FYS 100
Who am I? What am I?
Readings
Who is Julia?

Introduction
By Dennis Weiss

A beautiful model gets hit by a truck, completely destroying her glamorous body but miraculously sparing her brain. A plain and unremarkable housewife has a brain aneurism that makes pudding of her brain but spares her body. A brilliant and cutting-edge surgeon (pun intended) decides to make the most of the tragic situation and transplants the glamorous model’s brain into the plain-Jane’s body. Who wakes up? Is it the model or the housewife? This is the fascinating plot of Barbara Harris’s novel *Who is Julia?* Julia North is the beautiful model who, in act of courage, darts out into traffic in order to save the son of Mary Frances Beaudine, who had wandered into the street. Julia is killed, and almost simultaneously Mary Frances collapses from a brain aneurism. Following surgery, we are faced with something of a composite: the body of Mary Frances Beaudine and the brain of Julia North. So who is the resulting person? The rest of the novel explores this question. *Who is Julia?* was turned into a pretty cheesy movie-of-the-week.

Watch the first part of *Who is Julia?*

Click Here

Now that you’ve watched the first part of *Who is Julia?* how do you think it will end? How do you think it ought to end? Importantly, these are two distinct and different questions. Television writers aren’t always (maybe never?) the most philosophically astute writers and they aren’t setting out to educate but to entertain. Nonetheless, *Who is Julia?* presents us with an intriguing
philosophical thought experiment. Do you think the resulting person will be Julia North? Mary Frances Beaudine? Or maybe neither of them? Maybe the resulting person is a whole new person. But then who is she?

While the movie is melodramatic, the philosophical issues explored in the novel and movie will prove to be relevant to our exploration of self and identity. If you got a brain transplant, would you still be you?

This is not just the realm of science fiction. The PBS show NOVA has explored the implications of brain transplants and the implications for our sense of self, including cross-cultural implications, as the author, Peter Tyson explains:

One's very sense of self-hood would be at stake, Wolpe argues. In the West we tend to think of the brain as the locus of self, but culturally that is a very new idea, and it's still not shared in many cultures, he says. Consider Japan, where the locus of self is thoracic and abdominal. "That's why when you commit seppuku you disembowel yourself, you don't cut your head off, because you're attacking yourself at the seat of self-hood," [Paul Wolpe] told me.
The notion that if you put his head on someone else's body, that the resulting individual would be him and not the other person simply because the hybrid had his brain is, Wolpe says, "theory not fact, a philosophical position rather than a scientific reality. What you may end up finding is that when you transfer a brain from one body to another, the resulting organism is not solely what one would think of as the person whose brain it was but also has enormous components of the person into whose body it goes."

The issue of brain transfer and the case of Who is Julia? has been nicely investigated by philosopher Andreas Teuber in an extended thought experiment exploring the implications of brain transplants. If you are interested, you can READ IT HERE.

What do you think about the philosophical implications of brain transplants? Would you be interested in receiving a new, rejuvenated brain some day in the future? Would you preserve your identity if you received a new brain? Is who you are tied up with your brain? What if another person's brain were transplanted into your body (or someone else’s brain received a new body—which is it?)—would you still be you?
Who Am I?

By: RadioLab

Introduction

By Dennis Weiss

Radiolab is the incredibly inventive and fascinating radio show produced by WNYC and featuring Jad Abumrad and Robert Krulwich. As they describe their show on their web page, Radiolab explores the boundaries of science, philosophy, and human experience:

Radiolab believes your ears are a portal to another world. Where sound illuminates ideas, and the boundaries blur between science, philosophy, and human experience. Big questions are investigated, tinkered with, and encouraged to grow. Bring your curiosity, and we’ll feed it with possibility.

One of their very first shows was devoted to the topic of personal identity and they explored a number of stories that a student of philosophy contemplating the self and thinking about identity should find fascinating. You can listen to Radiolab’s podcasts on their web site, where they include this description of their episode devoted to the self:
The “mind” and “self” were formerly the domain of philosophers and priests. But in this hour of Radiolab, neurologists lead the charge on profound questions like “How does the brain make me?”

We stare into the mirror with Dr. Julian Keenan, reflect on the illusion of selfhood with British neurologist Paul Broks, and contemplate the evolution of consciousness with Dr. V. S. Ramachandran. Also: the story of woman who one day woke up as a completely different person.

**Listen to Radiolab podcast:** [Who Am I?](https://www.radiolab.org/podcasts/who-am-i/)
Introduction to the Reading
By Dennis Weiss

Who will you be in the future? Will you be the same person? What would it be like to hear from your “younger self”? Is that younger self you? If you got a letter from one of your past selves asking you to do something, would you be obliged to do it? Can your future self disavow your past self? These are some of the intriguing questions presented by the web site futureme.org, where you can post a letter to your future self. As the organization explains on their web site:

What the heck is this now?

Usually, it’s the future self that reflects back on the past. We think it’s fun to flip that all around.

So here at FutureMe, you write a letter to your future. And then we do a little bending of the space-time continuum and we deliver it for you.
It’s kinda cool to surprise your future self with a letter from the past. Some words of inspiration. A swift kick in the pants. Or just thoughts on where you’ll or what you’ll be up to in a year, two years...more?

FutureMe.org is based on the principle that memories are less accurate than emails. And we strive for accuracy.

So now that we are reflecting on the nature of self-identity, why don’t you take a few minutes and reflect not on your past self but your future self. And maybe you’ll be motivated to write your future self a letter. What will your future self be like? Will your future self still be you? What do you think it would be like to hear from one of your past selves?

To access the FutureMe web site, click here. You can read some public but anonymous FutureMe letters here.
Welcome to Second Life

What is Second Life?

Over the course of the semester we will be exploring different facets of self and identity. Today, many aspects of our self-identity and self-presentation are mediated by technology and one of the questions we will be exploring has to do with the impact of technology on our sense of self and the nature of our identity. From cell phones to social media, technology both shapes and refracts our sense of self, providing us with tools to control our virtual presentation of self and perhaps distort our sense of who we are.

In order to better explore and think about the manner in which technology mediates our self identity, one of the things you might like to do so this semester (as you work on being an engaged and responsible learner) is create an avatar in Second Life and spend the semester virtually interacting with the netizens of Second Life. What is Second Life? In essence, Second Life was one of the first MMORPG to debut on the web. Here’s how it is described on Wikipedia:

Second Life is an online virtual world, developed by San Francisco–based Linden Lab and launched on June 23, 2003; and which in 2014 had about 1 million regular users, according to Linden Lab, which own Second Life. In many ways, Second Life is similar to MMORPGs (Massively Multiplayer Online Role Playing Games); however, Linden Lab is
emphatic that their creation is not a game: "There is no manufactured conflict, no set objective".

The virtual world can be accessed freely via Linden Lab's own client programs, or via alternative Third Party Viewers. Second Life users (also called Residents) create virtual representations of themselves, called avatars and are able to interact with other avatars, places or objects. They can explore the world (known as the grid), meet other residents, socialize, participate in individual and group activities, build, create, shop and trade virtual property and services with one another. It is a platform that principally features 3D-based user-generated content. Second Life also has its own virtual currency, the Linden Dollar, which is exchangeable with real world currency. Second Life is intended for people aged 16 and over, with the exception of 13–15-year-old users restricted to the Second Life region of a sponsoring institution (e.g. school).

Click on the link above to go to Second Life and then click on the button to “Join Now.” Create your avatar and begin to explore Second Life!
The Case of Tookie Williams

By Dennis Weiss

Once you have had a few semesters of college under your belt, you may have the familiar experience of heading home to visit your family, maybe during the semester break or after a year away at school, and finding that somehow you have changed. Maybe you no longer feel like the same person who left for college just a couple of semesters ago. You’ve been living on your own, experiencing new things, pondering exciting and strange ideas, and now you’re back in your old room and your family members want to treat you as they always have. They don’t see that you’ve changed. You’re not the same person anymore. College can be a transformative experience, which is one of the reasons many people value their college years. And you’ve transformed. You no longer fit comfortably in the shoes of your old self. You’re new and
different. Hopefully improved. But different nonetheless.

And yet, you’re still the same self that left for school some nine months ago. Right? It’s not like when you show up at home your family looks at you and wonders who you are (well, at least I hope that hasn’t happened). They still let you in and give you back your old room and call you by the same name and still expect that you will take out the trash or mow the lawn or wash the dishes just like you did before you left for college. You’re different but you’re not.

Of course, it’s not just college that can transform us. Many of you may have had other transformative experiences that left you thinking that you’re no longer your same old self. Maybe you went through some other kind of life-altering experience. Perhaps you lost a significant amount of weight or had a religious conversion or faced a personal tragedy and came out the other side feeling like a new self.

These common experiences raise an immediate philosophical problem for us: what does it mean to be “the same self”? Consider that in your relatively short life so far, you have changed considerably. You’ve grown and matured, learned new things, and acquired new skills and abilities. Amidst all these changes you have undergone over the course of eighteen or twenty or forty years of life, what makes you the same person from one moment in life to the next? What makes the person sitting in a philosophy class today the same person that was taking their first steps many years ago, or uttering their first words, or blowing out the candles on their fourth birthday in those cheesy home videos? Is there some “thing,” a soul or spirit-entity, that makes you the same you in every moment in your life? Is it your body, despite all the changes it has gone through? Or maybe there is nothing that makes you the same person from one moment to the next and that unites all your various moments into one persisting self. Maybe there is no persisting self and you’re just a series of Kodak moments. Maybe the self is not enduring and your body has housed many different selves over time, or even many different selves at one time. As Walt Whitman observed: “I am large. I contain multitudes.”

These are the central issues philosophers explore when they turn to the question of the self and personal identity. As philosopher Eric Olson notes in an essay on personal identity in The Stanford Encyclopedia of Philosophy, a key question regarding the self is the persistence question: “Under what possible circumstances is a person who exists at one time identical with something that exists at another time?” Olson goes on to give us a good characterization of this question and its connection to another important matter: our concern for the afterlife. As he writes in his selection:

What does it take for a person to persist from one time to another—that is, for the same person to exist at different times? What sorts of adventures could you possibly survive, in the broadest sense of the word ‘possible’, and what sort of event would necessarily bring your existence to an end? What determines which past or future being is you? Suppose you point to a child in an old class photograph and say, “That’s me.” What makes you that one, rather than one of the others? What is it about the way she relates then to you as you are now that makes her you? For that matter, what makes it the case that anyone at all who existed back then is you? This is the question of personal identity over time….
Historically this question often arises out of the hope (or fear) that we might continue to exist after we die—Plato’s *Phaedo* is a famous example. Whether this could happen depends on whether biological death necessarily brings one’s existence to an end. Imagine that after your death there really will be someone, in the next world or in this one, who resembles you in certain ways. How would that being have to relate to you as you are now in order to be you, rather than someone else? What would the Higher Powers have to do to keep you in existence after your death? Or is there anything they could do? The answer to these questions depends on the answer to the Persistence Question.

People’s concerns for their disposition in the afterlife begin to suggest that the persistence question is more than simply a philosophical puzzle. For you, it may also be a very existential question concerning your survival in a future state. Bishop Joseph Butler takes up this matter in his essay on personal identity from *The Analogy of Religion*. But beyond one’s concern for the afterlife, self-identity and the persistence question can also be a real practical problem—even one that arises in the news. Consider a rather more extreme case of a young student who was fundamentally transformed by her college experience. To her mother, Katherine Ann Power was a happy, sweet, bookish child. As *Time* Magazine reported, she was a National Merit scholarship finalist, winner of a Betty Crocker Homemaker award, and the valedictorian of Marycrest, her Catholic high school in Colorado. But as a college student in the late 60’s, Power became radicalized by the Vietnam War and took up with a group of student activists. As *The New York Times* reported:

“Her senior year at college she rooms with Susan Saxe,” said Ms. Carroll. “They are so appalled at what’s going on in Vietnam that they want to do something. They decide to rob a bank and give the money away. But they don’t know anything about robbing a bank.”

According to court testimony, Ms. Power and Ms. Saxe hooked up with three ex-convicts and committed a series of crimes, including the theft of ammunition. One of the men, William Gilday, a former baseball player who acted as the lookout at the robbery, was later convicted of shooting Officer Schroeder.

Power spent 14 years on the Federal Bureau of Investigation’s most-wanted list. She went underground and changed her identity, taking the name Alice Metzinger. Once again, quoting from *Time* Magazine:

To friends and neighbors, she was mild-mannered Alice, who had moved to Oregon’s Willamette Valley 15 years ago with her infant son Jaime (she has never named the biological father). She became involved with a local meatcutter and bookkeeper, Ronley Duncan, and established herself as a valued consultant to the area’s gourmet restaurants. She trained cooks at M’s Tea & Coffee House, where she was famous for her Friday special -- black beans and rice with Martinican sauce.

But Alice was suffering from depression, and after 23 years of living underground and under
an assumed identity she ultimately decided to come forward and turn herself in. Her tale of transformation quickly sparked a media circus. Is the mild-mannered Alice, mother and restaurateur and neighbor, the same person as the bank-robbing 60’s radical? Is the bank-robbing 60’s radical the same person as the sweet and bookish child? How much transformation can one person go through and still continue to be the same person? Power’s case is made even more intriguing by questions about blame and punishment. Do we punish Alice Metzinger for the crimes of Katherine Power? In her essay, “Shame and Blame: The Self Through Time and Change,” philosopher Jennifer Radden raises this same dilemma in the case of John Demjanjuk:

Do our customary notions of shame, blame and guilt require us to adopt a particular view of the self’s singularity and invariance through time? Consider the intriguing case of John Demjanjuk, tried in Israel during 1987 and 1988 for the crimes of “Ivan the Terrible,” a concentration camp guard at Treblinka in Poland, during 1942–43. John Demjanjuk, a retired factory worker living in Cleveland, Ohio, appeared banal at his trial—old, quiet, ordinary and helpless; descriptions from survivors of Treblinka cast Ivan as monstrous in his vigorous brutality. Should John be found guilty and punished for Ivan’s crimes? This question takes us beyond any answers sought at the trial. Even if the spatio-temporal identity of the later John and earlier Ivan had been established conclusively, still the justice of punishing the later man for the earlier one’s crimes may be questioned. For a philosophical puzzle of personal identity lingers: is the later John the same person as the earlier Ivan? In cases such as this the passage of time and radical changes of character and personality seem to invite the notion that one self or person has succeeded another in the same body. If this were so, would—or should—culpability transfer undiminished from one self to another?

If you are interested, you can read more about the case of Katherine Ann Power here:

- “The Fugitive,” by Barbara Kantrowitz, Newsweek
- “The Return of the Fugitive,” by Margaret Carlson, Time
- “Alice Doesn’t Live Here Anymore,” by Pam Lambert, People

Let’s consider another fascinating case challenging our understanding of self and identity and raising what Olson has called “the persistence problem”: the case of Stanley “Tookie” Williams. In 2005 there was a firestorm of media interest in the case of Tookie Williams, a co-founder of the notorious Los Angeles gang the Crips, who was on death row, was soon to die, and was seeking a commutation of his death sentence from then California governor Arnold Schwarzenegger. The Los Angeles Times reported that the case provoked more controversy than any California execution in a generation, and became a magnet for media attention worldwide. The case generated so much attention in part because it was framed as a story of redemption and personal transformation:

“He went through a personal metamorphosis 15 years ago and since then has reached out to young people,” said Peter Fleming, the New York lawyer who is leading Williams’ clemency team. “He has made extraordinary efforts to reach out to young people to dissuade them from engaging in the kind of [gang] activity he did,” Fleming said in an
interview.

Williams was found guilty of brutally murdering four individuals, though Williams maintained his innocence. As the *LA Times* reported:

By the 1970’s, Williams was viewed as one of the more menacing toughs in South Los Angeles, weighing 300 pounds with biceps measuring 22 inches. In a move he said he regretted more than any other, he helped launch the Crips -- originally called the Cribs -- and began terrorizing the streets.

On Feb. 27, 1979, he and three cohorts smoked cigarettes laced with PCP and, armed with a 12-gauge shotgun and a .22-caliber handgun, set out on a late-night search for a place to rob, according to court documents. They wound up at the 7-Eleven where Owens, a father of two and Army veteran, was working the overnight shift. Owens was shot twice in the back. Less than two weeks later, Williams broke down the door at the Brookhaven Motel and killed the motel’s owners, Taiwanese immigrants Yen-I Yang, his wife, Tsai-Shai Chen Yang, and their daughter, Yu Chin Yang Lin, who was visiting. The two robberies netted $220.

But while in solitary, Williams began a transformation. Again, quoting from the *LA Times*:

While in solitary confinement, however, he began a transformation, Williams said. At first he read voraciously -- the Bible, the dictionary, philosophy, black history -- and struggled to understand his past.

By 1992, Williams was a changed man, he said, deeply remorseful for the bloody rampage the Crips had perpetrated across America -- and for the gang life that lured in one of his two sons.

In 1994, Williams left solitary confinement and declared himself a champion of peace. With the help of Becnel [a journalist and film producer who befriended Williams], he wrote a series of books warning youths away from violence and brokered gang truces in Los Angeles and New Jersey. Last year [2004], his life became the subject of a TV movie, “Redemption,” starring Jamie Foxx, and his imposing appearance gradually gave way to a graying beard and spectacles.

Ultimately, Williams became an advocate speaking out against and writing about gangs and gang violence. He renounced his gang past, started the Internet Project for Street Peace, wrote a series of children’s books subtitled “Tookie Speaks Out Against Gang Violence,” and was even nominated for the Nobel Peace Prize. As Williams said of his own case: “I have a despicable background.... I was a criminal. I was a co-founder of the Crips. I was a nihilist. But people forget that redemption is tailor-made for the wretched.” Governor Schwarzenegger ultimately decided against clemency and Tookie Williams was executed on December 13, 2005.

While the case of Tookie Williams raises a number of interesting philosophical questions,
including questions related to the death penalty, our focus is on the question of self-identity and the persistence question. Is it possible to be both the founder of the Crips and a nominee for the Nobel Peace Prize and yet be the same person, the same self? That is the relevant question posed by this complex case and the issue that will be central to your writing assignment.

To learn more about Stanley “Tookie” Williams, check out some of the following links:

- An interview with Williams conducted by Democracy Now
- Wikipedia’s entry on Stanley Tookie Williams
- Watch the trailer for F/X’s movie Redemption: The Stan Tookie Williams Story

As you may have already noticed, philosophers love hypothetical situations and thought experiments. In fact, over the course of the semester you will encounter a number of famous thought experiments, one going all the way back to John Locke’s 1690 Essay Concerning Human Understanding, where he imagines the consciousness of a prince changing places with the consciousness of a pauper. So thought experiments have a long and distinguished history in philosophy and especially in philosophical thinking about self and identity. So let’s engage in a little thought experiment of our own, building on the facts of the case of Stanley “Tookie” Williams.

Let’s imagine that you work as an aid for California Governor Arnold Schwarzenegger and he has convened a clemency hearing to discuss a petition from Williams’ attorneys for clemency and a stay of execution. Governor Schwarzenegger had previously stated that the decision to grant clemency was “the toughest thing when you are governor, dealing with someone’s life,” and the meeting has been contentious, as crowds consisting of both supporters of Williams and opponents of capital punishment have gathered outside the closed-door meeting. The Governor decides to call for a break in the discussions and you and various other aides are sitting around discussing some of the difficult issues the case raises. You mention that you were a student of philosophy in college and recall some interesting essays you read and discussed on the matter of personal identity. Governor Schwarzenegger, being an avid fan of all things philosophical, leans in to listen as you begin to discuss the persistence question and the issue of whether the 52-year old Williams sitting on death row in 2005 is the same person as the young 16-year old Williams who started the notorious Crips in South Central Los Angeles. Schwarzenegger listens in as you recall some of the details of theories espoused by Locke and Reid and Hume and other philosophers who debated the persistence question. Intrigued by your discussion, the Governor turns to you and asks in his inimitable accent, “So, what do you think?” Not one to shy away from a philosophical puzzle, especially one posed by the Governator, you organize your thoughts and begin to address this complex question. How do you reply? What do you think?
A Philosophical Identity Crisis

by Chris Durante, The Philosopher’s Magazine

Introduction to the Reading
By Dennis Weiss

What happens when you transplant a brain? What are the implications of two persons fusing? Can a person’s character change so fundamentally that they are no longer the same person? What does it even mean to be the “same person?” Through a series of puzzle cases and thought experiments, we have begun exploring the nature of the self and the questions and problems associated with personal identity. Most of you probably have an intuitive and common sense view of the self. But when that common sense view is confronted with these puzzle cases and thought experiments, we often find ourselves stumped and we may not know how to proceed. So what do we do now? One possible avenue to explore is to deepen our understanding of the
nature of the self and the problem of personal identity. Let’s begin to theorize. This is what philosophers do when confronted with puzzles that stump them.

As Chris Durante explains in this introductory account of distinct identity theories, philosophers have identified a number of competing ways to address these cases of brain transfers, fused persons, and psychological changes. As you read this essay, written for the popular magazine *The Philosopher’s Magazine* (check out its online site [HERE](http://example.com)), let’s practice our skills of identifying key information and analyzing important claims.

- Durante points out that philosophers have offered two major accounts of personal identity. Can you identify those two accounts and their key features?
- Thomas Reid identifies a weakness in one of those theories with his “Paradox of the Brave Officer.” What makes this a “paradox”?
- Proponents of the psychological theory of personal identity think they have a way of responding to the brave officer paradox. How?
- Durante seems to prefer an alternative to these two major accounts of personal identity. He names this alternative the “narrative identity theory.” What are the key features of this theory?

Now that we have some theories to work with, let’s return to some of our thought experiments: the case of Julia or Tookie Williams or Tuvix. Did you implicitly rely on some of these theories in your initial analysis of these cases? What does each theory have to say about each of these puzzle cases? As we work through thinking about these theories and these puzzle cases, we might begin to formulate an account of which theory seems to work best. Are you drawn to one of these theories over another?

A Philosophical Identity Crisis

Chris Durante asks himself just what makes him the person he used to be.

Stepping into a park I had frequented as a little boy, memories of my childhood began to flood my mind, each one a rich story of a distant past. As I continued to reminisce, each story flowed into the next, and I began to witness the development of an intricate character whom I refer to as ‘me’. All these stories that I had authored in my experience flowed together to give me a unique history. Yet reflecting on all the experiences, goals, traits, and values that I’ve had, it dawned on me that my identity seemed more elusive than one might usually believe. Ruminating over these strands of my past, at times it was as if I could watch my traits develop, values evolve, my goals be accomplished and recreated; but other moments I recollected appeared in my mind as if they were foreign elements in my mental landscape. Some of the stories seemed to be integral aspects
of who I am, while other memories seemed very distant, almost as if the main character was a different person.

Contemplating further, I began to wonder if there was more to my identity than common sense or intuition could account for. What struck me was the fact that I considered myself to be a single person with a single identity, yet viewing myself as always having been ‘me’ left something unresolved. The little boy, who shares my name and appears in my stories, seems to be so different from the person I am today, yet I tend to incorporate him into my identity as a single person. What is it exactly that makes me a single human person persisting through time with a single identity? Could it be my body – that I am and have been a single biological organism? Or, is it my mind – that my psychological states interconnect so that they constitute a single continuum? I also began to wonder, “At what point in my life did I begin to be a ‘person’? When did I attain ‘personhood’?” This got me thinking about a whole new series of questions: “Is that little boy truly the same person as I am today?”; “If I became severely demented, could I still be considered to be the same person as I was before?” Suddenly my pondering had led me to very serious metaphysical and philosophical problems. A dark storm of confusion and lightning-quick thoughts set in, only to give rise to a spectacular rainbow of insight in my psychic sky.

**Distinct Identity Theories**

We usually intuitively believe that our identities remain constant over long periods of time. We acknowledge changes in character traits, etc., yet maintain a belief in the singularity of people's ‘actual’ identities. If your good friend Greg were to claim that he was not the same person he was five years ago, we would not usually assume that Greg was now a numerically distinct person, we would take it as a figure of speech denoting that Greg has undergone some major event in his life, or that he has undergone some drastic change in his personality traits. Yet when asked “Just what is it that makes a person persist as the same person over time?” can we really say what it is that gives human beings the unique personal identities we assume them to have?

Many philosophers have attempted to tackle the issue of personal identity, generating a number of distinct theories. I shall provide a synopsis of the two major accounts, mentioning some of the major players, and proceed to reconcile these opposing views with a hybrid account of what constitutes a personal identity which persists over time as a numerically identical individual, or in other words, what makes a single person.

The two major – and rival – accounts of personal identity in philosophy have been *physical* or body-based theories, and *psychological* theories of persistent identity. The dominant of the two are those theories which adhere to some form of psychology-based criterion of continuing
personal identity. Yet before delving into this account I would like to summarize the physicalist approach.

The **bodily continuity** criterion for personal identity states that for a person at a particular time (t1) and a person at a later time (t2) to be numerically identical (meaning, retaining a single identity which has persisted over time), the person at t1 (P1) and the person at t2 (P2) must possess the same body. If it can be said that the body in question is indeed the same body despite any changes in regard to its individual parts or particular material composition, then P2 is indeed the same person as P1.

This view focuses upon a body in its entirety: a single human body which may be said to be the same physical thing as a previous body regardless of differences in some descriptive characteristics. Hence, if we follow the existence of the physical body which received the name ‘Greg’ at birth to the same grown body called ‘Greg’ at age twenty-five, then despite many differing physical traits, it may be said that this is indeed the same individual to whom the name ‘Greg’ was given in infancy. Therefore, on this theory, what matters for continuing personal identity is the continuing existence of a single physical entity. (More complex and elaborate versions of this theory have been put forth by David Wiggins and Eric Olsen.)

By contrast, **psychological** theories assert that the criterion for the persistence of personal identity over time is the intertwined relations of an individual’s psychological states. Initially, this theory was postulated by John Locke (1632-1704), often deemed the father of the personal identity problem. He employed memory as the sole criterion for identity. Later the theory was revised, by Lockeans and others, to include a plethora of psychological factors, not solely memories, as means of accounting for one’s singular personal identity over time. These more sophisticated theories focus primarily on either **psychological continuity** or **psychological connectedness**, or on a fusion of the two, and often rely on the idea of ‘person-stages’ (a person at t1 is a person-stage, at t2 is another person-stage, etc).

The psychological continuity theory typically states that in order for P1 at t1 to be identical to P2 at t2, some continuity of memory and personality must be recognizable between P1 and P2. The psychological connectedness theory, closely related to the psychological continuity theory, maintains that some type of psychological connectedness is necessary between person-stages for the two to have a single identity over time; but unlike memory-based theories of identity, the entirety of the contents of psychological states may be analysed and utilised to ascribe identity. To borrow a concise summary of Harold Noonan’s from his book *Personal Identity* (1989):
“One such connection is that which holds between an intention and the later act in which this intention is carried out. Other such direct psychological connections are those which hold when a belief, desire, or any other psychological feature, persists… In general any causal links between past factors and present psychological traits [not merely memories] can be subsumed under the notion of psychological connectedness.” (pp.10-11).

Objections and Persons
A classic refutation of Locke’s simple memory criterion for personal identity has been made by Thomas Reid (1710-96). His ‘Paradox of the Brave Officer’ essentially goes as follows. Consider a child who grows into a young man, and then into an old man. Based on a simple memory criterion alone, one could assert that the child is psychologically connected to the young man if the young man has a good portion of the memories of the child; and the young man is psychologically connected to the old man insofar as the old man has sufficient memories of being the young man. However, the old man may nevertheless be said to be psychologically discontinuous with, that is, unconnected with, the child, due to the fact that the old man has no memories of being the child. Yet, how is it possible for the child to be the young man, and the young man to be the old man, but for the child to be a different person from the old man?

“Obviously, these objections hit their target [the simple Lockean memory criterion], but they do not go deep” Noonan writes (p.55). That is, while powerful in its time, this objection fails to be an adequate objection to contemporary theories of psychological continuity, which say that as long as there is a continuous set of links of memories between the child and the old man, they may be said to be the same person. So (for example) as long as the old man can remember being the young man, and the young man can remember being the child, then the old man is the same person as the child.

One influential argument in favour of psychological rather than physical theories of identity has been put forward by Derek Parfit in *Reasons and Persons* (1984). It goes as follows. An individual enters a teleport machine on Earth, loses consciousness, and awakes in the teleport on Mars. The machine on Earth is the ‘scanner’ and the one on Mars is the ‘replicator’. Once the scanner has scanned the precise states of each molecule of the person’s body, it beams that information to the replicator on Mars and simultaneously completely destroys the body on Earth. Out of entirely new matter, the replicator on Mars creates a body which is an exact replica of the previous one. The person then steps out of the replicator with no thought that he is not continuous with the person on Earth, and thus he may be considered the same person. So this person has psychological but not bodily continuity with the person on Earth.
Despite their dominance in philosophy, there are objections to psychological theories of personal identity. One such objection might be called ‘the duplication problem’. It is conceivable that one day there will exist a machine which will be able to record everything about one’s psychological states and transfer this information into a new body, or even into more than one body. This case is akin to a variation on Parfit’s teleportation thought experiment, in which the replicator malfunctions and produces a number of exact replicas of the body being transported. In either case, more than one individual will be in possession of precisely the same psychological states, all of which are continuous and connected with one previous person. According to this critique, the psychological criteria for identity must therefore fail, for we shall be left with two or more embodied people who according to the psychological criteria may rightfully be considered continuous with the same person. Intuitively, this seems rather absurd.

Another Story of Identity

While the defenders of the psychological criteria and the advocates of the bodily criteria continue to duel, concocting amusing and intriguing science-fiction-inspired thought-experiments, neither group has successfully managed to take down their opponent.

Each camp of theorists has attempted to capture something of what makes a human being a person retaining a single identity. However, neither position seems to capture another integral element of our lived existences, namely, that we tend to define ourselves through the telling of stories. We get to know one another by learning about each other’s life histories, and we relate to others, identifying with them, based on their values, ideologies, beliefs, personalities, etc, all of which are transmitted via narratives, verbal, written, or otherwise. Hence, an alternative response to the philosophical identity crisis has been the proposal that a human self gains its identity through narration. This is often referred to as Narrative Identity Theory.

All Narrative Identity Theorists maintain in some form or another that the identities of persons are self-created narratives – claiming that narration, or story-telling, is the mode in which we represent ourselves to ourselves, present ourselves to others, and represent others around us. The narrative theorist is attempting to capture that element of experience in which we say, “Hey, tell me your story,” or “I know you, I’ve heard stories about you.”

On this account, who one is (and is not) is contingent upon the stories of one’s past, and the stories of who one wishes to become; the goals one possesses and the actions taken to arrive at those ends; the values inherited narratively or arrived at through reflection and self-story-telling; and one’s emplotment as a character in the story of one’s life, interacting with the stories of others. The narrative theorist thus takes human linguistic abilities and goal-orientation to play a
major role in someone’s acquisition of a unique identity as a person. Some theorists have maintained that the personal self is the product of an interactive unified narrative, others the virtual center of multiple narrative streams, while yet others maintain a more existential position, viewing the self as a constant becoming, evolving as we interact with our environments and reflect on our lives. Pre-eminent defenders of narrative identity include Daniel Dennett, Alasdair MacIntyre, and Paul Ricoeur. Although they differ in their approaches, they all attempt to capture features of the human condition which previous theorizing has ultimately left out, namely the importance of our life histories, story-telling, cultural immersion, goal-directedness, and self-creation.

Back To Life
While these theories of what make you continue to be you may seem obscure or abstract, they do indeed hold some bearing on human life and the concerns which arise on a daily basis – especially in medical settings, where we are faced with issues relating to brain death, permanent vegetative states, comas, advance directives and living wills, and many psychiatric dilemmas. All of these in one way or another evoke questions touching on the various theories presented.

Retiring from my sojourn in the park, having pondered the great mysteries of the human condition, I asked myself, “Could it not be that I am at once dependent upon my psychological connectedness, my biological persistence, and my life history, for my identity?” Although I did not accomplish a miraculous philosophical breakthrough during my stroll, I hope I have provided some food for thought with this précis of positions on the ‘Philosophical Identity Crisis’.

© Dr Chris Durante 2013

*Chris Durante has a PhD in Ethics, MA in Religious Studies, and MSc in Philosophy of Mental Disorder. His interests in theories of identity and personhood span a variety of fields, including bioethics, philosophy, comparative religion and sociopolitical theory. He teaches at McGill University in Montréal, and also at Marymount Manhattan College and St John’s University, both in New York City.*
Star Trek is one of my all-time favorite television franchises and over its various incarnations, from TV shows to movies to cartoons, it has raised many thoughtful and intriguing philosophical thought experiments—especially related to personal identity. Voyager was one of the darker series in the franchise but still delighted in addressing philosophical questions deep in the heart of the Delta quadrant.

We have been talking about the nature of the self, self-identity, and the relative roles of memory, body, brain, etc. as a criterion of personal identity. These issues are at play in the *Voyager* episode “Tuvix” in which a transporter malfunction fuses Neelix and Tuvok into Tuvix. The episode is available on Netflix and Hulu and was featured in an online debate you can review [HERE](#). You can find a transcript for the episode [HERE](#).

After you watch “Tuvix,” let’s consider the following. Imagine that you are the Captain of Voyager (You’re Janeway) and you are faced with the difficult decision regarding what to do with Tuvix. Do you separate him into his component parts, Neelix and Tuvok, or do you permit him to continue as the person he seems to be, Tuvix? Were you in Janeway’s position, what would you do and how would you philosophically defend your decision? There are a number of issues you might consider:

- The nature of Tuvix’s self-identity
- The relation of Tuvix’s self to Neelix and Tuvok
- The status of Neelix and Tuvok once they fuse (and presumably after they are separated)
- The status of Tuvix after he is separated into Neelix and Tuvok
- The criteria of personal identity that must be employed when making sense of these events

So, what would you do? And how would you justify it to the crew?
Introduction to the Reading
By Dennis Weiss

What makes you the same person from one moment to the next? Philosophers have come up with competing answers to this question: the soul, the mind, memory and consciousness, the body. While today there is a very developed literature on the issue of personal identity, Locke’s essay on identity is one of the most significant early accounts of personal identity and one of the most confusing. Part of this confusion is owing to the fact that Locke was really addressing new issues but doing so with an older vocabulary. Locke’s account of person identity is wrapped up with his account of what constitutes the same man and his attempts to deal with a substance view of personal identity.
In the early sections of his essay, not included here, Locke argues that our account of sameness and identity will vary depending on what we are talking about. What makes for the same tree is not the same thing as what makes for the same rock. So it is important, he cautions, that we be clear about which object we are talking. When talking about “the same man,” for instance, Locke argues that we should treat “the same man” as analogous to “the same animal: and suggests that in order to have the same man you must have the same physical body united to the same one life. It is in section 9 that Locke begins to address the issue of personal identity, what constitutes the same person or same self. These are the central sections wherein Locke defends his account of personal identity as depending on consciousness and memory. In section 9, Locke defines “person” (a thinking intelligent being...) and offers a preliminary account of what constitutes the same person (in this alone consists personal identity...). In section 10, Locke reiterates his view that what constitutes the same person is consciousness (one’s memories of one’s life) and not the same substance (not the same mind or soul or body). Read carefully Locke’s account of the role of memory (For as far as any intelligence being can repeat the idea...). All this suggests that what constitutes the same person for Locke is having the same memories. Try to reconstruct his argument for this position. Locke reiterates his position in sections 16 and 17. Locke also attempts to deal with obvious objections to his position. For instance, he raises the problem of forgetfulness. If my self-identity depends on my memories, what happens if I forget? This issue comes up in sections 10 and 20. Try to figure out how Locke responds to this problem. Locke is also trying to meet the objections of followers of Descartes. He is concerned with the issue of immaterial substances and personal identity because he is aware of Descartes’ argument (which perhaps you encountered in the unit on mind and consciousness) that what makes you the person you are is the soul or immaterial thinking substance. Maybe what makes you the same person from one day to the next is the existence of some permanent immaterial thinking thing or soul. How does Locke attempt to meet this objection? He addresses these issues primarily in sections 12 - 14. He develops an argument from analogy meant to persuade us that it is not the same substance that makes you the same person, rather it is consciousness or memory. In section 15 Locke raises one of the earliest thought experiments in this area when he imagines the soul of a prince being exchanged with the soul of a cobbler. In section 16 he imagines what would happen were he to remember the ark and Noah’s flood. Pay attention to these thought experiments as they play an important role in Locke’s overall theory. Locke also deals with the issue of accountability. When ought we to hold someone accountable or responsible for their actions. How does Locke’s reference to the way in which a court of law generally works help support his argument about personal identity? Note, too, that when dealing with the issue of accountability, we might have to distinguish between same man and same person.

An Essay Concerning Human Understanding
By John Locke

Book II, Chapter XXVII: Of Identity and Diversity
Paragraphs 9 – 20

9. **Personal identity.** This being premised, to find wherein personal identity consists, we must consider what person stands for:- which, I think, 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 only by that consciousness which is inseparable
from thinking, and, as it seems to me, essential to it: it being impossible for any one to
perceive without perceiving that he does perceive. When we see, hear, smell, taste, feel,
meditate, or will anything, we know that we do so. Thus it is always as to our present
sensations and perceptions: and by this every one is to himself that which he calls self:-
it not being considered, in this case, whether the same self be continued in the same or
divers substances. 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.

10. Consciousness makes personal identity. But it is further inquired, whether it be the
same identical substance. This few would think they had reason to doubt of, if these
perceptions, with their consciousness, always remained present in the mind, whereby
the same thinking thing would be always consciously present, and, as would be thought,
evidently the same to itself. But that which seems to
make the difficulty is this, that this consciousness being
interrupted always by forgetfulness, there being no
moment of our lives wherein we have the whole train
of all our past actions before our eyes in one view, but
even the best memories losing the sight of one part whilst they are viewing another;
and we sometimes, and that the greatest part of our lives, not reflecting on our past
selves, being intent on our present thoughts, and in sound sleep having no thoughts
at all, or at least none with that consciousness which remarks our waking thoughts,- I
say, in all these cases, our consciousness being interrupted, and we losing the sight of
our past selves, doubts are raised whether we are the same thinking thing, i.e. the same
substance or no. Which, however reasonable or unreasonable, concerns not personal
identity at all. The question being what makes the same person; and not whether it be
the same identical substance, which always thinks in the same person, which, in this
case, matters not at all: different substances, by the same consciousness (where they do
partake in it) being united into one person, as well as different bodies by the same life
are united into one animal, whose identity is preserved in that change of substances by
the unity of one continued life. For, it being the same consciousness that makes a man be
himself to himself, personal identity depends on that only, whether it be annexed solely
to one individual substance, or can be continued in a succession of several substances.
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: For it is by the consciousness
it has of its present thoughts and actions, that it is self to itself now, and so will be the
same self, as far as the same consciousness can extend to actions past or to come. and
would be by distance of time, or change of substance, no more two persons, than a man
be two men by wearing other clothes to-day than he did yesterday, with a long or a short
sleep between: the same consciousness uniting those distant actions into the same person,
whatever substances contributed to their production.

11. Personal identity in change of substance. That this is so, we have some kind of
evidence in our very bodies, all whose particles, whilst vitally united to this same
thinking conscious self, so that we feel when they are touched, and are affected by,
and conscious of good or harm that happens to them, as a part of ourselves; i.e. of our
thinking conscious self. Thus, the limbs of his body are to every one a part of Himself;
he sympathizes and is concerned for them. Cut off a
hand, and thereby separate it from that consciousness he
had of its heat, cold, and other affections, and it is then
no longer a part of that which is himself, any more than
the remotest part of matter. Thus, we see the substance
whereof personal self consisted at one time may be varied
at another, without the change of personal identity; there
being no question about the same person, though the limbs
which but now were a part of it, be cut off.

12. Personality in change of substance. But the question is, Whether if the same substance
which thinks be changed, it can be the same person; or, remaining the same, it can be
different persons? And to this I answer: First, This can be no question at all to those who
place thought in a purely material animal constitution, void of an immaterial substance.
For, whether their supposition be true or no, it is plain they conceive personal identity
preserved in something else than identity of substance; as animal identity is preserved
in identity of life, and not of substance. And therefore those who place thinking in an
inmaterial substance only, before they can come to deal with these men, must show why
personal identity cannot be preserved in the change of immaterial substances, or variety
of particular immaterial substances, as well as animal
identity is preserved in the change of material substances,
or variety of particular bodies: unless they will say, it is
one immaterial spirit that makes the same life in brutes, as
it is one immaterial spirit that makes the same person in
men; which the Cartesians at least will not admit, for fear
of making brutes thinking things too.

13. Whether in change of thinking substances there can be one person. But next, as to the
first part of the question, Whether, if the same thinking substance (supposing immaterial
substances only to think) be changed, it can be the same person? I answer, that cannot
be resolved but by those who know what kind of substances they are that do think; and
whether the consciousness of past actions can be transferred from one thinking substance
to another. I grant were the same consciousness the same individual action it could not:
but it being a present representation of a past action, why it may not be possible, that

Can you reconstruct the argument
that Locke is presenting in this
section? Keep in mind that
“thus” is a conclusion indicator
and so Locke is presenting a
conclusion to an argument. What
are the premises that support this
conclusion?

Locke takes up the perspective of
Rene Descartes. What claims does
he present in arguing against the
substance view of Descartes? Try
to reconstruct his reasoning.
that may be represented to the mind to have been which really never was, will remain to be shown. And therefore how far the consciousness of past actions is annexed to any individual agent, so that another cannot possibly have it, will be hard for us to determine, till we know what kind of action it is that cannot be done without a reflex act of perception accompanying it, and how performed by thinking substances, who cannot think without being conscious of it. But that which we call the same consciousness, not being the same individual act, why one intellectual substance may not have represented to it, as done by itself, what it never did, and was perhaps done by some other agent-why, I say, such a representation may not possibly be without reality of matter of fact, as well as several representations in dreams are, which yet whilst dreaming we take for true- will be difficult to conclude from the nature of things. And that it never is so, will by us, till we have clearer views of the nature of thinking substances, be best resolved into the goodness of God; who, as far as the happiness or misery of any of his sensible creatures is concerned in it, will not, by a fatal error of theirs, transfer from one to another that consciousness which draws reward or punishment with it. How far this may be an argument against those who would place thinking in a system of fleeting animal spirits, I leave to be considered. But yet, to return to the question before us, it must be allowed, that, if the same consciousness (which, as has been shown, is quite a different thing from the same numerical figure or motion in body) can be transferred from one thinking substance to another, it will be possible that two thinking substances may make but one person. For the same consciousness being preserved, whether in the same or different substances, the personal identity is preserved.

14. Whether, the same immaterial substance remaining, there can be two persons. As to the second part of the question, Whether the same immaterial substance remaining, there may be two distinct persons; which question seems to me to be built on this,- Whether the same immaterial being, being conscious of the action of its past duration, may be wholly stripped of all the consciousness of its past existence, and lose it beyond the power of ever retrieving it again: and so as it were beginning a new account from a new period, have a consciousness that cannot reach beyond this new state. 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 body, or informing any other body; and if they should not, it is plain experience would be against them. So that personal identity, reaching no further than consciousness reaches, a pre-existent spirit not having continued so many ages in a state of silence, must needs make different persons. Suppose a Christian Platonist or a Pythagorean should, upon God’s having ended all his works of creation the seventh day, think his soul hath existed ever since; and should imagine it has revolved in several human bodies; as I once met with one, who was persuaded his had been the soul of Socrates (how reasonably I will not dispute; this I know, that in the post he filled, which was no inconsiderable one, he passed for a very rational man, and the press has shown that he wanted not parts or learning;)—would any one say, that he, being not conscious of any of Socrates’s actions or thoughts, could be the same person with Socrates? Let any one reflect upon himself, and conclude that he has in himself an immaterial spirit, which is that which thinks in him, and, in the constant change of his body keeps him the same: and is that which he calls himself: let
him also suppose it to be the same soul that was in Nestor or Thersites, at the siege of Troy, (for souls being, as far as we know anything of them, in their nature indifferent to any parcel of matter, the supposition has no apparent absurdity in it), which it may have been, as well as it is now the soul of any other man: but he now having no consciousness of any of the actions either of Nestor or Thersites, does or can he conceive himself the same person with either of them? Can he be concerned in either of their actions? attribute them to himself, or think them his own, more than the actions of any other men that ever existed? So that this consciousness, not reaching to any of the actions of either of those men, he is no more one self with either of them than if the soul or immaterial spirit that now informs him had been created, and began to exist, when it began to inform his present body; though it were never so true, that the same spirit that informed Nestor’s or Thersites’ body were numerically the same that now informs his. For this would no more make him the same person with Nestor, than if some of the particles of matter that were once a part of Nestor were now a part of this man; the same immaterial substance, without the same consciousness, no more making the same person, by being united to any body, than the same particle of matter, without consciousness, united to any body, makes the same person. But let him once find himself conscious of any of the actions of Nestor, he then finds himself the same person with Nestor.

15. The body, as well as the soul, goes to the making of a man. And thus may we be able, without any difficulty, to conceive the same person at the resurrection, though in a body not exactly in make or parts the same which he had here,—the same consciousness going along with the soul that inhabits it. But yet the soul alone, in the change of bodies, would scarce to any one but to him that makes the soul the man, be enough to make the same man. For should the soul of a prince, carrying with it the consciousness of the prince’s past life, enter and inform the body of a cobbler, as soon as deserted by his own soul, every one sees he would be the same person with the prince, accountable only for the prince’s actions: but who would say it was the same man? The body too goes to the making the man, and would, I guess, to everybody determine the man in this case, wherein the soul, with all its princely thoughts about it, would not make another man: but he would be the same cobbler to every one besides himself. I know that, in the ordinary way of speaking, the same person, and the same man, stand for one and the same thing. And indeed every one will always have a liberty to speak as he pleases, and to apply what articulate sounds to what ideas he thinks fit, and change them as often as he pleases. But yet, when we will inquire what makes the same spirit, man, or person, we must fix the ideas of spirit, man, or person in our minds; and having resolved with ourselves what we mean by them, it will not be hard to determine, in either of them, or the like, when it is the same, and when not.

16. Consciousness alone unites actions into the same person. But though the same immaterial substance or soul does not alone, wherever it be, and in whatsoever state, make the same man; yet it is plain, consciousness, as far as ever it can be extended—should it be to ages past—unites existences and actions very remote in time into the
same person, as well as it does the existences and actions of the immediately preceding moment: so that whatever has the consciousness of present and past actions, is the same person to whom they both belong. *Had I the same consciousness that I saw the ark and Noah’s flood, as that I saw an overflowing of the Thames last winter, or as that I write now, I could no more doubt that who write this now, that saw’ the Thames overflowed last winter, and that viewed the flood at the general deluge, was the same self,- place that self in what substance you please- than that I who write this am the same myself now whilst I write (whether I consist of all the same substance, material or immaterial, or no) that I was yesterday. For as to this point of being the same self, it matters not whether this present self be made up of the same or other substances- I being as much concerned, and as justly accountable for any action that was done a thousand years since, appropriated to me now by this self-consciousness, as I am for what I did the last moment.

17. **Self depends on consciousness, not on substance.** 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 and pain, capable of happiness or misery, and so is concerned for itself, as far as that consciousness extends. Thus every one finds that, whilst comprehended under that consciousness, the little finger is as much a part of himself as what is most so. Upon separation of this little finger, should this consciousness go along with the little finger, and leave the rest of the body, it is evident the little finger would be the person, the same person; and self then would have nothing to do with the rest of the body. As in this case it is the consciousness that goes along with the substance, when one part is separate from another, which makes the same person, and constitutes this inseparable self: so it is in reference to substances remote in time. That with which the consciousness of this present thinking thing can join itself, makes the same person, and is one self with it, and with nothing else; and so attributes to itself, and owns all the actions of that thing, as its own, as far as that consciousness reaches, and no further; as every one who reflects will perceive.

18. **Persons, not substances, the objects of reward and punishment.** In this personal identity is founded all the right and justice of reward and punishment; happiness and misery being that for which every one is concerned for himself, and not mattering what becomes of any substance, not joined to, or affected with that consciousness. For, as it is evident in the instance I gave but now, 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 then it cannot but admit as its own now. Though, if the same body should still 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 own any of its actions, or have any of them imputed to him.
19. **Which shows wherein personal identity consists.** This may show us wherein personal identity consists: not in the identity of substance, but, as I have said, in the identity of consciousness, wherein if Socrates and the present mayor of Queinborough agree, they are the same person: *if the same Socrates waking and sleeping do not partake of the same consciousness, Socrates waking and sleeping is not the same person.* And to punish Socrates waking for what sleeping Socrates thought, and waking Socrates was never conscious of, would be no more of right, than to punish one twin for what his brother-twin did, whereof he knew nothing, because their outsides were so like, that they could not be distinguished; for such twins have been seen.

20. **Absolute oblivion separates what is thus forgotten from the person, but not from the man.** But yet possibly it will still be objected,- Suppose I wholly lose the memory of some parts of my life, beyond a possibility of retrieving them, so that perhaps I shall never be conscious of them again; yet am I not the same person that did those actions, had those thoughts that I once was conscious of, though I have now forgot them? To which I answer, that we must here take notice what the word I is applied to; which, in this case, is the man only. And the same man being presumed to be the same person, I is easily here supposed to stand also for the same person. But 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; which, we see, is the sense of mankind in the solemnest declaration of their opinions, human laws not punishing the mad man for the sober man’s actions, nor the sober man for what the mad man did,- thereby making them two persons: which is somewhat explained by our way of speaking in English when we say such an one is “not himself,” or is “beside himself”; in which phrases it is insinuated, as if those who now, or at least first used them, thought that self was changed; the selfsame person was no longer in that man.
“Thomas Reid” by Sir Henry Raeburn | Wikimedia Commons

For biographical information on Thomas Reid, click here.

**Introduction to the Reading**  
*By Dennis Weiss*

The following selection from Thomas Reid presents us with two challenges. The first is to figure out his own view of personal identity and the second is to figure out his critique of Locke and how this is connected to his own view. It is perhaps easier to start with his critique of Locke, which appears in the selection from chapter 6 of his work. He begins that section with a summary of Locke’s own view, which is always useful, focusing on Locke’s account of personal identity and remembrance, or memory. Then, he begins his critique of Locke’s account of personal identity and memory (“This doctrine has some strange consequences...). Reid mentions several strange consequences but we only get two in this selection. The first is in the same paragraph just cited and deals with the transfer of consciousness. The second goes by the name of “The Brave
Officer Paradox.” Consider these objections to Locke’s memory theory carefully. Do they indeed suggest that Locke’s account of personal identity is wrong? Could Locke defend himself against these objections?

If Reid critiques Locke’s version of the memory theory, what then does that imply about his own approach to personal identity? We get an account of Reid’s own approach to personal identity in the selection from chapter 4. In order to determine Reid’s approach, you will have to read this section carefully. Notice that in the first paragraph, Reid, like Locke, focuses on memory. Does this mean he supports a memory theory of personal identity? Reid attempts to improve upon Locke’s position by noting our conviction that there must be something that persists through time and that I call myself. It is presupposed in the very operations of reason. Reid seems to reach a preliminary conclusion (“From this it is evident....”) but then suggests that we focus on the distinct concepts of identity in general, personal identity, and the role of memory in personal identity. Reid spends several paragraphs talking about identity in general and concludes that identity “supposes an uninterrupted continuance of existence.” What forces him to this conclusion? He then turns to the more difficult significance of personal identity (“It is perhaps more difficult to ascertain...”). Pay close attention to the points that Reid is making in these paragraphs, through to the beginning of the next section. Why does he think that persons are indivisible? What are the implications of this for personal identity? What does Reid take the self to be? Pay special attention to the role that Reid suggests memory and remembrance play in personal identity. While Locke suggests that personal identity is constituted by memory, Reid suggests that the evidence I have for a permanent self is memory and remembrance. There is a slight but important distinction here. This distinction is especially evident in Reid’s discussion of what makes me the person that did an action. Ultimately, it is these comments that might help to distinguish Reid’s approach to personal identity from Locke’s.

Essays on the Intellectual Powers of Man
By Thomas Reid
Essay III, “Of Memory”
Chapter 4

The conviction which every man has of his identity, as far back as his memory reaches, needs no aid of philosophy to strengthen it; and no philosophy can weaken it, without first producing some degree of insanity.

The philosopher, however, may very properly consider this conviction as a phenomenon of human nature worthy of his attention. If he can discover its cause, an addition is made to his stock of knowledge; if not, it must be held as a part of our original constitution, or an effect of that constitution produced in a manner unknown to us.

We may observe, first of all, that this conviction is indispensably necessary to all exercise of reason. The operations of reason, whether in action or in speculation, are made up of successive parts. The antecedent are the
foundation of the consequent, and, without the conviction that the antecedent have been seen or
done by me, I could have no reason to proceed to the consequent, in any speculation, or in any
active project whatever.

There can be no memory of what is past without the conviction that we existed at the time
remembered. There may be good arguments to convince me that I existed before the earliest
thing I can remember; but to suppose that my memory reaches a moment farther back than by
belief and conviction of my existence is a contradiction.

The moment a man loses this conviction, as if he had drunk the water of Lethe, past things are
done away; and, in his own belief, he then begins to exist. Whatever was thought, or said, or
done, or suffered before that period, may belong to some other person; but he can never impute it
to himself, or take any subsequent step that supposes it to be his doing.

From this it is evident that we must have the conviction of our
own continued existence and identity, as soon as we are capable of
thinking or doing anything, on account of what we have thought,
or done, or suffered before; that is, as soon as we are reasonable
creatures.

That we may form as distinct a notion as we are able of this
phenomenon of the human mind, it is proper to consider what is
meant by identity in general, what by our own personal identity,
and how we are led into that invincible belief and conviction
which every man has of his own personal identity, as far as his memory reaches.

Identity in general I take be a relation between a thing which is known to exist at one time, and
a thing which is known to have existed at another time. If you ask whether they are one and the
same, or two different things, every man of common sense understands the meaning of your
questions perfectly. Whence we may infer with certainty, that every man of common sense has a
clear and distinct notion of identity.

If you ask a definition of identity, I confess I can give none; it is too simple a notion to admit
of logical definition: I can say it is a relation, but I cannot find words to express the specific
difference between this and other relations, though I am in no danger of confounding it with
any other. I can say that diversity is a contrary relation, and that similitude and dissimilitude are
another couple of contrary relations, which every man easily distinguishes in his conception from
identity and diversity.

I see evidently that identity supposes an uninterrupted continuance of existence. That which has
ceased to exist cannot be the same with that which afterwards begins to exist; for this would
be to suppose a being to exist after it ceased to exist, and to have had existence before it was
produced, which are manifest contradictions. Continued uninterrupted existence is therefore
necessarily implied in identity.
Hence we may infer, that identity, cannot in its proper sense, be applied to our pains and pleasures, our thoughts, or any operation of our minds. The pain felt this day is not the same individual pain which I felt yesterday, though they may be similar in kind and degree, and have the same cause. The same may be said of every feeling, ad of every operation of mind. They are all successive in their nature, like time itself, no two moments of which can be the same moment.

It is otherwise with the parts of absolute space. They always are, and were, and will be the same. So far, I think, we proceed upon clear ground in fixing the notion of identity in general.

It is perhaps more difficult to ascertain with precision the meaning of personality; but it is not necessary in the present subject: it is sufficient for our purpose to observe, that all mankind place their personality in something that cannot be divided or consist of parts.

A part of a person is a manifest absurdity. When a man loses his estate, his health, his strength, he is still the same person, and has lost nothing of his personality. If he has a leg or an arm cut off, he is the same person he was before. The amputated member is not part of his person, otherwise it would have a right to a part of his estate, and be liable for a part of his engagements. It would be entitled to a share of his merit and demerit, which is manifestly absurd. A person is something indivisible, and is what Leibnitz calls a monad.

My personal identity, therefore, implies the continued existence of that indivisible thing which I call myself. Whatever this self may be, it is something which thinks, and deliberates, and resolves, and acts, and suffers. I am not thought, I am not action, I am not feeling; I am something that thinks, and acts, and suffers. My thoughts, and actions, and feelings, change every moment; they have no continued, but a successive, existence; but that self, or I, to which they belong, is permanent, and has the same relation to all the succeeding thoughts, action, and feeling which I call mine.

Such are the notions that I have of my personal identity. But perhaps it may be said, this may all be fancy without reality. How do you know—what evidence have you—that there is such a permanent self which has a claim to all the thoughts, actions, and feelings which you call yours?

To this I answer, that the proper evidence I have of all this is remembrance. I remember that twenty years ago I conversed with such a person; I remember several things that passed in that conversation: my memory testifies, not only that this was done by me, I must have existed at that time, and continued to exist from that time to the present: if the identical person whom I call myself had not a part in that conversation, my memory is fallacious; it gives a distinct and positive testimony of what is not true. Every man in his senses believes what he distinctly remembers, and every thing he remembers convinces him that he existed at the time remembered.

Although memory gives the most irresistible evidence of my being the identical person that did such a thing, at such a time, I may have other good evidence of things which befell me, and which I do not remember: I know who bare me, and suckled me, but I do not remember these events.
It may here be observed (though the observation would have been unnecessary, if some great philosophers had not contradicted it), that it is not my remembering any action of mine that makes me to be the person who did it. This remembrance makes me to know assuredly that I did it; but I might have done it, though I did not remember it. That relation to me, which is expressed by saying that I did it, would be the same, though I had not the least remembrance of it. To say that my remembering that I did such a thing, or, as some choose to express it, my being conscious that I did it, makes me to have done it, appears to me as great an absurdity as it would be to say, that my belief that the world was created made it to be created.

When we pass judgment on the identity of other persons than ourselves, we proceed upon other grounds, and determine from a variety of circumstances, which sometimes produce the firmest assurance, and sometimes leave room for doubt. The identity of persons has often furnished matter of serious litigation before tribunals of justice. But no man of a sound mind ever doubted of his own identity, as far as he distinctly remembered.

The identity of a person is a perfect identity: wherever it is real, it admits of no degrees; and it is impossible that a person should be in part the same, and in part different; because a person is a monad, and is not divisible into parts. The evidence of identity in other persons than ourselves does indeed admit of all degrees, from what we account certainty, to the last degree of probability. But still it is true, that the same person is perfectly the same, and cannot be so in part, or in some degree only.

Chapter 6

In a long chapter upon Identity and Diversity, Mr. Locke has made some ingenious and just observations, and some which I think cannot be defended. I shall only take notice of the account he gives of our own personal identity. His doctrine upon this subject has been censured by Bishop Butler, in a short essay subjoined to his Analogy, with whose sentiments I perfectly agree.

Identity, as was observed (Chapter 4 of this Essay), supposes the continued existence of the being on which it is affirmed, and therefore can be applied only to things which have a continued existence. While any being continues to exist, it is the same being; but two beings which have a different beginning or a different ending of their existence cannot possibly be the same. To that, I think, Mr. Locke agrees.

He observes, very justly, that, to know what is meant by the same person, we must consider what the word person stands for; and he defines a person to be an intelligent being, endowed with reason and with consciousness, which last he thinks inseparable from thought.

From this definition of a person, it must necessarily follow, that, while the intelligent being continues to exist and to be intelligent, it must be the same person. To say that the intelligent
being is the person, and yet that the person ceases to exist while the intelligent being continues, or that the person continues while the intelligent being ceases to exist, it to my apprehension a manifest contradiction.

One would think that the definition of person should perfectly ascertain the nature of personal identity, or wherein it consists, though it might still be a question how we come to know and be assured of our personal identity.

Mr. Locke tell us, however, “that personal identity, that is, the sameness of a rational being, consists in consciousness alone, and, as far as this consciousness can extend backwards to any past action or thought, so far reaches the identity of that person. So that whatever has the consciousness of present and past actions is the same person to whom they belong.”

This doctrine has some strange consequences, which the author was aware of. Such as, that if the same consciousness can be transferred from one intelligent being to another, which he thinks we cannot show to be impossible, then two or twenty intelligent beings may be the same person. And if the intelligent being may lose the consciousness of his actions done by him, which surely is possible, then he is not the person that did those actions; so that one intelligent being may be two or twenty different persons, if he shall so often lose the consciousness of his former actions.

There is another consequence of this doctrine, which follows no less necessarily, though Mr. Locke probably did not see it. It is, that a man may be, and at the same time not be, the person that did a particular action.

Suppose a brave officer to have been flogged when a boy at school for robbing an orchard, to have taken a standard from the enemy in his first campaign, and to have been made a general in advanced life; suppose, also, which must be admitted to be possible, that, when he took the standard, he was conscious of his having been flogged at school, and that, when made a general, he was conscious of his taking the standard, but had absolutely lost the consciousness of his flogging.

These things being supposed, it follows, from Mr. Locke’s doctrine, that he who was flogged at school is the same person who took the standard, and that he who took the standard is the same person who was made a general. Whence it follows, if there be any truth in logic, that the general is the same person with him who was flogged at school. But the general’s consciousness does not reach so far back as his flogging; therefore, according to Mr. Locke’s doctrine, he is not the person who was flogged. Therefore, the general is, and at the same time is not, the same person with him who was flogged at school.

Leaving the consequences of this doctrine to those who have leisure to trace them, we may observe, with regard to the doctrine itself,

First, that Mr. Locke attributes to consciousness the conviction we have of our past actions, as if a man may now be conscious of what he did twenty years ago. It is impossible to understand the meaning of this, unless by consciousness be meant memory, the only faculty by which we have an immediate knowledge of our past actions.
Sometimes, in popular discourse, a man says he is conscious that he did such a thing, meaning that he distinctly remembers that he did it. It is unnecessary, in common discourse, to fix accurately the limits between consciousness and memory. This was formerly shown to be the case with regard to sense and memory: and therefore distinct remembrance is sometimes called sense, sometimes consciousness, without any inconvenience.

But this ought to be avoided in philosophy, otherwise we confound the different powers of the mind, and ascribe to one what really belongs to another. If a man can be conscious of what he did twenty years or twenty minutes ago, there is no use for memory, nor ought we to allow that there is any such faculty. The faculties of consciousness and memory are chiefly distinguished by this, that the first is an immediate knowledge of the present, the second an immediate knowledge of the past.

When, therefore, Mr. Locke’s notion of personal identity is properly expressed, it is, that personal identity consists in distinct remembrance; for, even in the popular sense, to say that I am conscious of a past action means nothing else than that I distinctly remember that I did it.

**Secondly,** it may be observed, that, in this doctrine, not only is consciousness confounded with memory, but, which is still more strange, personal identity is confounded with the evidence which we have of our personal identity.

It is very true, that my remembrance that I did such a thing is the evidence I have that I am the identical person who did it. And this, I am apt to think, Mr. Locke meant. But to say that my remembrance that I did such a thing, or my consciousness, makes me the person who did it, is, in my apprehension, an absurdity too gross to be entertained by any man who attends to the meaning of it; for it is to attribute to memory or consciousness a strange magical power of producing its object, though that object must have existed before the memory or consciousness which produced it.

Consciousness is the testimony of one faculty; memory is the testimony of another faculty; and to say that the testimony is the cause of the thing testified, this surely is absurd, if any thing be, and could not have been said by Mr. Locke, if he had not confounded the testimony with the thing testified.

When a horse that was stolen is found and claimed by the owner, the only evidence he can have, or that a judge or witness can have, that this is the very identical horse which was his property, is similitude. But would it not be ridiculous from this to infer that the identity of a horse consists in similitude only? The only evidence I have that I am the identical person who did such actions is, that I remember distinctly I did them; or, as Mr. Locke expresses it, I am conscious I did them. To infer from this, that personal identity consists in consciousness, is an argument which, if it had any force, would prove the identity of a stolen horse to consist solely in similitude.

**Thirdly,** is it not strange that the sameness or identity of a person should consist in a thing which is continually changing, and is not any two minutes the same?
Our consciousness, our memory, and every operation of the mind, are still flowing like the water of a river, or like time itself. The consciousness I have this moment can no more be the same consciousness I had last moment, that this moment can be the last moment. Identity can only be affirmed of things which have a continued existence. Consciousness, and every kind of thought, are transient and momentary, and have no continued existence; and, therefore, if personal identity consisted in consciousness, it would certainly follow, that no man is the same person any two moments of his life; and as the right and justice of reward and punishment are founded on personal identity, no man could be responsible for his actions.

But though I take this to be the unavoidable consequence of Mr. Locke’s doctrine concerning personal identity, and though some persons may have liked the doctrine the better on this account, I am far from imputing any thing of this kind to Mr. Locke. He was too good a man not have rejected with abhorrence a doctrine which he believed to draw this consequence after it.

**Fourthly**, there are many expressions used by Mr. Locke, in speaking of personal identity, which to me are altogether unintelligible, unless we suppose that he confounded that sameness or identity which we ascribe to an individual with the identity which, in common discourse, is often ascribed to many individuals of the same species.

When we say that pain and pleasure, consciousness and memory, are the same in all men, this sameness can only mean similarity, or sameness of kind. That the pain of one man can be the same individual pain with that of another man is no less impossible, then that one man should be another man: the pain felt by me yesterday can no more be the pain I feel to-day, than yesterday can be this day; and the same thing may be said of every passion and of every operation of the mind. The same kind or species of operation may be in different men, or in the same man at different times; but it is impossible that the same individual operation should be in different men, or in the same man at different times.

When Mr. Locke, therefore, speaks of “the same consciousness being continued through a succession of different substances”; when he speaks of “repeating the idea of a past action with the same consciousness we had of it at the first,” and of “the same consciousness extending to actions past and to come”; these expressions are to me unintelligible, unless he means not the same individual consciousness, but a consciousness that is similar, or of the same kind.

If our personal identity consists in consciousness, as this consciousness cannot be the same individually any two moments, but only of the same kind, it would follow, that we are not for any two moments the same individuals persons, but the same kind of persons.

As our consciousness sometimes ceases to exist, as in sound sleep, our personal identity must cease with it. Mr. Locke allows, that the same thing cannot have two beginnings of existence, so that our identity would be irrecoverably gone every time we ceased to think, if it was but for a moment.
Derek Parfit and Godfrey Vesey

BRAIN TRANSPLANTS AND PERSONAL IDENTITY
A Dialogue

Derek Parfit is an English philosopher who was educated and now teaches at Oxford University. He has made outstanding contributions to the subjects of ethical theory and the problem of personal identity. His major work is Reasons and Persons (1984). Godfrey Vesey was educated at Cambridge University and is a professor of philosophy at Open University. His principal works are Perception (1971) and Personal Identity (1974).

In this dialogue, Vesey introduces the problem of split-brain transplants. That is, a brain is divided into two, and half is put into each of two other people’s brainless heads. Does the original person survive? Parfit then responds by developing his ideas of personal identity as psychological identity.

BRAIN TRANSPLANTS

In 1973 in the Sunday Times there was a report of how a team from the Metropolitan Hospital in Cleveland under Dr. R. J. White had successfully transplanted a monkey’s head on to another monkey’s body. Dr. White was reported as having said, ‘Technically a human head transplant is possible’, and as hoping that ‘it may be possible eventually to transplant parts of the brain or other organs inside the head.’

The possibility of brain transplants gives rise to a fascinating philosophical problem. Imagine the following situation:

Two men, a Mr Brown and a Mr Robinson, had been operated on for brain tumours and brain extractions had been performed on both of them. At the end of the operations, however, the assistant inadvertently put Brown’s brain in Robinson’s head, and Robinson’s brain in Brown’s head. One of these men immediately dies, but the other, the one with Robinson’s body and Brown’s brain, eventually regains consciousness. Let us call the latter ‘Brownson’. Upon regaining consciousness Brownson exhibits great shock and surprise at the appearance of his body. Then, upon seeing Brown’s body, he exclaims incredulously ‘That’s me lying there!’ Pointing to himself he says ‘This isn’t my body; the one over there is!’ When asked his name he automatically replies ‘Brown’. He recognizes Brown’s wife and family (whom Robinson had never met), and is able to describe in detail events in Brown’s life, always describing them as events in his own life. Of Robinson’s past life he evinces no knowledge at all. Over a period of time he is observed to display all of the personality traits, mannerisms, interests, likes and dislikes, and so on, that had previously characterized Brown, and to act and talk in ways completely alien to the old Robinson.


The next step is to suppose that Brown’s brain is not simply transplanted whole into someone else’s brainless head, but is divided in two and half put into each of two other people’s brainless heads. The same memory having been coded in many parts of the cortex, they both then say they are Brown, are able to describe events in Brown’s life as if they are events in their own lives, etc. What should we say now?

The implications of this case for what we should say about personal identity are considered by Derek Parfit in a paper entitled ‘Personal Identity’. Parfit’s own view is expressed in terms of a relationship he calls ‘psychological continuity’. He analyses this relationship partly in terms of what he calls ‘q-memory’ (‘q’ stands for ‘quasi’). He sketches a definition of ‘q-memory’ as follows:

I am q-remembering an experience if (1) I have a belief about a past experience which seems in itself like a memory belief, (2) someone did have such an experience, and (3) my belief is dependent upon this experience in the same way whatever that is) in which a memory of an experience is dependent upon it.1

The significance of this definition of q-memory is that two people can, in theory, q-remember doing what only one person did. So two people can, in theory, be psychologically continuous with one person.

Parfit’s thesis is that there is nothing more to personal identity than this ‘psychological continuity’. This is not to say that whenever there is a sufficient degree of psychological continuity there is personal identity, for psychological continuity could be a one- two, or ‘branching’, relationship, and we are able to speak of ‘identity’ only when there is a one-one relationship. It is to say that a common belief—in the special nature of personal identity—is mistaken.

In the discussion that follows I began by asking Parfit what he thinks of this common belief.

PERSONAL IDENTITY

Vesey: Derek, can we begin with the belief that you claim most of us have about personal identity? It’s this: whatever happens between now and some future time either I shall still exist or I shan’t. And any future experience will either be my experience or it won’t. In other words, personal identity is an all or nothing matter: either I survive or I don’t. Now what do you want to say about that?

Parfit It seems to me just false. I think the true view is that we can easily describe and imagine large numbers of cases in which the question, ‘Will that future person be me—or someone else?’, is both a question which doesn’t have any answer at all, and there’s no puzzle that there’s no answer.

Vesey: Will you describe one such case.

Parfit One of them is the case discussed in the correspondence material, the case of division in which we suppose that each half of my brain is to be transplanted into a new body and the two resulting people will both seem to remember the whole of my life, have my character and be psychologically continuous with me in every way. Now in this case of division there were only three possible answers to the question: ‘What’s going to happen to me?’ And all three of them seem to me open to very serious objections. So the conclusion to be drawn from the case is that the question of what’s going to happen to me, just doesn’t have an answer. I think the case also shows that that’s not mysterious at all.

Vesey: Right, let’s deal with these three possibilities in turn.

Parfit Well, the first is that I’m going to be both of the resulting people. What’s wrong with that answer is that it leads very quickly to a contradiction.

Vesey: How?

Parfit The two resulting people are going to be different people from each other. They’re going to live completely different lives. They’re going to be as different as any two people are. But if they’re different people from each other it can’t be the case that I’m going to be both of them. Because if I’m both of them, then one of the resulting people is going to be the same person as the other.

Vesey: Yes. They can’t be different people and be the same person, namely me.

Parfit Exactly. So the first answer leads to a contradiction.

Vesey: Yes. And the second?

Parfit Well, the second possible answer is that I’m not going to be both of them but just one of them. This

Keep track of how Parfit develops his argument. What are the three possible answers and how does he analyze them?

Parfit’s analysis of this first possibility has to do with the nature of identity, where a = a and identity is one-to-one.
doesn’t lead to a contradiction, it’s just wildly implausible. It’s implausible because my relation to each of the resulting people is exactly similar.

Vesey: Yes, so there’s no reason to say that I’m one rather than the other?

Parfit: It just seems absurd to suppose that, when you’ve got exactly the same relation, one of them is identity and the other is nothing at all.

Vesey: It does seem absurd, but there are philosophers who would say that sort of thing. Let’s go on to the third.

Parfit: Well, the only remaining answer, if I’m not going to be both of them or only one of them, is that I’m going to be neither of them. What’s wrong with this answer is that it’s grossly misleading.

Vesey: Why?

Parfit: If I’m going to be neither of them, then there’s not going to be anyone in the world after the operation who’s going to be me. And that implies, given the way we now think, that the operation is as bad as death. Because if there’s going to be no one who’s going to be me, then I cease to exist. But it’s obvious on reflection that the operation isn’t as bad as death. It isn’t bad in any way at all. That this is obvious can be shown by supposing that when they do the operation only one of the transplants succeeds and only one of the resulting people ever comes to consciousness again.

Vesey: Then I think we would say that this person is me. I mean we’d have no reason to say that he wasn’t.

Parfit: On reflection I’m sure we would all think that I would survive as that one person.

Vesey: Yes.

Parfit: Yes. Well, if we now go back to the case where both operations succeed . . .

Vesey: Where there’s a double success . . .

Parfit: It’s clearly absurd to suppose that a double success is a failure.

Vesey: Yes.

Parfit: So the conclusion that I would draw from this case is firstly, that to the question, ‘What’s going to happen to me?’, there’s no true answer.

Vesey: Yes.

Parfit: Secondly, that if we decide to say one of the three possible answers, what we say is going to obscure the true nature of the case.

Vesey: Yes.

Parfit: And, thirdly, the case isn’t in any way puzzling. And the reason for that is this. My relation to each of the resulting people is the relation of full psychological continuity. When I’m psychologically continuous with only one person, we call it identity. But if I’m psychologically continuous with two future people, we can’t call it identity. It’s not puzzling because we know exactly what’s going to happen.

Vesey: Yes, could I see if I’ve got this straight? Where there is psychological continuity in a one-one case, this is the sort of case which we’d ordinarily talk of in terms of a person having survived the operation, or something like that.

Parfit: Yes.

Vesey: Now what about when there is what you call psychological continuity—that’s to say, where the people seem to remember having been me and so on—in a one-two case? Is this survival or not?

Parfit: Well, I think it’s just as good as survival, but the block we have to get over is that we can’t say that anyone in the world after the operation is going to be me.

Vesey: No.

Parfit: Well, we can say it but it’s very implausible. And we’re inclined to think that if there’s not going to be anyone who is me tomorrow, then I don’t survive. What we need to realize is that my relation to each of those two people is just as good as survival. Nothing is missing at all in my relation to both of them, as compared with my relation to myself tomorrow.

Vesey: Yes.

Parfit: So here we’ve got survival without identity. And that only seems puzzling if we think that identity is a further fact over and above psychological continuity.

Vesey: It is very hard not to think of identity being a further fact, isn’t it?

Parfit: Yes, I think it is. I think that the only way to get rid of our temptation to believe this is to consider many more cases than this one case of division. Perhaps I should give you another one. Suppose that the following is going to happen to me. When I die in a normal way, scientists are going to map the states of all the cells in my brain and body and after a few

Psychological continuity is not the same as identity.

Parfit's key idea is to separate identity and psychological continuity.

Parfit concludes from this case.

Here is what Parfit concludes from this case.

Here’s another thought experiment: the case of the duplicate. Philosophers love thought experiments.
months they will have constructed a perfect duplicate of me out of organic matter. And this duplicate will wake up fully psychologically continuous with me, seeming to remember my life with my character, etc.

Vesey: Yes.

Parfit: Now in this case, which is a secular version of the Resurrection, we’re very inclined to think that the following question arises and is very real and very important. The question is, ‘Will that person who wakes up in three months be me or will he be some quite other person who’s merely artificially made to be exactly like me?’

Vesey: It does seem to be a real question. I mean in the one case, if it is going to be me, then I have expectations and so on, and in the other case, where it isn’t me, I don’t.

Parfit: I agree, it seems as if there couldn’t be a bigger difference between it being me and it being someone else.

Vesey: But you want to say that the two possibilities are in fact the same?

Parfit: I want to say that those two descriptions, ‘It’s going to be me’ and ‘It’s going to be someone who is merely exactly like me’, don’t describe different outcomes, different courses of events, only one of which can happen. They are two ways of describing one and the same course of events. What I mean by that perhaps could be shown if we take an exactly comparable case involving not a person but something about which I think we’re not inclined to have a false view.

Vesey: Yes.

Parfit: Something like a club. Suppose there’s some club in the nineteenth century . . .

Vesey: The Sherlock Holmes Club or something like that?

Parfit: Yes, perhaps. And after several years of meeting it ceases to meet. The club dies.

Vesey: Right.

Parfit: And then two of its members, let’s say, have emigrated to America, and after about fifteen years they get together and they start up a club. It has exactly the same rules, completely new membership except for the first two people, and they give it the same name. Now suppose someone came along and said: ‘There’s a real mystery here, because the following question is one that must have an answer. But how can we answer it? The question is, ‘Have they started up the very same club—is it the same club as the one they belonged to in England—or is it a completely new club that’s just exactly similar?’

Vesey: Yes.

Parfit: Well, in that case we all think that this man’s remark is absurd; there’s no difference at all. Now that’s my model for the true view about the case where they make a duplicate of me. It seems that there’s all the difference in the world between its being me and its being this other person who’s exactly like me. But if we think there’s no difference at all in the case of the clubs, why do we think there’s a difference in the case of personal identity, and how can we defend the view that there’s a difference?

Vesey: I can see how some people would defend it. I mean, a dualist would defend it in terms of a soul being a simple thing, but . . .

Parfit: Let me try another case which I think helps to ease us out of this belief we’re very strongly inclined to hold.

Vesey: Go on.

Parfit: Well, this isn’t a single case, this is a whole range of cases. A whole smooth spectrum of different cases which are all very similar to the next one in the range. At the start of this range of cases you suppose that the scientists are going to replace one per cent of the cells in your brain and body with exact duplicates.

Vesey: Yes.

Parfit: Now if that were to be done, no one has any doubt that you’d survive. I think that’s obvious because after all you can lose one per cent of the cells and survive. As we get further along the range they replace a larger and larger percentage of cells with exact duplicates, and of course at the far end of this range, where they replace a hundred per cent, then we’ve got my case where they just make a duplicate out of wholly fresh matter.

Vesey: Yes.

Parfit: Now on the view that there’s all the difference in the world between its being me and its being this other person who is exactly like me, we ought in consistency to think that in some case in the middle of that range, where, say, they’re going to replace fifty per cent, the same question arises: it is going to be me
or this completely different character? I think that even the most convinced dualist who believes in the soul is going to find this range of cases very embarrassing, because he seems committed to the view that there’s some crucial percentage up to which it’s going to be him and after which it suddenly ceases to be him. But I find that wholly unbelievable.

**Vesey:** Yes. He’s going to have to invent some sort of theory about the relation of mind and body to get round this one. I’m not quite sure how he would do it. Derek, could we go on to a related question? Suppose that I accepted what you said, that is, that there isn’t anything more to identity than what you call psychological continuity in a one-one case. Suppose I accept that, then I would want to go on and ask you, well, what’s the philosophical importance of this?

**Parfit** The philosophical importance is, I think, that psychological continuity is obviously, when we think about it, a matter of degree. So long as we think that identity is a further fact, one of the things we’re inclined to think is that it’s all or nothing, as you said earlier. Well, if we give up that belief and if we realize that what matters in my continued existence is a matter of degree, then this does make a difference in actual cases. All the cases that I’ve considered so far are of course bizarre science fiction cases. But I think that in actual life it’s obvious on reflection that, to give an example, the relations between me now and me next year are much closer in every way than the relations between me now and me in twenty years. And the sorts of relations that I’m thinking of are relations of memory, character, ambition, intention—all of those. Next year I shall remember much more of this year than I will in twenty years. I shall have a much more similar character. I shall be carrying out more of the same plans, ambitions and, if that is so, I think there are various plausible implications for our moral beliefs and various possible effects on our emotions.

**Vesey:** For our moral beliefs? What have you in mind?

**Parfit** Let’s take one very simple example. On the view which I’m sketching it seems to me much more plausible to claim that people deserve much less punishment, or even perhaps no punishment, for what they did many years ago as compared with what they did very recently. Plausible because the relations between them now and them many years ago when they committed the crime are so much weaker.

**Vesey:** But they are still the people who are responsible for the crime.

**Parfit** I think you say that because even if they’ve changed in many ways, after all it was just as much they who committed the crime. I think that’s true, but on the view for which I’m arguing, we would come to think that it’s a completely trivial truth. It’s like the following truth: it’s like the truth that all of my relatives are just as much my relatives. Suppose I in my will left more money to my close relatives and less to my distant relatives; a mere pittance to my second cousin twenty-nine times removed. If you said, ‘But that’s clearly unreasonable because all of your relatives are just as much your relatives’, there’s a sense in which that’s true but it’s obviously too trivial to make my will an unreasonable will. And that’s because what’s involved in kinship is a matter of degree.

**Vesey:** Yes.

**Parfit** Now, if we think that what’s involved in its being the same person now as the person who committed the crime is a matter of degree, then the truth that it was just as much him who committed the crime, will seem to us trivial in the way in which the truth that all my relatives are equally my relatives is trivial.

**Vesey:** Yes. So you think that I should regard myself in twenty years’ time as like a fairly distant relative of myself?

**Parfit** Well, I don’t want to exaggerate; I think the connections are much closer.

**Vesey:** Suppose I said that this point about psychological continuity being a matter of degree—suppose I said that this isn’t anything that anybody denies?

**Parfit** I don’t think anybody does on reflection deny that psychological continuity is a matter of degree. But I think what they may deny, and I think what may make a difference to their view, if they come over to the view for which I’m arguing—what they may deny is that psychological continuity is all there is to identity. Because what I’m arguing against is this further belief which I think we’re all inclined to hold even if we don’t realize it. The belief that however much we change, there’s a profound sense in

---

**Think about what Parfit is saying here. How is this applicable to the case of Tookie Williams?**
which the changed us is going to be just as much us. That even if some magic wand turned me into a completely different sort of person—a prince with totally different character, mental powers—it would be just as much me. That’s what I’m denying.

Vesey: Yes. This is the belief which I began by stating, and I think that if we did lose that belief that would be a change indeed.
Introduction to the Reading

By Dennis Weiss

As you begin to make your way through this reading, you will find Hume pointing to some of his basic philosophical principles: his commitment to empiricism, his skepticism, his distinction between ideas and impressions, his attempt to trace back all ideas to one or more impressions. Perhaps you encountered some of these ideas in other units of this course. Hume in fact starts this selection by raising the question of whether we have an idea or an impression of the self. Given that the self is supposedly a constant, simple, unchanging thing, do we ever have an impression of our selves? Hume suggests not. Why? What is his reasoning? But if the self does not exist, what am I? How does Hume describe the object of my awareness when I am aware of my self? You should connect Hume’s view to both his account of the bundle of perceptions
and his account of the mind as theatre. If the self does not exist, what gives rise to the idea of the self in the first place? This issue occupies Hume from paragraph 5 through to the end of this selection. He actually attempts to do two things. First, he wants to clarify what we mean by identity. Why does he consider all the examples of a ship, animals and vegetables, a noise, a rebuilt church? Hume is attempting to clarify our notion of “identity.” When Hume turns to a discussion of personal identity and attempts to explain what gives rise to our understanding of personal identity, he cites resemblance, causation, and memory as playing a role. Can you explain in your own words their role in our understanding of personal identity?

**A Treatise on Human Nature**  
*By David Hume*  

**Book I, Part 4, Section VI: Of Personal Identity**

1. There are some philosophers, who imagine we are every moment intimately conscious of what we call our SELF; that we feel its existence and its continuance in existence; and are certain, beyond the evidence of a demonstration, both of its perfect identity and simplicity. The strongest sensation, the most violent passion, say they, instead of distracting us from this view, only fix it the more intensely, and make us consider their influence on self either by their pain or pleasure. To attempt a farther proof of this were to weaken its evidence; since no proof can be derived from any fact, of which we are so intimately conscious; nor is there any thing, of which we can be certain, if we doubt of this.

2. Unluckily all these positive assertions are contrary to that very experience, which is pleaded for them, nor have we any idea of self, after the manner it is here explained. For from what impression coued this idea be derived? This question it is impossible to answer without a manifest contradiction and absurdity; and yet it is a question, which must necessarily be answered, if we would have the idea of self pass for clear and intelligible, It must be some one impression, that gives rise to every real idea. But self or person is not any one impression, but that to which our several impressions and ideas are supposed to have a reference. If any impression gives rise to the idea of self, that impression must continue invariably the same, through the whole course of our lives; since self is supposed to exist after that manner. But there is no impression constant and invariable. Pain and pleasure, grief and joy, passions and sensations succeed each other, and never all exist at the same time. It cannot, therefore, be from any of these impressions, or from any other, that the idea of self is derived; and consequently there is no such idea.

3. But farther, what must become of all our particular perceptions upon this hypothesis? All these are different, and distinguishable, and separable from each other, and may be separately considered, and may exist separately,
and have no Deed of tiny thing to support their existence. After what manner, therefore, do they belong to self; and how are they connected with it? For my part, when I enter most intimately into what I call myself, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch myself at any time without a perception, and never can observe any thing but the perception. When my perceptions are removed for any time, as by sound sleep; so long am I insensible of myself, and may truly be said not to exist. And were all my perceptions removed by death, and coued I neither think, nor feel, nor see, nor love, nor hate after the dissolution of my body, I should be entirely annihilated, nor do I conceive what is farther requisite to make me a perfect non-entity. If any one, upon serious and unprejudiced reflection thinks he has a different notion of himself, I must confess I call reason no longer with him. All I can allow him is, that he may be in the right as well as I, and that we are essentially different in this particular. He may, perhaps, perceive something simple and continued, which he calls himself; though I am certain there is no such principle in me.

4. But setting aside some metaphysicians of this kind, I may venture to affirm of the rest of mankind, that they are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement. Our eyes cannot turn in their sockets without varying our perceptions. Our thought is still more variable than our sight; and all our other senses and faculties contribute to this change; nor is there any single power of the soul, which remains unalterably the same, perhaps for one moment. The mind is a kind of theatre, where several perceptions successively make their appearance; pass, re-pass, glide away, and mingle in an infinite variety of postures and situations. There is properly no simplicity in it at one time, nor identity in different; whatever natural propension we may have to imagine that simplicity and identity. The comparison of the theatre must not mislead us. They are the successive perceptions only, that constitute the mind; nor have we the most distant notion of the place, where these scenes are represented, or of the materials, of which it is composed.

5. What then gives us so great a propension to ascribe an identity to these successive perceptions, and to suppose ourselves possest of an invariable and uninterrupted existence through the whole course of our lives? In order to answer this question, we must distinguish betwixt personal identity, as it regards our thought or imagination, and as it regards our passions or the concern we take in ourselves. The first is our present subject; and to explain it perfectly we must take the matter pretty deep, and account for that identity, which we attribute to plants and animals; there being a great analogy betwixt it, and the identity of a self or person.

6. We have a distinct idea of an object, that remains invariable and uninterrupted through
a supposed variation of time; and this idea we call that of identity or sameness. We have also a distinct idea of several different objects existing in succession, and connected together by a close relation; and this to an accurate view affords as perfect a notion of diversity, as if there was no manner of relation among the objects. But though these two ideas of identity, and a succession of related objects be in themselves perfectly distinct, and even contrary, yet it is certain, that in our common way of thinking they are generally confounded with each other. That action of the imagination, by which we consider the uninterrupted and invariable object, and that by which we reflect on the succession of related objects, are almost the same to the feeling, nor is there much more effort of thought required in the latter case than in the former. The relation facilitates the transition of the mind from one object to another, and renders its passage as smooth as if it contemplated one continued object. This resemblance is the cause of the confusion and mistake, and makes us substitute the notion of identity, instead of that of related objects. However at one instant we may consider the related succession as variable or interrupted, we are sure the next to ascribe to it a perfect identity, and regard it as enviable and uninterrupted. Our propensity to this mistake is so great from the resemblance above-mentioned, that we fall into it before we are aware; and though we incessantly correct ourselves by reflection, and return to a more accurate method of thinking, yet we cannot long sustain our philosophy, or take off this bias from the imagination. Our last resource is to yield to it, and boldly assert that these different related objects are in effect the same, however interrupted and variable. In order to justify to ourselves this absurdity, we often feign some new and unintelligible principle, that connects the objects together, and prevents their interruption or variation. Thus we feign the continued existence of the perceptions of our senses, to remove the interruption: and run into the notion of a soul, and self, and substance, to disguise the variation. But we may farther observe, that where we do not give rise to such a fiction, our propension to confound identity with relation is so great, that we are apt to imaginei something unknown and mysterious, connecting the parts, beside their relation; and this I take to be the case with regard to the identity we ascribe to plants and vegetables. And even when this does not take place, we still feel a propensity to confound these ideas, though we are not able fully to satisfy ourselves in that particular, nor find any thing invariable and uninterrupted to justify our notion of identity.

7. Thus the controversy concerning identity is not merely a dispute of words. For when we attribute identity, in an improper sense, to variable or interrupted objects, our mistake is not confined to the expression, but is commonly attended with a fiction, either of something invariable and uninterrupted, or of something mysterious and inexplicable, or at least with a propensity to such fictions. What will suffice to prove this hypothesis to the satisfaction of every fair enquirer, is to shew from daily experience and observation, that the objects, which are variable or interrupted, and yet are supposed to continue the same, are such only as consist of a succession of parts, connected together by resemblance, contiguity, or causation. For as such a succession answers evidently to our notion of diversity, it can only be by mistake we ascribe to it an identity; and as the relation of

Precisely how does Hume define "identity or sameness"?

Hume offers a hypothesis as to the origins of our idea of a soul or a self. What precisely is his hypothesis? What do you think about the hypothesis?
parts, which leads us into this mistake, is really nothing but a quality, which produces an association of ideas, and an easy transition of the imagination from one to another, it can only be from the resemblance, which this act of the mind bears to that, by which we contemplate one continued object, that the error arises. Our chief business, then, must be to prove, that all objects, to which we ascribe identity, without observing their invariableness and uninterruptedness, are such as consist of a succession of related objects.

8. [Not Included]
9. [Not Included]
10. [Not Included]
11. [Not Included]
12. [Not Included]
13. [Not Included]
14. [Not Included]

15. We now proceed to explain the nature of personal identity, which has become so great a question ill philosophy, especially of late years in England, where all the abstruser sciences are studied with a peculiar ardour and application. And here it is evident, the same method of reasoning must be continued. which has so successfully explained the identity of plants, and animals, and ships, and houses, and of all the compounded and changeable productions either of art or nature. The identity, which we ascribe to the mind of man, is only a fictitious one, and of a like kind with that which we ascribe to vegetables and animal bodies. It cannot, therefore, have a different origin, but must proceed from a like operation of the imagination upon like objects.

16. But lest this argument should not convince the reader; though in my opinion perfectly decisive; let him weigh the following reasoning, which is still closer and more immediate. It is evident, that the identity, which we attribute to the human mind, however perfect we may imagine it to be, is not able to run the several different perceptions into one, and make them lose their characters of distinction and difference, which are essential to them. It is still true, that every distinct perception, which enters into the composition of the mind, is a distinct existence, and is different, and distinguishable, and separable from every other perception, either contemporary or successive. But, as notwithstanding this distinction and separability, we suppose the whole train of perceptions to be united by identity, a question naturally arises concerning this relation of identity; whether it be something that really binds our several perceptions together, or only associates their ideas in the imagination. That is, in other words, whether in pronouncing concerning the identity of a person, we observe some real bond among his perceptions, or only feel one among the ideas we form of them. This question we might easily decide, if we would
recollect what has been already proud at large, that the understanding never observes any real connexion among objects, and that even the union of cause and effect, when strictly examined, resolves itself into a customary association of ideas. For from thence it evidently follows, that identity is nothing really belonging to these different perceptions, and uniting them together; but is merely a quality, which we attribute to them, because of the union of their ideas in the imagination, when we reflect upon them. Now the only qualities, which can give ideas an union in the imagination, are these three relations above-mentioned. There are the uniting principles in the ideal world, and without them every distinct object is separable by the mind, and may be separately considered, and appears not to have any more connexion with any other object, than if disjoined by the greatest difference and remoteness. It is, therefore, on some of these three relations of resemblance, contiguity and causation, that identity depends; and as the very essence of these relations consists in their producing an easy transition of ideas; it follows, that our notions of personal identity, proceed entirely from the smooth and uninterrupted progress of the thought along a train of connected ideas, according to the principles above-explained.

17. The only question, therefore, which remains, is, by what relations this uninterrupted progress of our thought is produced, when we consider the successive existence of a mind or thinking person. And here it is evident we must confine ourselves to resemblance and causation, and must drop contiguity, which has little or no influence in the present case.

18. To begin with resemblance; suppose we coued see clearly into the breast of another, and observe that succession of perceptions, which constitutes his mind or thinking principle, and suppose that he always preserves the memory of a considerable part of past perceptions; it is evident that nothing coued more contribute to the bestowing a relation on this succession amidst all its variations. For what is the memory but a faculty, by which we raise up the images of past perceptions? And as an image necessarily resembles its object, must not the frequent placing of these resembling perceptions in the chain of thought, convey the imagination more easily from one link to another, and make the whole seem like the continuance of one object? In this particular, then, the memory not only discovers the identity, but also contributes to its production, by producing the relation of resemblance among the perceptions. The case is the same whether we consider ourselves or others.

19. As to causation; we may observe, that the true idea of the human mind, is to consider it as a system of different perceptions or different existences, which are linked together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other. Our impressions give rise to their correspondent ideas; said these ideas in their turn produce other impressions. One thought chases another, and draws after it a third, by which it is expelled in its turn. In this respect, I cannot compare the soul more properly to any thing than to a republic or commonwealth, in which the several

Hume discusses the various powers of mind that lead us to suppose that the self exists as an uninterrupted thing.

Here Hume offers us a second analogy for thinking about the self: it is like a republic or commonwealth. In what ways is the self like a republic? Do you think this is a strong analogy?
members are united by the reciprocal ties of government and subordination, and give rise to other persons, who propagate the same republic in the incessant changes of its parts.

And as the same individual republic may not only change its members, but also its laws and constitutions; in like manner the same person may vary his character and disposition, as well as his impressions and ideas, without losing his identity. Whatever changes he endures, his several parts are still connected by the relation of causation. And in this view our identity with regard to the passions serves to corroborate that with regard to the imagination, by the making our distant perceptions influence each other, and by giving us a present concern for our past or future pains or pleasures.

20. As a memory alone acquaints us with the continuance and extent of this succession of perceptions, it is to be considered, upon that account chiefly, as the source of personal identity. Had we no memory, we never should have any notion of causation, nor consequently of that chain of causes and effects, which constitute our self or person. But having once acquired this notion of causation from the memory, we can extend the same chain of causes, and consequently the identity of car persons beyond our memory, and can comprehend times, and circumstances, and actions, which we have entirely forgot, but suppose in general to have existed. For how few of our past actions are there, of which we have any memory? Who can tell me, for instance, what were his thoughts and actions on the 1st of January 1715, the 11th of March 1719, and the 3rd of August 1733? Or will he affirm, because he has entirely forgot the incidents of these days, that the present self is not the same person with the self of that time; and by that means overturn all the most established notions of personal identity? In this view, therefore, memory does not so much produce as discover personal identity, by shewing us the relation of cause and effect among our different perceptions. It will be incumbent on those, who affirm that memory produces entirely our personal identity, to give a reason why we cm thus extend our identity beyond our memory.

21. The whole of this doctrine leads us to a conclusion, which is of great importance in the present affair, viz. that all the nice and subtile questions concerning personal identity can never possibly be decided, and are to be regarded rather as gramatical than as philosophical difficulties. Identity depends on the relations of ideas; and these relations produce identity, by means of that easy transition they occasion. But as the relations, and the easiness of the transition may diminish by insensible degrees, we have no just standard, by. which we can decide any dispute concerning the time, when they acquire or lose a title to the name of identity. All the disputes concerning the identity of connected objects are merely verbal, except so fax as the relation of parts gives rise to some fiction or imaginary principle of union, as we have already observed.
Are you a human being? A person? A soul? Something else altogether? Were you once a fetus? If so, think about the implications. This is the challenge that so-called “animalists” pose to our thoughts about personal identity. We have examined a variety of approaches to thinking through the problem of personal identity. But what if what you are is fundamentally not a person, at least in John Locke’s sense of being a consciousness. Consider your start in life as a fetus. Presumably each one of you was at some point in time a fetus. Now, of course, you’re 18 or 19 and your considerably bigger, heavier, and (we hope) wiser than you were when you were a fetus. Are you the same being now that started life out 18 or 19 years ago as a fetus? Think about the challenge this question poses to Locke’s view of personal identity. After all, the fetus that you were was not a person—it had no memories, no thoughts, likely no consciousness. So what makes you the same being as that fetus? Might it be that you are the same animal and that you what are most fundamentally is not a person or consciousness but
a particular human animal? In this reading, Paul Snowdon explores the animalism thesis that what you are is an animal. How does this present challenges to how we have been thinking about the core questions “who am I?” and “what am I?”?

The animal you are

In his discussion of personal identity, written in the seventeenth century but still, probably, the first piece about the problem that students read, Locke is concerned to draw a distinction between what makes something the same man and what makes something the same person (or self). Locke summarises his theory of persons in the slogan that “... personal identity consists .... In the identity of consciousness”. This is usually taken to mean that according to him a person’s life stretches back only to periods the person can remember, as it is said, “from the inside”. Locke is remarkably cagey about what the identity of a man consists in. This uncertainty derives from Locke’s, commendable for his time, professed ignorance as to whether there are spirits (or souls), and if there are, how such entities are involved in human life. However, Locke is clear that whatever is involved in human identity it is not a matter of consciousness. Locke’s theory then seems committed to what has recently been called the man/person distinction.

According to him, as one might put it, where I, the person, am there is a man (or human), but the person (or self) is not the same thing as the man, since they have different conditions of persistence. A thing’s “condition of persistence” is what is required for that thing to remain in existence. For example, it is a condition for a house to remain in existence that the bricks making it up stay together and are not scattered over the ground. It is generally assumed that people as a sort have their own persistence conditions, and so have animals.

Locke’s defence of the man/person distinction was so successful that philosophers simply took the contrast for granted, losing all interest in the notion of man or animal when discussing their own persistence conditions, and so representing any disagreement, to whatever extent, they had with Locke as disagreements about the concept of personal identity.

In what might be thought of as the recent classic period of discussion of personal identity, culminating in Parfit’s brilliant Reasons and Persons, together with the ensuing debate about it, the notion of an animal (or man) is virtually invisible. Now, this invisibility ended in the 1980s when a number of philosophers were independently struck by how unsatisfactory this neglect is. This insight struck, amongst others, David Wiggins, Michael Ayers, Eric Olson and myself in Britain, and Peter van Inwagen and William Carter in the United States. We responded to this thought in different ways, given the contrasting metaphysical frameworks we worked within, but the idea of us as animals had returned to a debate from which it might be said Locke had removed it.

Having noted how central the idea of man (or the human animal) is to Locke’s discussion, but how invisible it became in subsequent discussions of personal identity, it needs to be asked whether Locke’s conception of the relation between the person and the human animal is correct. The choice is between
holding that the human animal (or man) is a distinct thing from the person, and holding that the animal and the person are the same thing. The view that, contrary to the longstanding Lockean framework, they are the same thing, is the view currently called “animalism”. Before presenting some reasons which indicate, to varying degrees of strength, that adopting the distinctness thesis has real drawbacks, I want to clarify animalism and the questions to which it is offered as an answer.

Animalism is an identity thesis; it says that a certain thing is the same as another thing. Now, normally when philosophers discuss the problem of personal identity they are searching for an informative specification of what is involved by way of links across time in the persistence of a person. So a typical (but I do not mean to imply plausible) candidate answer would be: a person P who does X at time t remains in existence so long as there is a person P* who can recall, from the inside, doing X at t. Clearly, the animalist identity thesis does not have that structure. It does not explicitly pick out links over time. However, it does so indirectly, because we have what we might call a shared proto-theory of animal persistence, and animalism implies an answer to the normal problem of personal identity by requiring that persons fall under and conform to that proto-theory.

Of course there are disagreements among animalists as to what the correct theory of animals is. But what seems clear is that even normally minded animals do not cease to exist if their mental capacities are destroyed, so long as they remain physically intact and alive. In such circumstances they remain in existence even though damaged. So the animalist identity thesis implies that any claim that the survival of people requires the presence of mental links over time is incorrect. It also seems clear that no one could think that some processes which purport to preserve mental links over time, without preserving that physically substantial thing which is the animal, could seriously count as preserving the actual animal. The animalist should think, therefore, that simply generating mental links cannot be enough to ground real animal persistence, and hence not enough to preserve the persistence of the person as they think of the person. This explains how a thesis with the structure of animalism can imply substantial things about personal persistence.

A second clarification is needed. So far I have adopted a way of expressing animalism that regards it as saying that the person is the human animal. But if we are prepared to allow there might be entities which merit being described as persons who are not human – say God, or angels, or Martians, or robots, – then animalism should not rule them out. The content of animalism is better formulated as the claim that we, the persons hereabouts discussing these issues, are human animals. That is the nature that we have. Indeed, it is fairly clear, I suggest, that questions about the nature of persons as they interest us really are questions about the nature of ourselves. I have argued that we should understand animalism as a thesis about ourselves, rather than a thesis about persons in general, and also that, despite its form as an identity thesis, it has significant implications about the conditions of what is called personal identity.

The third clarification is that the animalist identity claim (we are identical to animals) should not be thought of as solely, or mainly, a thesis about what is known as “personal identity over time”. One reason for not saying that it is a theory of personal identity full stop is that, as we have seen, animalists can disagree about the persistence conditions of animals. More importantly, we should think of animalism as a general, so to say, metaphysical characterisation of our nature, with myriad implications.

For example, if we are animals, it is also true to say that we are subjects of experience, and so it implies that there need be no more unity to the mind of a subject than there need be to the mind of an animal. On the face of it, an animal’s mind can be radically dis-unified – imagine the sorts of surgical interventions that can break connections consistent with the animal remaining in existence and having
experiences. So animalism also leads to a distinctive way of thinking about subjects and psychological unity.

Now I want to present some reasons that incline some of us to accept animalism. There is something artificial in dividing the reasons into those in favour and then subsequently considering those against. Part of the appeal of animalism, assuming it has an appeal, derives from a sense that the grounds that have convinced philosophers that animalism is hopeless are not as strong as their standard assessment supposes. However, in a preliminary presentation things simply have to be artificial. Now, I am assuming, as I hope any reader will allow me to, that where each reader of this article is, there is also a human animal. Animalism claims that the person reading and the animal (presumably also reading) are one and the same thing. Now, if the man/person distinction is accepted it seems to follow that the same space is occupied simultaneously by two distinct things – the animal and the person.

This is sometimes called the thesis of the possibility of coincidence – the idea that two distinct things can coincide in the space they occupy. It is a thesis that strikes some as absurd. How can there be two things precisely coinciding in a single space (possibly, indeed, throughout their entire histories)? Now, if it is absurd then the conventional man/person distinction would have to be wrong, and there would be considerable pressure in favour of animalism. However, although this is a line of argument that has persuaded some, it rests on the conviction that coincidence is not a possibility, which is a conviction that is not obviously correct. So this reason strikes me as unreliable.

What is more striking, I believe, is that on any normal conception of them, there is a massive degree of similarity between the individual human animal and the individual person who occupy the same place. They seem to have the same spatial extension, the same parts, the same history (starting at the same point, and ending at the same point), the same causal roles – they seem to be the same in more or less every specific way. For example, if the animal is breathing then the person is breathing, and vice versa. Now, the most obvious explanation for this virtually exception-less similarity is that the animal and the person are, in fact, the same thing.

The weakness of this argument, though, is obvious. Opponents of animalism hold that there are differences between the human animal and the person. What these are alleged to be I’ll consider shortly, but they will reject the basic premise of the present argument. There is, according to them, no complete similarity, and so no explanation is needed. This reservation does mean that there is no proof here, but it can be suggested that a weaker conclusion looks plausible. Given the manifest extent of the similarity between person and animal, the animalist claim seems to be the view that counts as the default view. Discussion should start from it, and if we are not to accept it some strong reasons need to be provided. If that is correct the neglect of the view for quite a period must count as extremely regrettable.

A third reason or type of reason in favour of animalism is that there are difficulties in the idea of the man/person distinction. There are different ways to develop this idea, and here I shall sketch two closely related versions of the problem. The problem starts from the question whether the animal has mental states. It is obvious that the person or self does, but does the animal?

It surely seems quite clear that the human animal does have mental states, so initially I want to assume that is the correct answer to our question. Indeed, it is this conviction that fuels the appeal of animalism. What mental states does the animal have? It is hard to deny that the animal itself has the same complex array of mental states that the person has. The human animal can think about itself, can talk, can reason, as well as have experiences and perform actions. If that is so then maintaining the man/person
dis1nc1on seems to carry the implication that there are two subjects of experience and two mental lives being lived. Now, that is not a verdict that we accept as we start thinking about the nature of persons (or ourselves).

But the problem seems worse than this. On normal conceptions of what qualifies someone as a person, it is that they have (or are capable of having) the higher mental functions – such as self-consciousness and reason. These are the mental capacities that Locke cites in his famous and resonant elucidation of what persons are. But it seems that we have to credit these higher capacities to the human animal, in which case accepting the man/person distinction implies that there are two persons – as one might say, the non-animal person to be distinguished from the animal, and the personal animal. It would seem highly paradoxical to hold that there are two persons. We seem to have derived two paradoxical consequences from the man/person distinction. Clearly the opponents of animalism need to do some serious explaining at this point.

Here is a final line of thought. I have suggested that the question about the relation between the person and the animal can be viewed as a question about ourselves. How am I related to this animal? That is in effect what I want to know. We can get some purchase on the question if we can determine what “I” refers to, as used in that question. Does it refer to the animal or not? But it seems to help with that question to ask how it comes about that there are “I” users in the first place. As we might say, the “I” in the crucial question refers to that thing, whatever it is, that “I” talk exists in order to refer to. But it would seem to be a plausible principle that such cognitive/linguistic devices have emerged in the course of evolution because they enable advanced animals to self-refer. We naturally think of such cognitive devices as akin to, say, other devices that natural selection has generated and preserved; they are preserved because they yield functions for animals that animals benefit from. This means that they are devices for self-reference by the animals that are lucky enough to possess them. In this way, reflections on the emergence of self-reference indicates that the subject of self-reference is the animal. Which is to say that it indicates that the animalist identity should be accepted.

This amounts, then, to a sketch of an argument from general biology to metaphysics. This may seem a surprising route to a metaphysical claim, but philosophers often appeal to considerations about language to support substantive (non-linguistic) conclusions, and in this case we are relying on empirical considerations about the origins of language, rather than brute semantic intuitions. Plainly, questions can be put to this line of thought, but it is not, I suggest, absurd.

There are then strong reasons in favour of animalism. However, I think that it is true to say that most philosophers do not accept it. Why? One way to think of the ground of opposition is to notice that animalism is an identity thesis. The person is the animal (where the person is). Now, an identity cannot be accepted if there is a (detectable) property difference between the two items. The ground for opposition is the general conviction that there are property differences between the person (or self) and the animal.

As we have seen there is a massive overlap or congruence of properties if we survey what we might think of as the ordinary features that our lives exhibit. But what philosophers have tended to think is that there are detectable differences in relation to what is possible for the person and for the animal. These differences come out when considering what might be called possible, even if not actual, dissociation cases. There are fundamentally two kinds of such cases. One sort of possible case is where we start with the person and the animal occupying the same space at the same time, but as the case develops the person is counted as ceasing to exist whereas the animal remains. One candidate for such a case would be where someone suffers a traumatic injury which makes them what we sometimes call a “human
vegetable". There is life (perhaps artificially sustained), and so surely an animal, but no mental capacity. About such cases many judge that the person no longer exists.

Another type of example that has been suggested fitting this pattern is what used to be called cases of multiple personality disorder. As popularly conceived in such occurrences there is associated with a certain human animal a definite personality, linked to a battery of memories and values, and a name to which they answer. After some time this psychological syndrome vanishes and is replaced in the same animal by contrasting personality, set of memories and values, and self professed new name. There is one animal but it is proposed that one person (or self) has been replaced by another. Cases of this sort can be called \([A \text{ and not } P]\) cases – indicating that they are supposed to be examples where there is the same animal but not the same person.

But the sorts of cases that have had most influence on philosophers are ones which can be called \([P \text{ and not } A]\) cases. These purport to be cases where we start with a (coincident) person and animal but which develop in such a way that the person remains but the animal is no longer there. The example that has seemed most convincing is that of brain transplants. The standard example was first set out by Sydney Shoemaker. In his example we start with two people, Brown and Robinson. Both have their brains removed (for medical reasons) but, by mistake, Brown’s brain is returned to Robinson’s body, and to avoid extra complications Robinson’s brain perishes. It seems that we now have one person whom we can name Brownson. But is it Brown or is it Robinson?

Shoemaker suggests that the correct judgement is that Brownson is in fact Brown. Assuming that our psychological character is preserved in the brain, we can say that Brownson will remember Brown’s life, have Brown’s character and beliefs, and Brown’s values, and so on, and will be sure that he is Brown. This psychological similarity and continuity between Brown and Brownson, according to Shoemaker and most philosophers, makes it the case that Brownson is Brown. In which case we have the same person, but on the face of it Brown is no longer housed in or attached to the same human animal. So there is the same person but not the same animal. Now, there are, of course, other candidate \([P \text{ and not } A]\) examples – for example, head transplants – but brain transplants have seemed very powerful anti-animalist cases.

The initial anti-animalist argument derives then from the conviction, encouraged by reflection on such cases, that the person and the animal can “come apart”. If so, they can hardly be the same thing. If the arguments on both sides have struck you as powerful and plausible then we have to acknowledge that we have the task of evaluating them and also assembling, if we can find it, new evidence to determine what the truth is about our relation to the animals where we are. I’ll now raise some questions about the anti-animalist case that has been sketched. On the whole, in these anti-animalist examples, it is not really in doubt what the correct thing to say about the animal is. We feel confident that we can trace it in the developing story. The issue that is more unclear is what is happening to the person. What is it right to say about the person in such cases?

What of the \([A \text{ and not } P]\) examples? There is no general demonstration against the possibility of such cases, and so it is a matter of deciding what to say case by case. It helps, it seems to me, to consider such stories by imagining that they involve people who matter to you. So imagine your father suffers the accident. Would you seriously deny that he, your father, was himself still alive and in the hospital bed? When the hospital authorities ask you when you will be visiting your father would you really react by asking who they are talking about, since your father no longer exists? It seems to me you would not. Next, imagine in the second example that it is your brother. Do you seriously have any inclination to think that your brother’s whereabouts are unknown and that talking to you is a totally new person?
Rather, we think of such abnormal psychological developments as illnesses that befall people. Our aim is to cure the people to whom they happen. More needs to be said, but despite their popularity in philosophical folk-lore such examples do not really seem like genuine [A and not P] cases. The [not P] description is implausible.

Such [A and not P] cases seem, on reflection, fairly weak because we have an established way of thinking of them according to which they really concern the same person, even if it is a way we tend to lose sight of when philosophising. With suggested [P and not A] cases the situation is different. We do not have an established way of thinking about them, since they do not occur. The fundamental question is whether opponents of animalism are right to be confident as to what is correct to say about such examples. Once we ask this, possible reasons for being less confident suggest themselves. First, even if more or less everyone initially regards the person as going with the brain, it has to be acknowledged that the final verdict needs to be made in the light of all the evidence, and so it may be that some initial and quite strong convictions have to be discarded. I have not investigated here lines of reply to the pro-animalist arguments – and that is a weakness in the overall strategy of argument here – but they also surely seem strong and convincing. So there is no entitlement to insist that the brain transplant intuition just has to be accepted.

Second, we can ask whether we are entitled to trust what I am calling our brain transplant intuitions. We can distinguish two broad conceptions of our conceptual practice of tracing ourselves over time. On one conception we do so in virtue of having an understanding of what our kind of persistence involves, an understanding that allows us to generate reliable verdicts about all sorts of imaginary cases. The second conception is that we are entities that are able to trace themselves in the world we find ourselves in, but that we must fix what the nature of our persistence is by investigating what kinds of things we in fact are tracing. Then, but only then, can we determine verdicts about such merely imaginable cases. This is a rough contrast, but people who feel confident about what to say in brain transplant cases need to consider whether such confidence is based on a proper understanding of our conceptual practices.

Third, an alternative way of thinking about such cases can be proposed, and it may not be without merit. Organs can be transplanted. What was one animal’s liver can become another animal’s liver. Similarly, the organ of cognition (or mentation) which sustained cognition for one thing might be able to become the organ of cognition for another creature. On this way of thinking it is not that brain transplants transfer people; rather a single organ for doing something goes from one person to another. Robinson may get a new brain, just as he can get a new liver. Philosophers who attach supreme importance to thought and cognition may find this an unnatural way of thinking, but that may not be how it does, and should, strike everyone.

I have been trying to put some question marks against some anti-animalist arguments. But it is obvious that these moves constitute merely the first moves in a complex debate, and in some respects I have not even sketched the first moves of the debate. What, I believe, we need to cultivate in explorations of our own nature and persistence conditions is the ability to resist being swept away from solid and clear ways of thinking into realms of fantasy, where more or less anything goes. But it is too early yet to say whether animalism is the view that avoids fantasy.

Paul Snowden is Grote Professor of Mind and Logic at University College London. His forthcoming books include a collection of his own papers entitled Essays on Perceptual Experience, and, jointly edited with Stephen Blatti, Essays on Animalism. Both will be published by Oxford University Press in 2012.