, we could feasibly refer to the two resulting selves as ‘future selves’ which are related to a single ‘past self’. This would a) allow the original self to survive as a past self and b) allow two future selves to persist without implying the notion of identity to each other: “Since psychological connectedness is not transitive, and is a matter of degree, the relations ‘being a past self of’ and ‘being a future self of’ should themselves be treated as relations of degree” (qtd. in Perry 214). As a result we now have a way to trace the self through these branches of separation.

In sum, Parfit has offered us a feasible resolution regarding the issue of the branching problem. Instead of holding onto the notion of tracing identity through the various splits, we begin to see a more promising account emerge if we acknowledge the possibility that ‘connectedness’ can too imply identity, as opposed to solely ‘continuity’. Instead, if we vie for a view that champions relational degrees rather than strict identity, we are able to incorporate the notion of branching under Locke’s general framework without running into the obstacles that arose because of such stringency. He successfully avoids using identity as the ideal by suggesting that we think of additional branches from ourselves as ‘future selves’ and branches from which we might arise as ‘past selves.’ To reiterate, we are not identical to our immediate past or future selves, rather, we have strong relational degrees. This can be held in opposition to one that is several branches away from their original self. In this case, they would have a relatively weak relational degree.

## 2 The Fickle Nature of Memory

In this section, my aim is to further explore the issue that the Brave Officer example raises and ultimately suggest a few possible avenues in an effort to move forward with a memory-based theory of personal identity. I will begin with a brief refresher of the issue at hand: There are three time periods, T1, T2, and T3 (sorted chronologically) each of which correspond to a certain Person A, B, and C. Now, imagine Person A at T1 to be a 10 year-old boy that steals some candy or commits a petty act. In 15 years, this boy is now Person B at T2 who happens to be an officer in the military and wins a major battle. This is the ‘brave officer’. At this point (T2), Person B remembers the act that he committed as Person A at T1. From T2, fast forward to Person C at T3, 40 years later. Person C is now 65 years old and a retired general. Person C can distinctly remember being 25 and winning the major battle, however, at this point he has lost the memory from when he was 10. Now, according to Locke, Person C is identical to Person B because he can recall winning the battle and Person B is identical to Person A because Person B can remember stealing the candy. However, Person C and Person A have no memory-based connection. This essentially breaks the notion of transitivity and exposes a contradiction:  $A = B, B = C, A \neq C$ . Though much has been said about the Brave Officer, I would like to take the general framework and apply it to what I deem to be a much more current and relevant topic.

## 3 Severe Memory Loss

One of the most pressing matters in bioethics surrounds the concept of severe memory loss, specifically concerning diseases such as dementia generally, and Alzheimer’s specifically.

In terms of the widespread nature of this disease, the World Health Organization states that 50 million people globally suffer from dementia and that there are 10 million new cases every year. Additionally, the WHO claims that Dementia is, “one of the major causes of disability and dependency among older people worldwide” (WHO, “Dementia”). The NIH defines dementia as follows:

“Dementia is the loss of cognitive functioning—thinking, remembering, and reasoning—and behavioral abilities to such an extent that it interferes with a person's daily life and activities. These functions include memory, language skills, visual perception, problem solving, self-management, and the ability to focus and pay attention. Some people with dementia cannot control their emotions, and their personalities may change. Dementia ranges in severity from the mildest stage, when it is just beginning to affect a person's functioning, to the most severe stage, when the person must depend completely on others for basic activities of living” (NIA, “Dementia Symptoms and Diagnosis”)

Those that suffer from Dementia are at an increased risk of suffering from severe memory loss. Personally, both of my grandmothers suffered from dementia and would often wake up ready to work at the citrus packing houses where they were once employed. The fact of the matter was that they hadn't worked at said establishments for almost 50 years. They also tragically began to forget the names of their children and were enveloped in a perpetual mental fog. Because of the increasing prevalence of such diseases, it is important to understand how this might affect our sense of self so that we as a society have a basis from which to make healthcare-related decisions.

Were we to simply jump into a conversation regarding the ramifications of Locke's theory concerning those that suffer from severe memory loss, we might run the risk of losing much of the nuance that the discussion requires. One might simply think, “Well, according to Locke's theory, anyone that has significant memory loss and cannot remember broad swaths of their past experience would thereby cease to be their previous self.” They might in this case be a

new person. It is worrisome to imagine a case in which Person B remembers Person A, but fails to remember the various stages in between. Does a corrupted memory necessarily spell a fragmented individual? This is a hugely important question and one that I believe we are currently ill equipped to answer. It may be advantageous to think through a specific issue before tackling such a broad question.

### 3.1 Advance Euthanasia Directive

One topic that has garnered much attention in recent years has been that of advance directives. Eight states in the United States have what are referred to as ‘Death with Dignity Laws’ which “allow mentally competent adult state residents who have a terminal illness with a confirmed prognosis of having 6 or fewer months to live to voluntarily request and receive a prescription medication to hasten their inevitable, imminent death” (Death With Dignity, “Death with Dignity Acts). The process is certainly robust and requires confirmation from two separate physicians regarding mental competence, residency, diagnosis, prognosis and the voluntariness of request. The ethics behind such treatment are wrought with debate, however, that is not the objective of its mention in this paper. In fact, I would like to further narrow the focus to a specific type of physician-assisted euthanasia: advance euthanasia directive.

Dutch law allows for the use of what are called advance euthanasia directives. These are essentially authorizations given in advance by patients whose condition has worsened to a point at which they would be able to request physician-assisted suicide according to their physical condition and outlook, they might not be able to due to ‘mental incompetence’. This essentially allows them to state in advance, “I, of sound body and mind, wish to be euthanized once my

symptoms progress past a certain threshold.” In fact, AEDs were developed specifically for patients that have been diagnosed with advanced dementia for his very reason. In a recent study conducted by several Dutch researchers, it was found that most healthcare professionals fear that allowing for physician-assisted suicide in patients with advanced dementia runs the risk of not adhering to the criteria of due care because the patient in question may not be in a mental state fit enough to make such a decision (Kouwenhoven et al.) Therefore, AEDs may provide a potential solution for those that do not wish to suffer the later stages of Dementia whilst adhering to healthcare guidelines. In the United States, AEDs are not universally available yet.

Now that a basis for AEDs has been established, we may consider one case that spells trouble for AEDs and it goes as follows: Imagine that there is a patient that has been diagnosed with advanced dementia and the disease is considered terminal. In wishing to avoid as much suffering as possible, the patient opts into an AED which would guarantee physician-assisted suicide once their dementia progresses past a certain threshold. A significant amount of time lapses and the patient has begun to experience advanced symptoms of the disease, the most noticeable being memory loss. The memory loss is significant, they no longer remember that they are suffering from Dementia, they often think they hold occupations that they haven't held in years, and they even begin to forget the names and faces of family members. The patient's doctors deem that the disease has now progressed past the specified 'suffering' threshold and it is an appropriate time to perform the physician-assisted suicide. As is procedure in this particular hospital, they would give the patient a week's notice before performing the euthanasia. Upon receiving this news, however, the patient is thrown into a frenzy. They can no longer remember requesting the AED and the notion that they will be euthanized in a week's time is incredibly troubling and as a result they request to opt out of their AED.

This case complicates matters simply because it is a case of one person's testimony over the other. There was Person A at Time A, when the AED was signed and there is now Person B at Time B that now wants to opt out of their scheduled euthanasia. However, the troubling detail is that there is little connection between Person A and Person B. Person A never knew Person B but implicitly when making the decision to sign the AED, they were extrapolating their future desire in making the decision. Additionally, we know that Person B has no connection with Person A because they can no longer remember making the decision to sign the AED in the first place. Our patient, therefore, can be thought of as two separate people under Locke's theory. The question remains regarding how to proceed.

Intuitively, it seems morally dubious to value the word of a past self over that of a current self even if the current self's mental state has been compromised. I would posit that, even though Person A had a measurably higher level of mental competence, this does not necessarily give them the power to preside over their future self if their future self desires different outcomes. In fact, I believe that because Person B does not remember Person A, we should think about them as different people for the purposes of this case. This is because if person B does not remember person A and the decision that they made, there is no sense of deference that they would feel indebted to. In a sense, person B would be acting as a separate agent completely severed from the decisions of a past self. Therefore, they should be free to act according to whatever they feel most comfortable with.

Critics might be concerned about the notion that Person A's decision has been completely discarded and in essence their final wishes have failed to be fulfilled. While I recognize these real concerns, I believe that this is the most plausible way to go about handling this specific case because the ethical dilemma of renegeing on the wishes of a previous person or self is

significantly less than euthanizing someone based on the wishes of a self that arguably no longer exists.

Ultimately, we will see that the literature on AEDs is extremely varied and will undergird any final decision on the matter.

### *Chapter Three*

Over the course of this final chapter, I review much of the popular literature regarding the issue of the self and AEDs. Ultimately, my goal is to conduct a thorough investigation regarding the potential options of those who have signed AEDs but encounter conflict when the time comes to exact the AED. As we will see, the issue is twofold. On the one hand we must make a decision regarding the effectiveness of AEDs in making end-of-life decisions, however, that goal may remain out of reach without a solidified theory of self.

#### 1. Centering the Issue

In chapter two, I began by addressing two significant issues that arise when one considers a memory-based theory of personal identity. These would be the branching problem and advanced euthanasia directives (AEDs) respectively. However, in moving forward, I will focus primarily on AEDs as this seems to be a more pertinent issue with regards real-world concerns. In analyzing AEDs we saw that there is an inherent dilemma in planning death in advance, especially when dealing with a disease such as dementia. Because of the way in which dementia naturally degrades memory and various mental faculties, this forces us to think of a single person in different instances across time as different people. That is to say, we should rationally think of a person in the early stages of dementia as one person and someone in the late stages of dementia as a different person. Additionally, we should think of the various stages in between as having some level of autonomy as well. The notion of autonomy is one that I will discuss more in-depth as much of the literature suggests a notion of autonomy that is both problematic and fails to

consider the well-being of the patient in question. Over the course of this chapter I will critically analyze AEDs within the relevant literature, assess the viability of such a directive, and finally decide whether it is safe enough to use in practice.

At the heart of the argument for and against the use of AEDs is the well-being of the person in question. At first glance, this certainly seems like a straightforward issue, however, there are two aspects that make this a complex quandary. The complications arise when the terms ‘well-being’ and ‘person’ are called into question. We must ask what exactly is meant by well-being. In various papers on this topic, it has been taken to mean either one of two things: well-being as beneficence or well-being as in the autonomy of an individual. Well-being as beneficence centers around mercy; whatever is merciful to the patient is that which guarantees the most well-being. Autonomy concerns the wishes of the person with early stage dementia over those of the same being with later stage dementia, given that the earlier person agreed to an AED. It is the idea that, because the person that agreed to the AED did so without significant mental impairment, their decision should take precedence over the same being with later stage dementia. In *Dementia and Dignity: Towards a New Method of Surrogate Decision Making*, a source I will analyze later, Elysa Koppelman provides a very compelling definition of autonomy suited to this context:

“It refers to the capacity to choose identity-defining values and beliefs and to base choices on them. Autonomy refers to one's ability (and right) to express, confirm, or create a self (determine who one is) through choice and action. Being autonomous, then, presupposes the idea of an authentic self: one who sees and embraces herself as subject or agent and who controls the creation and expression of herself.” (Koppelman, “Death and Dignity”)

Though the philosophers who have written on the subject of AEDs diverge in many respects, they all operate on the implicit assumption that there exist two different people, person A at time

A and person B at time B and also that because of the disease person A cannot exist while person B does. In this sense, because person A made their decision without the mental degradation that dementia entails, not only should their previous decision prevail. But it should do so even in cases in which person B objects to the AED. It is also important to remember that because of this specific case regarding dementia, there is the distinct possibility that person B does not even remember agreeing to the AED in the first place. According to a memory-based theory of personal identity, they would be considered two different people and if the AED were to be exacted against person B, it would be as if they were killed by the wishes of person A, a person that, because they have no recollection, can be thought of as a different person all together.

Those who endorse a notion of beneficence believe that in some cases while one is in the midst of battling dementia they suffer a great amount. Therefore, if one were to equate well-being with beneficence, it would fall in line that to euthanize the patient would be in their best interest for cases in which the suffering was bad enough. However, it is dubious to assume that all dementia patients suffer. And, while it may be a source of pride for person A to maintain a certain level of cognition throughout their life, and perhaps, to end their life once they no longer have that level of cognition, in essence a new and distinct person inhabits their body and person A ceases to exist.

## 2. The Irrational Weight of the Past

Intuitively, we all want to control what happens in the future, at least to a certain extent. It can be said that the uncertainty associated with what the future holds has always been a perennial source of anxiety. With that being said, it may help to think of AEDs as a means of attempting to quell such anxiety. Regarding dementia, it is almost as if at a certain, ambiguous point in the

future, the patient ceases to be ‘themselves’ and ‘another’ inhabits their body. One fails to remember the essence of who they are and inevitably becomes a different version of who they once were in the eyes of friends and loved ones. However, when considering AEDs there is one undeniable truth: though there may persist a different person in lieu of the one that once was, they are still a person. And, though they do not maintain the same cognitive capabilities and memories as they once did, they still maintain a certain level of autonomy insofar as they are able to think and communicate. Jennifer Hawkins summarizes the discomfort in the idea of essentially executing someone following strict AED procedure:

“It strikes many as extremely bizarre to say that it is best for a person to die when that same individual is enjoying her life and claims not to want to die. Such intuitions are not decisive alone, but then, in such a case, neither is the fact that in the past she formed an autonomous desire to avoid dementia.” (Hawkins, 514)

If a) the patient in question is able to think and communicate and b) is either afraid to die or does not feel that they should die then it follows that they should receive adequate treatment and that any resulting death without proper precaution should be viewed as untimely. To exact an AED on such a person would be morally reprehensible and would only be fulfilling insofar as it achieves the wishes of a person that is essentially dead, assuming that the dementia has progressed to a point at which there is significant memory loss.

Now, to some it may seem as though this assessment is uncharitable. Some might argue that AEDs are indeed fulfilling to the extent that they achieve the wishes of a mentally competent person before the effects of an insidious disease takes place. Many might also argue that AEDs are the only true means by which such people can maintain dignity through this disease in the eyes of their family and friends. To clarify, I am in no way attempting to refute any of these claims as some may be true. However, the main point that I would like to stress is that any

patient, even those that have severe dementia are themselves people. Albeit, perhaps a different one from they who requested the AED. Therefore, once we accept the late-stage patient as a person who is fundamentally different from the one that existed pre-AED, we must also grant them autonomy as well. The issue with the notion of autonomy as it is discussed in the literature is that it is associated with pre-dementia cognitive capabilities. However, if we rely too heavily on this distinction then we run the risk of dehumanizing the patient. Though these patients may not have the same cognitive capabilities as they once did, so long as they are not in a vegetative state they should be granted a certain level of freedom when it comes to decisions regarding their health and outcomes.

There is also another point I would like to address as it is often found in the literature on this subject and it is called the “life-object approach”. Hawkins summarizes it as follows:

Significantly, a life can be viewed as a whole in two different senses. In one unproblematic sense, a life is a collection of events that occur over time. The “whole” life is simply the sum of these events. And if we limit ourselves to this perspective when thinking about the value of the life, we will naturally think of the value of the whole as the sum of the values of the parts. However, we can also look at a life in a way that pays close attention to properties of the whole such as shape, trajectory, narrative structure, or thematic unity.<sup>16</sup> If we allow that such properties add value to a life, then we may be tempted to say that the value of the life as a whole is not simply the sum of the value of the parts—since this further kind of value is not ‘contained in’ or ‘associated with’ any single part but is rather a property of the whole.

When philosophers adopt this kind of perspective on a life, analogies with aesthetic value are often employed, although importantly they are intended merely as analogies. For example, many philosophers have found it helpful to talk about the narrative dimensions of a life. Dworkin appears to focus more on thematic unity—on the idea that the parts of a life should harmonize with one another as opposed to the idea that a life should have a certain kind of plot. But whether one thinks in terms of narrative or theme, the approach assumes that early parts of a life can affect what counts as a good choice later on in the life. That is why, on this view, past desires matter. (Hawkins, 519)

I believe that this view, while perhaps appealing at first, fails to add nuance to the picture we’re trying to paint and rather muddies the water. It is at once appealing to think of a life as a narrative arc, one in which events that occur later were influenced by those that happened in the

past. Or that a life is a series of actions and reactions which occur and are influenced by each other. On the whole, I would agree with this. However, that is not what we're talking about here and this is where I believe a memory-based theory of self similar to Locke's can help us understand what is truly at stake in this situation. When talking about dementia, it would be dubious to speak of a life in terms of one narrative arc as is Dworkin's argument. In my estimation, the correct way to think about such a life would be one that begins with a singular narrative arc but overtime as the disease progresses, said arc eventually dissolves and as the original person fades due to memory loss a new discontinuous arc replaces it. This new arc can sometimes be informed by the old, but it is a corrupted version at best and therefore should not be tethered strictly to the old. This arc should be viewed as distinct and able to express original or dissenting views relative to the old, most importantly as to whether they wish to live or die.

### 3. Four philosophers discuss AEDs

Simply iterating arguments from several different philosophers can confuse their stances on our central inquiry. In an effort to quell such confusion, I will formulate a practical example concerning an AED, a case that could mirror one we might find in reality, and extrapolate how each of the respective philosophers that we've considered would suggest we proceed.

Let us suppose there is a person, Ted, and he is diagnosed with dementia. Ted is in the early stages of dementia and as a result his symptoms are mild in nature. He sometimes forgets where he put his keys, forgets what day of the week it is, and struggles to keep score when he plays tennis. He still recognizes and knows all the names of his family members and fundamental facts about his life. The essence of his person is still intact. However, Ted is very concerned

about this diagnosis, he has seen the effects of late-stage dementia and wants to ensure he lives out his life on his own terms. Ted resides in a country in which advance euthanasia directives are available to those diagnosed with dementia and is interested in signing one. One of the primary requirements for being able to authorize an advance euthanasia directive, or AED, is that the signee must be of sound mind at the time of signing. Ted passes the required test, learns more about what is entailed, and signs. The AED Ted signed stipulates that should Ted's dementia progress pass a certain threshold and should he contract any form of illness or suffer any type of accident, no potentially life-saving medical treatment will be administered. If Ted survives, then he will continue as per usual. However, even if it becomes apparent that Ted's condition will ultimately kill him and is treatable, doctors are to do nothing.

Fast forward five years: Ted is still alive, though the dementia has progressed to a stage at which his doctor has chosen to 'activate' the parameters of the AED. In fact, the disease has progressed so much that the AED is now 'active' and should Ted contract an illness he will not receive life-saving treatment. A few months go by and it becomes apparent that Ted is very happy. He lives in a nursing home with other elderly patients. He's talkative, active in the nursing home's card game scene, and loves visits from his children and grandchildren, though he consistently mistakes them for others. Then, Ted contracts pneumonia. Thanks to modern medicine, this is an easily treatable illness, however, if left untreated it can often become fatal. Ted is obviously unwell from the pneumonia and suffering. He is bed-ridden and has expressed several times to the nurses that he loves his life at the home and is afraid of dying. John's doctor reminds him of the AED he signed but because of his dementia, he can't remember signing it and states he'd never sign an AED. He wants the AED nullified and to receive the life-saving treatment. How should the situation proceed?

### 3.1 Dworkin

According to Ronald Dworkin's view, human desires fall into one of two categories: critical or experiential. Critical desires are outcomes that someone wants to realize because of desires that arise out of what someone deems crucial to their well-being. Experiential interests are only desired insofar as they would be pleasant were they to occur. He assumes that critical interests matter more than experiential ones, that they have a greater value. Dworkin might situate the narrative as a conflict between the desire to die (past desire) and a new desire to live (new desire). He would categorize the older desire as critical and the newer desire as experiential, therefore the older desire is the one that should be fulfilled. Dworkin would also likely acknowledge the dubious nature of weighing the desires of the past against those of the present, however, he would say that the person at the time that they authorized the AED had more autonomy than the person we are now considering and that therefore, we ought to respect the integrity of the greater autonomous decision.

### 3.2 Hawkins

If tasked with determining whether Ted should receive potentially life-saving treatment or be left to die, Jennifer Hawkins would say that Ted should indeed receive such treatment. While Hawkins acknowledges that autonomy should be valued somewhat and that the desire to have a cohesive narrative structure in life is valid, the notion that these should outweigh current desires is wrong. Hawkins' central question is which would provide more well-being for the person in question and letting Ted die after he has explicitly expressed the wish to carry on is not

in his best interest. Therefore, Ted should be able to opt out of his AED and receive treatment.

She sums it up quite well, “Past desires that a person no longer has are not in any way determinative of current good.” (Hawkins, 541)

### 3.3 Koppelman

Elysa Koppelman, in response to our inquiry, would likely offer a hybrid between Hawkins and Dworkin. In her work, *Dementia and Dignity: Towards a New Method of Surrogate Decision Making*, she vies for a middle ground that veers slightly toward the ideas championed by Dworkin. She states, “Autonomous decisions, I believe, are decisions that reflect the self which makes them.” (Koppelman, “Dementia and Dignity”) Ultimately, she would like to modify Dworkin’s ‘character view’ in order to further define what we mean by autonomy and how we might go about truly achieving it through surrogates. Additionally, central to Koppelman’s thesis is the idea of dignity which she takes to mean others’ perception of us. Specifically, that their perception will fall in line with how we want them to perceive us. We need people to respect our autonomy in order to gain dignity. “Decisions regarding the care of the AD patients will either be consistent or inconsistent with her conception of the good life, rational life plan, with her character. Only a decision that is consistent with these things, one that is reflective of the self which is most affected, is a good decision” (Koppelman, “Dementia and Dignity”). In this sense, a person acting in the interests of a person's dignity and therefore their autonomy would be a suitable surrogate. Concerning this specific case, Koppelman would likely state that there is no way to achieve the best-case scenario without this surrogate decision-maker. She would likely not choose any of the available avenues currently present.

### 3.4 Locke

Tying the case of Ted back in with our foundational theory of self from John Locke, the former might claim that according to his theory, Ted at time B is numerically different from Ted at time A. Locke would argue that because Ted at time B has suffered significant memory loss, he is fundamentally different than he was. Locke would also likely focus on the fact that Ted at time B could not remember authorizing the AED. Because Ted at time A and Ted at time B are categorically different, Locke would likely say that the decisions of Ted at time A should have no bearing on the outcomes of Ted at time B. To say that Ted at time B should die because of something that Ted at time A decided would be ludicrous and wrong. It would be akin to putting someone to death at the word of another. Undoubtedly, Locke would disapprove of Ted's death and would support him receiving the life-saving treatment.

### 4 Altered AED?

Though we have seen how AEDs can be problematic when it comes to patients with dementia, the wish to die on one's own terms is one that should be honored if possible. In trying to exact such wishes, perhaps there would be a way to alter the structure of AEDs to accommodate such wishes while avoiding the issues previously discussed.

One such alteration was discussed earlier in the paper wherein, an AED would be signed by the dementia patient before the disease truly took hold and doctors would monitor them closely. If they ever reached a point where they were experiencing overt suffering or there was any indication that the conditions for euthanasia were met, the patient would essentially be asked

whether or not they still wanted to go through with the AED. In this case they would be reminded of the option of euthanasia and they would give a consenting order that it be carried out. For this option, however, we must ask whether someone who is enveloped by dementia truly has the ability and wherewithal to truly consent. Many definitions merely list consent as ‘giving permission’ or something of that nature, however, we well know that there is significant nuance when it comes to what some might call ‘true consent.’

No doctor should ask a patient that lacks mental competency whether they would abide by a previous AED because the patient may not be fully aware of what is going on. However, if they do express the wish to not die then this should certainly be honored. This is simply due to the fact that the stakes in the former example are far higher than those in the latter.

Given what we now know about memory and its relation to personal identity, we cannot respect the stipulations of AEDs. To do so would risk harming dementia patients and subjecting them to the desires of a distinct past self. Now, there are several concerning implications that arise from this conclusion, one of which being the question of property rights. If we cannot respect AEDs because the person in question is essentially different from they who authorized the AED, then wouldn't they no longer lay claim to their home or possessions? If person A is different from person B, then wouldn't person B not be entitled to the possessions of person A? I believe that the answer to this question lies in the fact that decisions about life and death are different from those concerning possession. It would be bizarre to state that the person who paid for and lived in a home for fifty years would not be entitled to continue living in it because of memory loss. Therefore, I suggest that the implications of memory-loss as it pertains to decision-making only apply to life and death decisions. It would not be permissible to take away someone's possessions on this basis.

## 5 In Conclusion

I began this paper with the goal of analyzing various theories of self in an effort to solve the issue that AEDs present. Ultimately, my analysis demonstrates that John Locke's memory-based theory of self is the most promising account when it comes to locating where the self resides. As we have seen, it is not in an immaterial substance such as a soul and it is not in our physical manifestation. Rather, the self resides in memory and is therefore informed by the experiences we have. By extension, this fact has ramifications regarding how we think about persons with severe memory loss. As we've seen, dementia is one such case that is common in older individuals and entails substantial memory loss. Under Locke's view, such memory loss would constitute a change in person and would therefore heavily influence the ways in which we should think about such persons and their end-of-life treatment. Specifically, I chose to focus on AEDs because of their potential in exacting final wishes of those diagnosed with dementia.

Having discussed the issues regarding AEDs at length, it seems as though they cannot be implemented as they currently stand without posing a serious risk to the well-being of the patient involved. Because of the nature of dementia and how it augments the patient's ability to think and make decisions, signing a contract to either euthanize them or let them die in the case of a treatable illness would risk killing said patient unduly. As we have seen, Locke's theory concerning memory-based personal identity is crucial in showing how a being before and after dementia should be thought of as two different people and therefore should be treated as such. This would mean that the wishes of a past person could not rationally extend to dictate the course of action or fate of a distinct future self and that we should not honor the binding nature of AEDs.

## Works Cited

David Wiggins, *Identity and Spatio-Temporal Continuity* (Oxford, 1967), p. 50

“Dementia.” *World Health Organization*, World Health Organization, [www.who.int/news-room/fact-sheets/detail/dementia#:~:text=Worldwide%2C%20around%2050%20million%20people,dependency%20among%20older%20people%20worldwide](http://www.who.int/news-room/fact-sheets/detail/dementia#:~:text=Worldwide%2C%20around%2050%20million%20people,dependency%20among%20older%20people%20worldwide). Hawkins, Jennifer. “Well-Being, Time, and Dementia.” *Ethics*, vol. 124, no. 3, 2014, pp. 507–542. *JSTOR*, [www.jstor.org/stable/10.1086/675365](http://www.jstor.org/stable/10.1086/675365). Accessed 24 Mar. 2021.

Koppelman ER. Dementia and dignity: towards a new method of surrogate decision making. *J Med Philos.* 2002 Feb;27(1):65-85. doi: 10.1076/jmep.27.1.65.2971. PMID: 11961687.

Kouwenhoven, P.S., Raijmakers, N.J., van Delden, J.J. *et al.* Opinions about euthanasia and advanced dementia: a qualitative study among Dutch physicians and members of the general public. *BMC Med Ethics* 16, 7 (2015). <https://doi.org/10.1186/1472-6939-16-7>

Locke, John (1690). Of identity and diversity (book II, chapter XXVII). In *An Essay Concerning Human Understanding*.

National Institute on Aging. “What Is Dementia? Symptoms, Types, and Diagnosis.” *National Institute on Aging*, U.S. Department of Health and Human Services, [www.nia.nih.gov/health/what-dementia-symptoms-types-and-diagnosis](http://www.nia.nih.gov/health/what-dementia-symptoms-types-and-diagnosis).

Perry, John (ed.) (1975). *Personal Identity*. University of California Press.

“Resources.” *Death With Dignity*, 22 Mar. 2021, [www.deathwithdignity.org/learn/](http://www.deathwithdignity.org/learn/).

World Health Organization. “Dementia.” *World Health Organization*, World Health Organization, [www.who.int/news-room/fact-sheets/detail/dementia#:~:text=Although%20dementia%20mainly%20affects%20older,60%E2%80%9370%25%20of%20cases](http://www.who.int/news-room/fact-sheets/detail/dementia#:~:text=Although%20dementia%20mainly%20affects%20older,60%E2%80%9370%25%20of%20cases).