



## SELF-IDENTITY

*"I'm not myself today, you see," Alice said to the caterpillar.*

*"I don't see," said the caterpillar.*

LEWIS CARROLL

"Just be yourself!" How often have you heard that? What is it to be a "self"? And what does it mean to be a particular self? In the abstract, these questions seem obscure. But in everything that we do, we adopt some conception of our identity, both as a person and as an individual, whether we are called upon to articulate it or not. As a student, you walk into a classroom with certain conceptions of your own abilities and intelligence, your status among other students, your role vis-à-vis the professor, some haunting memories, perhaps some embarrassment or a certain vanity about your looks, your clothes, your grades, or just your new pair of shoes.

If you had to identify yourself as an individual, describe what makes you you, how would you do it? What features are essential to being a person? What are essential to being the person you are and what features distinguish you from other persons? Think for a moment of yourself in an office, applying for a job or a scholarship for professional school or just filling out one of those dozens of forms that bombard you during the year. You dutifully fill in your birthdate, where you were born, your grades in school, your service in the military if any, awards you have received, arrests and other troubles, whether you're married or not, male or female, perhaps your race and religion. This list of facts about yourself would be one way of identifying "you." But at some point, I am sure you have felt that sense of absurdity and rebellion, "this isn't me!" or "this is all irrelevant!" What may seem more relevant to your self-identity are your political views, your tastes in art and music, your favorite books and movies, your loves and hates, habits and beliefs, or just the fact, perhaps,

that you think your own thoughts. These more personal, or “internal,” features as well as the “cold facts,” or “external” features, about yourself are important for identifying you as an individual different from other individuals.

In one sense, your **self-identity** is the way you characterize yourself as an individual. The philosophical problem of self-identity is thus concerned in part with what these characterizing qualities are. Are they just concerned with status and roles among other people? Or is there something that can truly be called “your self,” your “essence,” or maybe even your soul, without reference to anyone else? Should we think of ourselves as individuals? Or should we instead view ourselves as mere components of a larger organism—society, mankind, or perhaps the world as a whole? How should we identify ourselves?

There is another sense in which we might talk about a person’s self-identity. You may know someone who has experienced a religious conversion, or who has just undergone treatment for alcoholism; or perhaps you can think of someone you see again after a long time who says, “I’m not the same person.” What does this mean, and how is it possible to say you are not the same person? When someone says, “I’m not the same person I was,” he or she is pointing to the fact that some significant aspects of himself or herself have changed; this person has a new self-identity. Yet, in another sense, this is still the same person; the old identity and the new identity are both identities of the same person.

So we have a second sense of “self-identity.” Here your self-identity is what makes you the same person over time. Thus, the second philosophical problem, which is the one that has most concerned philosophers, is how to identify an individual as the *same* individual over time. What is it about you without which you wouldn’t be you? Presumably you would still be you if you changed your religious beliefs or tastes in music. But what if you had a sex-change operation? Or what if you completely lost your memory, all recollection of your family and friends? Or what if you physically disappeared altogether, remaining only a wispy consciousness, a spirit, or a ghost without a body? Would it then make sense to say that you are still you?

There is yet a third sense in which we talk about self-identity. What is it that *allows* us to be individual people at all? Interestingly, the fact of consciousness—of having thoughts and feelings—seems the most private and individual thing a person possesses, and yet is the very thing that we tend to believe *all* people possess. Whatever our attempts to answer the problems of self-identity, they begin with a single “fact”—our own consciousness. We will see that this was the logical beginning for Descartes, who used the fact of his own consciousness as the starting point for his whole philosophy. It was true for Locke, as well, who argued that our identity is to be found in the continuity of our consciousness rather than in the continuity of our bodies. We will see that it was even used by Hume, who used his own consciousness as the basis of his denial that there is any such thing as the self!

Of course, these questions assume some metaphysical claims that can and have been questioned. Later in this chapter, we will look at some of these assumptions and some criticisms of them. Is self-identity a matter of nature? Are we born with an identity? Or is it a matter of personal choice or of environmental factors such as our

upbringing or education? Must a person have only one “self,” or might he or she in fact have several or many selves? Is there a “self” at all? Or is the self, as many Eastern philosophers have argued, an illusion? Should self-lessness be our ideal self-identity? Ought we even to conceive of ourselves as individuals? Or ought we instead think of ourselves as organic components of a larger community or society?

In this chapter, we explore a number of different questions and conceptions about self-identity. We consider the question of what makes you *you*, both in the sense of what makes you a person, and in the sense of what makes you the particular person you are. Freud, the great nineteenth-century psychologist, once said that every man is in some ways like all other men, in some ways like some other men, and in some ways like no other man. Philosophers who investigate the problem of self-identity attempt to figure out exactly how and why this might be so. We begin with an all-important and still influential tradition in philosophy, from Descartes and Locke to Hume and Kant, which focuses on self-consciousness as the sole key to personal identity.

## A. CONSCIOUSNESS AND THE SELF: FROM DESCARTES TO KANT

---

Sometimes we act out our identities without being aware of it at all. At other times, particularly when we talk about ourselves or are placed in a situation where we are forced to “look at ourselves as others see us,” we are very much aware—even painfully aware—of our identities. At such times we say we are **self-conscious**. In general, most modern philosophers and psychologists would argue that you can’t have a concept of who you are unless you are also sometimes (not necessarily always) self-conscious. Conversely, you can’t even be self-conscious unless you have some sense of identity, no matter how crude. The two concepts, in other words, go hand in hand and cannot be separated from each other.

Many philosophers have argued that not only is self-consciousness crucial to having a concept of one’s own individuality, but it is also crucial for establishing that one is an enduring self, that is, the same person over time. Descartes is an example. Remember how he characterized himself:

*But what then am I? A thing which thinks. What is a thing which thinks? It is a thing which doubts, understands, affirms, denies, wills, refuses, which also imagines and feels.*

He goes on to say, I am a thing with desires, who perceives light and noise and feels heat. Clearly, Descartes’ concept of “self” is of thought, or consciousness—a human essence—which each person has and with which each person identifies himself or herself. He also claims to show by his method of doubt that all of this might be so even if he were not to have a body at all. Perhaps, he argues, I am fooled about my

“having” a body just as I might be fooled about all sorts of other things. Therefore, he concludes, it is not my body that provides me with an identity or with the self from which I begin my philosophy. It follows from this that the particular aspects of my self—whether I am male or female, black or white, tall or short, handsome or ugly, strong or weak—are associated with my body only and cannot be essential to my identity. My self-identity is in my mind, in my thinking, doubting, feeling, perceiving, imagining, and desiring. I am, essentially, “a thing which thinks.”

FROM “**MEDITATION VI,**”  
BY **RENÉ DESCARTES**

Therefore, just because I know certainly that I exist, and that meanwhile I do not remark that any other thing necessarily pertains to my nature or essence, excepting that I am a thinking thing, I rightly conclude that my essence consists solely in the fact that I am a thinking thing [or a substance whose whole essence or nature is to think]. And although possibly (or rather certainly, as I shall say in a moment) I possess a body with which I am very intimately conjoined, yet because, on the one side, I have a clear and distinct idea of myself inasmuch as I am only a thinking and unextended thing, and as, on the other, I possess a distinct idea of body, inasmuch as it is only an extended and unthinking thing, it is certain that this I [that is to say, my soul by which I am what I am], is entirely and absolutely distinct from my body, and can exist without it.<sup>1</sup>

It is important to appreciate the kind of step Descartes has taken here. What he is saying is that self-identity depends on consciousness. Our identity does not depend in any way on our body remaining the same, and so human identity is different from the identity of anything else in the world.

John Locke, like Descartes, sees self-consciousness as the key to self-identity. But unlike Descartes, he argues that this identity does not depend on our remaining the same thinking substance, that is, on our having the same soul. Indeed, in the course of our life our soul might be replaced with new souls just as in the course of a tree's growth its cells are replaced with new cells. What makes the tree the same tree is the fact that the same life is present in spite of changes in its physical structure; and what makes a person the same person is that the same *consciousness* and memories are present. Thus Locke differs from Descartes in distinguishing between the soul (a substance) and consciousness. It is our consciousness that we call our “self.” In *An Essay Concerning Human Understanding*, Locke argues:

---

<sup>1</sup> René Descartes, “Meditation VI,” in *Meditations on First Philosophy*, in *The Philosophical Works of Descartes*, trans. Elizabeth S. Haldane, and G. R. T. Ross (Cambridge: Cambridge University Press, 1911).

# ON PERSONAL IDENTITY, BY JOHN LOCKE

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.

*Consciousness makes personal Identity.*—But it is further inquired, whether it be the same identical substance. This few<sup>2</sup> 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

---

<sup>2</sup>Locke refers here to Descartes.

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.

*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, are 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.



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. 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 the pre-existing 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 of learning;)—would any one say, that he, being not conscious of any of Socrates' 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*.

*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.

*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 I 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.

*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 the 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.<sup>3</sup>

The main thesis of Locke's argument is this: Personal self-identity is based upon self-consciousness, in particular, upon memories about one's former experiences. In this, he argues, man is different from animals, whose identity (that is, "the same dog" or "the same horse") is based on the continuity of the body, just as you would say that you have had "the same car" for ten years even if almost every part except the chassis has been replaced during that time. The identity of a "person," that is, "personal" identity, depends on self-consciousness.

Locke's idea that memory is what constitutes a self-identity is inspired by the distinctly Cartesian notion that a person's relationship to her own thoughts is unique. You cannot think my thoughts and I cannot think yours. Since memory is a species of thought, it follows that you cannot remember my experiences, nor I yours. For example, you may remember your first day of school. Because *you* are remembering that experience as one that happened to *you*, you are self-identical to the person who had the earlier experience. According to Locke, then, memory provides an infallible link between what we might call different "stages" of a person. Memory seems to guarantee the identity of the person who is now remembering with the person who was then having the experience.

While Locke's theory has the advantage over Descartes' of allowing us to understand self-identity in terms of consciousness without requiring that we posit the existence of a persisting immaterial soul, it is not without its own difficulties. First, much of what we experience, we later forget. Do you remember everything that has ever happened to you? Undoubtedly, the answer is no. Even a person who has a very good memory does not remember being born, learning to walk, or what he had for breakfast on June 3, 1982. Probably you have completely forgotten some fairly long stretches of your past. Does it follow that you did not exist during those stretches? According to Locke's theory, it is not clear that *you* did.

Second, our memories are not always accurate. Sometimes we remember things that never happened. For example, you might remember very clearly lending your copy of *Introducing Philosophy* to a friend, only to find later that you had, in fact,

<sup>3</sup>John Locke, *An Essay Concerning Human Understanding*, ed. A. C. Fraser (Oxford: Clarendon Press, 1894).



been using it as a doorstep. Even more disturbing, though fortunately less frequent, cases of inaccurate memory occur when a person sincerely remembers experiences which, in fact, happened, but didn't happen to him. There are people who now remember delivering the Gettysburg Address, discovering radium, or singing "I wanna hold your hand" with George, John, Paul, and Ringo at Carnegie Hall. What are we to say of the memories of the deluded and the deranged? They are not *genuine* memories, but are only *apparent* memories. Clearly, Locke did not intend merely apparent memories to count among those which guarantee identity. We must, then, find a way to distinguish between those cases in which a memory is genuine and those in which it is not. But to do this, it seems that we would have to say that the memories are in fact the correct memories *of that person*. If this is so, it would appear that the Memory Theory is circular.

A genuine memory, as opposed to a merely apparent one, is, of course, a memory of an experience the rememberer actually had. The person who is having the memory must be the one who had the experience. Now you can see that in distinguishing genuine from apparent memory, we have presupposed the existence of a persisting, self-identical person. That would be all very well were it not for the fact that the concept of self-identity is precisely what we are trying to explain. We cannot use the concept of memory to explain self-identity and then use the concept of self-identity to explain memory. Moreover, once we reflect on the nature of genuine memory, we can see that Locke was, indeed, putting the cart before the horse. When a person says, "I remember when I learned to ride my bike," the truth of his statement presupposes, rather than establishes, that he is self-identical to the little boy with the scabby knees.

Nevertheless, modern theories of personal identity have by and large appealed to some notion of memory, self-consciousness, or psychological continuity. It is worth mentioning that ancient philosophers, Aristotle for example, did not believe this and probably would not even have understood much of what Descartes and Locke were arguing about. For Aristotle, self-identity was essentially bodily identity, without any particular reference to self-consciousness. But neither must it be thought that the modern thesis has been without its critics, in fact, its devastating critics.

Hume completely undercuts Descartes' and Locke's view of self-identity. Relying on his belief that any idea must be derived from an impression, Hume argues that when we are self-conscious we are only aware of fleeting thoughts, feelings, and perceptions; we do not have an impression of the self or a thinking substance. He concludes that the idea of the self is simply a fiction. Moreover, since we are never aware of any enduring self, we are never justified in claiming we are the same person we were a year or a minute ago.

### **"THERE IS NO SELF,"**

BY DAVID HUME

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 deriv'd 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.

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 explain'd. For from what impression cou'd this idea be deriv'd? This question 'tis impossible to answer without a manifest contradiction, and absurdity; and yet 'tis a question, which must necessarily be answer'd, if we wou'd 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 suppos'd to have a reference. If any impression gives rise to the idea of self, that impression must continue invariably the same, thro' the whole course of our lives; since self is suppos'd 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 deriv'd; and consequently there is no such idea.

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 consider'd, and may exist separately, and have no need of any 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 remov'd 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 remov'd by death, and cou'd I neither think, nor feel, nor see, nor love, nor hate after the dissolution of my body, I shou'd be entirely annihilated, nor do I conceive what is further requisite to make me a perfect nonentity. If any one upon serious and unprejudiced reflexion, thinks he has a different notion of *himself*, I must confess I can 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 continu'd, which he calls *himself*; tho' I am certain there is no such principle in me.

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, repass, 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 compos'd.

What then gives us so great a propension to ascribe an identity to these successive perceptions, and to suppose ourselves possessors of any invariable and uninterrupted existence thro' the whole course of our lives?<sup>4</sup>

To answer this question, Hume draws an analogy between the fictitious identity we ascribe to persons and the equally fictitious identity we ascribe to things. Just as we can never find an impression of the self that will explain human identity, so we can never find an impression of an object or substance to explain the identity of plants, animals, and things. According to Hume, then, we are never justified in claiming that, for example, a tree we see now is the same tree we saw five years ago or even five minutes ago. The cells and parts of the tree are continuously being replaced so that at no time is it ever literally the same tree. But Hume's argument goes further than this; even if that were not so, we would still have no way of justifying our belief that this tree is the same one we saw some time ago, rather than another, reasonably similar to it, but yet different. How do we know, for example, that someone has not come along and replaced it with another?

The temptation to ascribe identity to things and persons, Hume thinks, arises in part from the spatiotemporal continuity of the thing; the tree is in the same place at different times. People, however, have the troublesome habit of moving around, going to Europe for the summer or college for the semester; we still see the continuity of his or her movement, receive postcards from the appropriate places at the appropriate times, and so we conclude that it is the same person. In addition to spatiotemporal continuity, we ordinarily rely on resemblance as a criterion of identity. We tolerate small changes, a haircut or a new scar, perhaps even a lost leg or a bit of plastic surgery. As long as there is a strong resemblance between two individuals, for example before and after a haircut, we think of them as the same. Only when there is a great change, as from Dr. Jekyll to Mr. Hyde, do we question the identity of the two individuals.

Hume, however, argues that spatiotemporal continuity and resemblance do not in fact guarantee identity:

---

<sup>4</sup>David Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge (Oxford: Oxford University Press, 1888).

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 his objects.

We have a distinct idea of an object, that remains invariable and uninterrupted thro' a suppos'd 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 tho' these two ideas of identity, and a succession of related objects be in themselves perfectly distinct, and even contrary, yet 'tis 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 to thought requir'd 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 continu'd object. This resemblance is the cause of the confusion and mistake, and makes us substitute the notion of identity, instead of that related objects.

Our last resource is to . . . 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 continu'd 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 imagine 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, tho' we are not able fully to satisfy ourselves in that particular, nor find any thing invariable and uninterrupted to justify our notion of identity. . . .

Suppose any mass of matter, of which the parts are contiguous and connected, to be plac'd before us; 'tis plain we must attribute a perfect identity to this mass, provided all the parts continue uninterruptedly and invariably the same, whatever motion or change of place we may observe either in the whole or in any of the parts. But supposing some very *small* or *inconsiderable* part to be added to the mass, or subtracted from it; tho' this absolutely destroys the identity of the whole, strictly speaking; yet as we seldom think so accurately, we scruple not to pronounce a mass of matter the same, where we find so trivial an alteration. The passage of the thought from the object before the change to the object after it, is so smooth and easy, that we scarce perceive the

transition, and are apt to imagine, that 'tis nothing but a continu'd survey, of the same object.<sup>5</sup>

Hume's argument is familiar to us from our discussion of empiricism in Chapter 2. All we perceive, he says, is a sequence of impressions, and nowhere do we encounter an impression either of a substance (an enduring object) or of the self. What right do we have, therefore, to identify the object of this impression with the object of another? What right do we have to identify the person we are now with someone in the past?

But Hume's argument that "I never can catch myself" suffers from a peculiar but obvious form of self-contradiction. He can't even deny that there is a self without in some sense pointing to himself in order to do it. This point was not missed by Kant. Kant agrees with Hume that the enduring self is not to be found in self-consciousness. The enduring self is not an object of experience—Hume was right on this point and both Descartes and Locke were mistaken. In Kant's words, the enduring self is not empirical. It is transcendental.

By "transcendental" Kant means what is a necessary condition for the possibility of *any* experience. Kant saw that if there were a different self at each moment of consciousness, we would not be able to perceive anything. In order to experience an object, we must be able to combine our various impressions of it in a unified consciousness. Thus, if we do in fact experience objects, we must assume that we have a unified consciousness that combines these impressions into the perception of an object. Or, to take a different example, Hume talks of different sorts of relations of impressions, for example, succession. In order for an individual to perceive two impressions as successive, these impressions must have been perceived by the *same* consciousness.

The self of "I" for Kant, then, is the necessary logical subject of any thought, perception, feeling, and so on. It is not an object of experience but transcends and is presupposed by all experience.

Hume's error, as in other matters, was in confusing the supposed experience of self-consciousness with the transcendental rules with which we tie these various experiences together. Accordingly, Kant argues, Descartes and Locke are both correct in equating self-identity and self-consciousness, but it must not be thought that the self is therefore a "thing" (as Descartes said) that we find in experience.

According to Kant, the self is the activity of consciousness, in particular the activity of organizing our various experiences. Kant borrows Hume's argument, but he turns it toward the opposite conclusion: True, I never find a self "in" my experiences, but I can always find myself in that "I" that has the experience. Kant's "self," in other words, is the act of having experiences rather than anything that we experience itself. But for Kant this self is not merely the passive recipient of experiences, and here is where the notion of self as activity becomes all-important. The self is the activity of

---

<sup>5</sup>Hume, *A Treatise of Human Nature*.

applying the rules by which we organize our experience. Moreover, Kant argues that one of the most basic rules of this activity is that the self organize its experience as its own experience. The rule is that we must always “synthesize” our various experiences into a unity, for we could not come to have any knowledge whatever of a scattering of various impressions and sensations without this synthesis. (Think, for example, of the slightly painful and generally meaningless sensations you have when someone sets off a camera flashbulb in your eyes.) This basic rule of synthesis allows Kant to say that not only is the self the activity that applies the various rules to experience but its existence as a unified self with a unified synthesis of experiences is itself a rule.

Kant gives this curious idea of the self as a rule a formidable name, “the transcendental unity of apperception.” What is important in this concept is that the self for Kant is indeed essential to self-consciousness, but it is not “in” self-consciousness. Metaphorically, it is often said that it is “behind” self-consciousness, that is, it is the activity of bringing our various experiences together in accordance with the basic rules of our experience. Accordingly, Kant refers to this self as the **transcendental ego**, “transcendental” because, as we discussed in Chapter 3, it is basic and necessary for all possible human experience. The difference between Hume and Kant is sometimes illustrated in this way: Hume looks for the self among our experiences and doesn’t find it; Kant agrees with Hume but argues that he looked in the wrong place. The self, Kant says, is the thread that ties together our various experiences. Accordingly, the self is not in the bundle of our experiences; it is rather the “transcendental” thread that holds them all together and is as real as any experience.

Kant returns to Descartes and challenges his main theses, even while agreeing with parts of them. First, while Descartes thought that we had to be self-conscious all the time, Kant insists that it is only necessary for “the ‘I think’ to be able to accompany all experiences.” It is not necessary to be always conscious of our selves but only to be, at any point in our experience, capable of becoming self-conscious; we can turn our attention when we want to from whatever we are doing and watch ourselves doing what we are doing. This is an important point: Our concern with self-consciousness is given impetus just because we are often not self-conscious. In fact, several philosophers (and many mystics) have argued that self-consciousness is bad, a useless thing, and should be avoided as much as possible. According to Descartes, this is not possible, for to exist at all as a human being is to exist self-consciously. According to Kant, on the other hand, to exist as a human being is “to be able” to be self-conscious.

Second, Kant objects to Descartes’ belief that the thinking self is a thinking thing. He objects to this, first of all, because of his insistence (as a result of Hume’s argument) that the self (or “transcendental ego”) is not in our experience but rather “behind” it and responsible for it. More literally, he says that the self must be thought of as an activity. You can see what a radical move this is when you recall the traditional doctrine of the soul in Plato, in Christianity, and in much of modern thought. The soul, quite simply, is the self conceived of as a thing, an enduring thing that can

survive the death of the body. By saying that the self is an activity, Kant undermines (as Hume had intended to undermine) the traditional concept of soul.<sup>6</sup>

### AGAINST THE SOUL, BY IMMANUEL KANT

Pure reason requires us to seek for every predicate of a thing its own subject, and for this subject, which is itself necessarily nothing but a predicate, its subject, and so on indefinitely (or as far as we can reach). But hence it follows that we must not hold anything at which we can arrive to be an ultimate subject.

Now we appear to have this substance in the consciousness of ourselves (in the thinking subject), and indeed in an immediate intuition; for all the predicates of an internal sense refer to the *ego*, as a subject, and I cannot conceive myself as the predicate of any other subject. Hence completeness in the reference of the given concepts as predicates to a subject—not merely an Idea, but an object—that is, the absolute subject itself, seems to be given in experience. But this expectation is disappointed. For the *ego* is not a concept, but only the indication of the object of the inner sense, so far as we know it by no further predicate. Consequently it cannot indeed be itself a predicate of any other thing; but just as little can it be a definite concept of an absolute subject, but is, as in all other cases, only the reference of the inner phenomena to their unknown subject. Yet this idea (which serves very well as a regulative principle totally to destroy all materialistic explanations of the internal phenomena of the soul) occasions by a very natural misunderstanding a very specious argument, which infers its nature from this supposed knowledge of the substance of our thinking being. This is specious so far as the knowledge of it falls quite without the complex of experience.

But though we may call this thinking self (the soul) “substance,” as being the ultimate subject of thinking which cannot be further represented as the predicate of another thing, it remains quite empty and without significance if permanence—the quality which renders the concept of substances in experience fruitful—cannot be proved of it.

But permanence can never be proved of the concept of a substance as a thing in itself, but for the purposes of experience only.

If, therefore, from the concept of the soul as a substance we would infer its permanence, this can hold good as regards possible experience only, not of the soul as a thing in itself and beyond all possible experience. Life is the subjective condition of all our possible experience; consequently we can only infer the

---

<sup>6</sup>It is worth mentioning that Kant held onto the Christian concept of soul; to do so, he defended it as a “postulate of practical reason,” in other words, as a strictly moral claim, much as he had defended his belief in God. Later philosophers borrowed Kant’s arguments to get rid of the concept of the soul altogether.

permanence of the soul in life, for the death of a man is the end of all experience which concerns the soul as an object of experience, except the contrary be proved—which is the very question in hand. The permanence of the soul can therefore only be proved (and no one cares to do that) during the life of man, but not, as we desire to do, after death. The reason for this is that the concept of substance, so far as it is to be considered necessarily with the concept of permanence, can be so combined only according to the principles of possible experience, and therefore for the purposes of experience only.<sup>7</sup>

Third, Kant argued that we need two very different conceptions of self. He saw that this conception of self as self-consciousness was not sufficient to do the whole job that philosophers had wanted it to do. One part of Descartes' enterprise was to find out what was essential to his existence, what could not be doubted and so could serve as a first premise for his *Meditations*. So too Locke and Hume had tried to find (though Locke did and Hume didn't) that self that defined us through our various changes, which identified Jekyll and Hyde and identifies us from year to year, day to day, and mood to mood. But the function of the self was also to serve as a way of identifying ourselves in distinction from other people and other things. Thus Descartes' concept of the self as a thinking thing was not sufficient to tell us what made one person different from another, and he found it necessary to supplement his concept of self with an account of how a person was composed of a self and a body in some special way.

Similarly, Locke distinguishes between personal identity and identity as a man (that is, as a biological example of the species *Homo sapiens*) and tells us that both are necessary for us to understand how one particular person is different from another particular person. We can now clearly see that the question of self-identity divides into two questions: (1) What is essential to being a self? and (2) What is essential to being a particular self? Kant's conception of self as that which has experiences, the transcendental self, only answers the first question. Nothing in the notion of transcendental self allows us to distinguish between different people and tell them apart. Accordingly, he identifies another "self" that he calls the **empirical ego**, which includes all of those particular things about us that make us different people. Differences in our bodies, our looks, our size, our strength would be such differences. So too would our different personalities, our different thoughts and memories. It is the empirical self that identifies us as individual persons. The transcendental self makes us human.

---

<sup>7</sup>Immanuel Kant, *Prolegomena to Any Future Metaphysics*, trans. Lewis White Beck (New York: Bobbs-Merrill, 1950). Ludwig Wittgenstein, for example, used a Kantian thesis to deny the existence of the subject altogether. In his *Tractatus Logico-Philosophicus*, he tells us: The thinking, presenting subject—there is no such thing. *In an important sense* there is no subject. The subject does not belong to the world, but is a limit of the world. There is [therefore] really a sense in which in philosophy we can talk nonpsychologically of the I. The I occurs in philosophy through the fact that the "world is my world." The philosophical I is not the man, not the human body, or the human soul of which psychology treats, but the metaphysical subject, the limit—not a part of the world.



These have been the traditional answers to the philosophical problem of self-identity. What is it that makes one “the same person” from moment to moment and year to year? The spatiotemporal **continuity** of the body would seem to be a part of the answer. But philosophers since (and before) Descartes have seen quite clearly that this is never enough, that it is also self-consciousness that provides the key to self-identity. But even this has not solved the problem. To see this, consider these bizarre but illuminating examples:

Jones has an emergency brain operation. His own brain is removed and replaced by the brain of Brown (who is recently deceased). “Jones” still looks like Jones, still carries the same driver’s license and lives in the same house, but all of his memories and personality traits are those of Brown. Is he still Jones? Or is he Brown in Jones’s body? Suppose you had been (or are) Brown, could you claim to be still alive in Jones’s body?

Or, suppose Smith undergoes a personality split of the most radical kind imaginable. Like the one-celled animal, the amoeba, Smith splits, head to toe, each half of him forming an exact duplicate of the original—same memories and personality, same habits, knowledge, likes and dislikes, skills, and so on. Which of the two is Smith? Does it make any sense to say that both are? Suppose that you are one of the two resultant Smiths. Would you—could you—intelligibly say that the other Smith is also you? But, given your common origins and exact similarities, could you say that the other Smith was merely someone else?

In the movie *All of Me*, the soul of a woman who has recently died (played by Lily Tomlin) winds up occupying the body of a male lawyer (played by Steve Martin). The resulting character, played by Martin, still looks, acts, and thinks like the Steve Martin character, but now also possesses the memories and personality traits of Lily Tomlin’s original character. Who is this dual character?

To consider the bizarre complications, here is Massachusetts philosopher Meredith Michaels, who has written extensively on the problem of personal identity:

### **“PERSONAL IDENTITY,”**

BY MEREDITH MICHAELS

While they are illuminating, particularly in relation to one another, these traditional answers to the philosophical problem of self-identity raise as many questions as they answer. To see this, let us travel to a not very distant make-believe world.

One night, after a serious bout with the library, you and your best friend Wanda Bagg (or Walter, if you prefer) decide to indulge yourselves at the College Haven. Before you can stop her, Wanda steps out in front of a steamroller that happens to be moving down Main Street. Wanda is crushed. Witnessing the horror of the accident, you have a stroke. Fortunately, Dr. Hagendaas, the famous neurosurgeon who has been visiting the campus, is also on the way to the College Haven. Taking charge, he rushes you and Wanda to the Health Center, where he performs a “body transplant.” He takes Wanda’s brain, which

miraculously escaped the impact of the steamroller, and puts it in the place of yours, which was, of course, severely damaged by the stroke. After several days, the following battle ensues: Wanda's parents claim that they are under no obligation to continue paying tuition. After all, Wanda was killed by a steamroller. Your parents claim that they are under no obligation to continue paying tuition. After all, you died of a stroke. It is clear, then, that a basic question is in need of an answer: who is the person lying in bed in the Health Center? Is it Wanda? Is it you? Is it someone else altogether? For the sake of discussion, let us call the person lying in the bed Schwanda. What reasons do we have for believing that Schwanda is Wanda? Given that one's self-consciousness, one's thoughts, beliefs and feelings are all mental phenomena, we might naturally conclude that a person goes wherever her brain goes (on the assumption that our mental characteristics are more likely "located" in the brain than in, say, our smallest left toe). Schwanda will remember having set off for the College Haven with you; she will remember receiving the college acceptance letter addressed: "Dear Wanda, We are happy to inform you that . . ."; she'll remember being hugged by Wanda's mother on the afternoon of her first day of school. That is, Schwanda will *believe* that she's Wanda.

Nevertheless, the fact that Schwanda believes herself to be Wanda does not in itself guarantee that she is. Do we have any basis for insisting that Schwanda is Wanda and not someone who is *deluded* into thinking that she's Wanda? How can we determine whether Schwanda's Wanda memories are genuine and not merely apparent? As we came to realize in our discussion of Locke's Memory Theory, it is not legitimate at this point to appeal to the self-identity of Schwanda and Wanda, since that is precisely what we're trying to determine. In other words, in attempting to establish that Schwanda's Wanda memories are genuine memories, we cannot argue that they are genuine on the grounds that Schwanda *is* Wanda.

Perhaps it is possible to stop short of circularity. Why couldn't we say that Schwanda's Wanda memories are genuine because the *brain* that is remembering is the same as the brain that had the original experiences. Thus, the experiences are preserved in the very organ that underwent them. Though there is an initial plausibility to this response, it fails to solve our problem. Suppose that Schwanda is Wanda—remembering the experience of learning to ride a bicycle. Though the brain in question is indeed the same, it is nonetheless clear to all of us that brains alone do not learn to ride bicycles. Nor, indeed, do brains alone remember having done so. *People* learn to ride bicycles and *people* remember having done so. And the question we are trying to answer is whether Schwanda (who is remembering) is the same person as Wanda (who did the bicycling). The appeal to the fact that the same brain is involved in each event does not provide us with a way out of the Lockean circle.

It is at this point that philosophers begin to reconsider the Aristotelian position, mentioned earlier, that self-identity is essentially *bodily* identity. If the Body Theory of Personal Identity is true, then the person lying in bed at

the Health Center is you, deluded into believing that you are Wanda. That is, Schwanda is self-identical to you.

You might wonder, at this point, whether there are any positive reasons for endorsing the Body Theory, or whether it is simply a place to which one retreats only in defeat? The following case is designed to persuade you that there is at least *some* plausibility to the Body Theory. Suppose that an evil scientist, Dr. Nefarious, has selected you as his prime subject for a horrible experiment. You are dragged into his office. He says, "Tomorrow at 5:00, you will be subjected to the most terrible tortures. Your nails will be pulled out one by one. Rats will be caged around your head. Burning oil will drip slowly on your back. The remainder I leave as a surprise."

Are you worried about what will happen to you at 5:00 tomorrow? If you have any sense, you are. You think of the excruciating pain and suffering you will undergo and would surely do just about anything to avoid it.

But now, Dr. Nefarious says, "Tomorrow at 4:55, I will use my Dememorizer to erase your memory of this conversation." Are you still anxious about what is going to happen to you tomorrow at 5:00? Surely you are. After all the fact that you won't, between 4:55 and 5:00, be anticipating your torture doesn't entail that the torture itself will be any less painful. When you forget that your Calculus professor told the class there would be a test on Friday, you aren't thereby spared the experience of taking the test (in fact, in that case the experience is made worse by your not having had the opportunity to anticipate it).

Now, Dr. Nefarious says, "Tomorrow at 4:57, I will use my Dememorizer to erase *all* of your memories." Are you still anxious about what will happen tomorrow at 5:00? Isn't it natural to describe the situation as one in which you will undergo horrible torture, though you won't know who you are or why this is happening to you? *You* will still experience *your* fingernails being pulled out, *your* back being burned, *your* face being eaten up by rats. Surely, those experiences are ones you would like to avoid.

Finally, Dr. Nefarious says to you, "Tomorrow at 4:58, I am going to use my Rememorizer to implant in your brain all of Ronald Reagan's memories." Though this may not please you for personal or political reasons, the relevant question remains this: are you still worried about what is going to happen tomorrow at 5:00? Isn't it again perfectly natural to describe the situation as one in which you will undergo horrible torture, all the while believing that you are Ronald Reagan. Do you not *now* remain concerned that you are Ronald Reagan. Do you not *now* remain concerned that you will experience excruciating pain and intolerable suffering? Look at your fingernails while you consider your answer to this question.

What this story demonstrates is not the conclusive superiority of the Body Theory over the Memory (or Brain) Theory, but rather the importance of our bodies to our self-identity. This is something that tends to get lost in the traditional conceptions of personal identity. Furthermore, returning to the case of Schwanda, we can now see that it is not altogether preposterous to argue that

Schwanda is indeed you, deluded into believing that she is Wanda. In other words, anyone who wishes to dismiss the possibility must also dismiss the possibility that the person who undergoes the torture is indeed you, deluded into believing that you are Ronald Reagan.

While it is true that we tend to identify ourselves with and by our thoughts, beliefs, inclinations and feelings, our discussion of the Body Theory should remind us that there are reasons for believing that our bodies are, at the very least, important to who we are. Some philosophers would argue that our bodies *are* who we are, that self-identity *is* bodily identity.

In considering these admittedly fanciful problem cases, we have seen that we lack a concept of self-identity that allows us to predict when we would or wouldn't persist through time. This might suggest to us that our concept of self-identity is not an all-or-nothing one, that, in fact, our concept is one which admits of degrees. If so, we are no longer talking about identity *per se*, which is an all-or-nothing concept, but rather about some other relation of psychological and physical connectedness. Nevertheless, we can now see first, that the answer to the question "Who ought to pay Schwanda's tuition?" will depend upon which theory of personal identity we are inclined to endorse and second, that the answer may not be as clear and unequivocal as we would like it to be.<sup>8</sup>

## B. EXISTENTIALISM: SELF-IDENTITY AND FREE CHOICE

The idea of a multiple, or nonself introduced a very important alternative to Western, essentially Judeo-Christian conceptions. In addition to the concern for survival in Heaven or Hell, this Western conception is typically concerned with striving and ambition, status and planning for the future, "making something of yourself." Existentialism is one form of this conception. The American dream and the Protestant and capitalist ethics in general are another. But from the conception of nonself comes a very different picture—of unqualified acceptance of things as they are rather than struggling to change them, which involves a rejection of such notions as "status" and "making something of yourself."

Once again we must not confuse this rejection of traditional Western conceptions of the self with a rejection of the philosophical question of self-identity. "Who am I?" is as important a question for the mystic as it is for the Western philosopher. It is just that the mystic's answer to the question is radically different. Does it make sense to say that the one is more "correct" than the other? This too is part of the question of self-identity. Must there be a single "correct" answer for everyone? There is nothing necessary about this, or even desirable. The problem of self-identity, both in each

<sup>8</sup>This essay was written by Meredith Michaels for the third edition of *Introducing Philosophy*.

individual case and as a general problem, is the problem of deciding which of the many possible characteristics (not necessarily one) should be chosen as our own standards for self-identity, and that choice is not just one that philosophers make, but one which each person makes at some point or points during his or her life.

A self-identity isn't simply a label you throw on yourself in the casual discussion of a philosophy class. It is a mask and a role that you wear in every social encounter (though perhaps slightly different masks and roles for significantly different encounters). It is the way you think of yourself and the standards by which you judge yourself in every moment of reflection and self-evaluation. It is the self-image you follow in every action, when you decide that one thing is "worth doing" more than another or when you decide how to act in a given circumstance. Because of it, you feel proud, guilty, ashamed, or delighted after you have done something. The problem of self-identity is not just a problem for philosophers; it is a problem we all face, either explicitly or implicitly, every self-conscious minute of our lives.

But, you might say, why make it sound as if there is any single notion or goal of a correct self-identity? The way in which a contemporary Chinese farmer thinks of himself and judges himself is very different from that of a contemporary American college student. And a very handsome but stupendously dumb bully will surely have a very different conception of self-identity than an extremely intelligent and talented college mathematics major. This may not stop the medieval scholastic, who will immediately declare all of that irrelevant and insist that "before God" all of us are the same and our identities are to be judged accordingly. And most of us, despite the glib relativism we usually defend, would insist on a category that transcends all such individual considerations; we call it "being a good person." Ultimately we would judge the dumb bully and the budding young artist according to the same criterion, and, in doing so, we would think that they should share the criterion "being a good person." Even where cultural differences would seem to demand entirely different conceptions of self-identity, we might still insist on applying the same criterion. For example, a South Sea Islander might well think of himself or herself in terms that would be wholly unacceptable to us, but we can always reduce any variance from our norms to mere "accidental differences," insisting that we are essentially the same. Of course people are different and think differently of themselves; but it does not follow that those differences are essential, nor does it follow that relativism is true. Ultimately when you say that all people are "essentially the same," then you believe that there are, indeed, universal criteria for self-identity and that the differences between people, though we need not deny them, are merely superficial.

One of the most powerful schools of contemporary thought, however, has been dedicated to the idea that self-identity, in every case, is a matter of individual choice. This school, which we have briefly met before, is **existentialism**. Its most powerful advocate is the French philosopher Jean-Paul Sartre. According to Sartre, there are no set standards for self-identity, either for individuals or for people in general. There is, he argues, no such thing as "human nature," and what we are—and what it means to be a human being—are always matters of decision. There is no correct choice; there are only choices, he claims. In a well-known essay from the late 1940s, he argues:

## ON EXISTENTIALISM, BY JEAN-PAUL SARTRE

What existentialists have in common is simply the fact that they believe that *existence* comes before *essence*—or, if you will, that we must begin from the subjective. What exactly do we mean by that?

If one considers an article of manufacture—as, for example, a book or paper-knife—one sees that it has been made by an artisan who had a conception of it; and he has paid attention, equally, to the conception of a paper-knife and to the pre-existent technique of production which is a part of that conception and is, at bottom, a formula. Thus the paper-knife is at the same time an article producible in a certain manner and one which, on the other hand, serves a definite purpose, for one cannot suppose that a man would produce a paper-knife without knowing what it was for. Let us say, then, of the paper-knife that its essence—that is to say the sum of the formulae and the qualities which made its production and its definition possible—precedes its existence. The presence of such-and-such a paper-knife or book is thus determined before my eyes. Here, then, we are viewing the world from a technical standpoint, and we say that production precedes existence.

When we think of God as the creator, we are thinking of him, most of the time, as a supernal artisan. Whatever doctrine we may be considering, whether it be a doctrine like that of Descartes, or of Leibniz himself, we always imply that the will follows, more or less, from the understanding or at least accompanies it, so that when God creates he knows precisely what he is creating. Thus, the conception of man in the mind of God is comparable to that of the paper-knife in the mind of the artisan: God makes man according to a procedure and a conception, exactly as the artisan manufactures a paper-knife, following a definition and a formula. Thus each individual man is the realisation of a certain conception which dwells in the divine understanding. In the philosophic atheism of the eighteenth century, the notion of God is suppressed, but not, for all that, the idea that essence is prior to existence; something of that idea we still find everywhere, in Diderot, in Voltaire and even in Kant. Man possesses a human nature; that “human nature,” which is the conception of human being, is found in every man; which means that each man is a particular example of an universal conception, the conception of Man. In Kant, this universality goes so far that the wild man of the woods, man in the state of nature and the bourgeois are all contained in the same definition and have the same fundamental qualities. Here again, the essence of man precedes that historic existence which we confront in experience.

♦ ♦ ♦

What do we mean by saying that existence precedes essence? We mean that man first of all exists, encounters himself, surges up in the world—and defines himself afterwards. If man as the existentialist sees him is not definable, it is because to begin with he is nothing. He will not be anything until later, and then he will be what he makes of himself. Thus, there is no human nature, because there

is no God to have a conception of it. Man simply is. Not that he is simply what he conceives himself to be, but he is what he wills, and as he conceives himself after already existing—as he wills to be after that leap towards existence. Man is nothing else but that which he makes of himself. That is the first principle of existentialism. And this is what people call its “subjectivity,” using the word as a reproach against us. But what do we mean to say by this, but that man is of a greater dignity than a stone or a table? For we mean to say that man primarily exists—that man is, before all else, something which propels itself towards a future and is aware that it is doing so. Man is, indeed, a project which possesses a subjective life, instead of being a kind of moss, or a fungus or a cauliflower. Before that projection of the self nothing exists; not even in the heaven of intelligence: man will only attain existence when he is what he purposes to be. Not, however, what he may wish to be. For what we usually understand by wishing or willing is a conscious decision taken—much more often than not—after we have made ourselves what we are. I may wish to join a party, to write a book or to marry—but in such a case what is usually called my will is probably a manifestation of a prior and more spontaneous decision. If, however, it is true that existence is prior to essence, man is responsible for what he is. Thus, the first effect of existentialism is that it puts every man in possession of himself as he is, and places the entire responsibility for his existence squarely upon his own shoulders. And, when we say that man is responsible for himself, we do not mean that he is responsible only for his own individuality, but that he is response for all men.

When we say that man chooses himself, we do mean that every one of us must choose himself; but by that we also mean that in choosing for himself he chooses for all men. For in effect, of all the actions a man may take in order to create himself as he wills to be, there is not one which is not creative, at the same time, of an image of man such as he believes he ought to be.<sup>9</sup>

But this existentialist doctrine of choice doesn't make the problem of self-identity any easier. In fact, it complicates it enormously. Earlier in this section, we began by asking whether the facts about a person are sufficient to determine his or her identity. We said surely not all of them are necessary; some are more essential than others. But this isn't yet an answer to the question, for it may be that all the essential facts are still not sufficient to determine a person's identity.

According to the existentialist, this is made even more complex by the fact that a person chooses which facts are to be considered as essential. Are the facts alone ever sufficient to determine our identity? Sartre's answer, which he adapted from German existentialist Martin Heidegger, is “never!” The facts that are true of a person are always, at least so long as a person is alive, only indicative of what a person has been and done so far. In judging a person's identity, we must always consider more than

---

<sup>9</sup> Jean-Paul Sartre, *Existentialism As a Humanism*, trans. Phillip Mairet (New York: Philosophical Library, 1949).

the facts that are true of him or her (which Sartre and Heidegger collectively name, somewhat technically, a person's **facticity**); we must also consider their projections into the future, their ambitions, plans, intentions, hopes, and fantasies. (Sartre calls these considerations a person's **transcendence**. Notice that this is the third different way in which "transcendence" has been used, so be careful.) This way of viewing the person makes the question of self-identity impossibly complex, in fact, irresolvable. For example, consider Sartre's example of what he calls "bad faith" in one of his most important works. *Being and Nothingness* (1943).

**Bad faith**, quite simply, is refusing to accept yourself.<sup>10</sup> This can happen in two different ways. Either you can refuse to accept the facts and actions as relevant to your self-identity (for example, denying that your repeated cowardly behavior establishes your identity as a coward). Or you can go too far in the opposite direction, believing that your actions conclusively and unalterably establish your self-identity (for example, denying that you could ever alter your cowardly self-identity through an act of heroism).

**ON BAD FAITH,  
BY SARTRE**

Let us take an example: A homosexual frequently has an intolerable feeling of guilt, and his whole existence is determined in relation to this feeling. One will readily foresee that he is in bad faith. In fact it frequently happens that this man, while recognizing his homosexual inclination, while avowing each and every particular misdeed which he has committed, refuses with all his strength to consider himself "*a homosexual*." His case is always "different," peculiar; there enters into it something of a game, of chance, of bad luck; the mistakes are all in the past; they are explained by a certain conception of the beautiful which women cannot satisfy; we should see in them the results of a restless search, rather than the manifestations of a deeply rooted tendency, *etc., etc.* Here is assuredly a man in bad faith who borders on the comic since, acknowledging all the facts which are imputed to him, he refuses to draw from them the conclusion which they impose. His friend, who is his most severe critic, becomes irritated with this duplicity. The critic asks only one thing—and perhaps then he will show himself indulgent: that the guilty one recognize himself as guilty, that the homosexual declare frankly—whether humbly or boastfully matters little—"I am a homosexual." We ask here: Who is in bad faith? The homosexual or the champion of sincerity?

The homosexual recognizes his faults, but he struggles with all his strength against the crushing view that his mistakes constitute for him a *destiny*. He does

<sup>10</sup>This is, however, the main concept of *Being and Nothingness* and takes well over seven hundred pages to analyze correctly.



not wish to let himself be considered as a thing. He has an obscure but strong feeling that a homosexual is not a homosexual as this table is a table or as this red-haired man is red-haired. It seems to him that he has escaped from each mistake as soon as he has posited it and recognized it; he even feels that the psychic duration by itself cleanses him from each misdeed, constitutes for him an undetermined future, causes him to be born anew. Is he wrong? Does he not recognize in himself the peculiar, irreducible character of human reality? His attitude includes then an undeniable comprehension of truth. But at the same time he needs this perpetual rebirth, this constant escape in order to live; he must constantly put himself beyond reach in order to avoid the terrible judgment of collectivity. Thus he plays on the word *being*. He would be right actually if he understood the phrase, "I am not a homosexual" in the sense of "I am not what I am." That is, if he declared to himself, "To the extent that a pattern of conduct is defined as the conduct of a paederast and to the extent that I have adopted this conduct, I am a homosexual. But to the extent that human reality cannot be finally defined by patterns of conduct, I am not one." But instead he slides surreptitiously toward a different connotation of the word "being." He understands "not being" in the sense of "not-being-in-itself." He lays claim to "not being a homosexual" in the sense in which this table *is not* an inkwell. He is in bad faith.<sup>11</sup>

Sartre then lays bare the heart of his theory. Bad faith points to the most important single fact about personal self-identity—there isn't any. In somewhat paradoxical terminology, Sartre tells us, "one is what one is not, and one is not what one is." In other words, whatever the facts about you, you are always something more than those facts. The homosexual in Sartre's example *is* a homosexual to the extent that all his past actions and desires are those of a homosexual. He falls into bad faith by refusing to see that his past actions point to his having a self-identity as a homosexual. Yet at the same time, there is a genuine sense in which he is *not* a homosexual: In the future, he may radically alter his lifestyle. It would, then, also be bad faith were he to totally accept his self-identity as a homosexual, denying that he could be anything else. As long as a person is alive, he or she is identified by intentions, plans, dreams, and hopes as much as by what is already true by virtue of the facts. Given this complexity, the problem of deciding "who I am" takes on dramatic and extravagant complications. Consider the following scene from Sartre's famous play, *No Exit*, in which one of the characters (now dead and "living" in hell) tries to justify his image of himself as a hero, despite the facts of his life, which would indicate that he was a coward.

---

<sup>11</sup> Jean-Paul Sartre, *Being and Nothingness*, trans. Hazel E. Barnes (New York: Philosophical Library, 1956).

FROM **NO EXIT**,  
BY **SARTRE**

GARCIN: They shot me.

ESTELLE: I know. Because you refused to fight. Well, why shouldn't you?

GARCIN: I—I didn't exactly refuse. [*In a far-away voice*] I must say he talks well, he makes out a good case against me, but he never says what I should have done instead. Should I have gone to the general and said: "General I decline to fight"? A mug's game; they'd have promptly locked me up. But I wanted to show my colors, my true colors, do you understand? I wasn't going to be silenced. [*To ESTELLE*] So I—I took the train. . . . They caught me at the frontier.

ESTELLE: Where were you trying to go?

GARCIN: To Mexico. I meant to launch a pacifist newspaper down there. [*A short silence.*] Well, why don't you speak?

ESTELLE: What could I say? You acted quite rightly, as you didn't want to fight. [*GARCIN makes a fretful gesture.*] But, darling, how on earth can I guess what you want me to answer?

INEZ: Can't you guess? Well, *I* can. He wants you to tell him that he bolted like a lion. For "bolt" he did, and that's what's biting him.

GARCIN: "Bolted," "went away"—we won't quarrel over words.

ESTELLE: But you *had* to run away. If you'd stayed they'd have sent you to jail, wouldn't they?

GARCIN: Of course. [*A pause.*] Well, Estelle, am I a coward?

ESTELLE: How can I say? Don't be so unreasonable, darling. I can't put myself in your skin. You must decide that for yourself.

GARCIN: [*warily*]: I can't decide.

ESTELLE: Anyhow, you must remember. You must have had reasons for acting as you did.

GARCIN: I had.

ESTELLE: Well?

GARCIN: But were they the real reasons?

ESTELLE: You've a twisted mind, that's your trouble. Plaguing yourself over such trifles!

GARCIN: I'd thought it all out, and I wanted to make a stand. But was that my real motive?

INEZ: Exactly. That's the question. Was that your real motive? No doubt you argued it out with yourself, you weighed the pros and cons, you found good reasons for what you did. But fear and hatred and all the dirty little instincts one keeps dark—they're motives too. So carry on, Mr. Garcin, and try to be honest with yourself—for once.

GARCIN: Do I need you to tell me that? Day and night I paced my cell, from the window to the door, from the door to the window. I pried into my heart, I sleuthed myself like a detective. By the end of it I felt as if I'd given my

whole life to introspection. But always I harked back to the one thing certain—that I had acted as I did, I'd taken that train to the frontier. But why? Why? Finally I thought: My death will settle it. If I face death courageously, I'll prove I am no coward.

INEZ: And how did you face death?

GARCIN: Miserably. Rottenly. [INEZ *laughs*.] Oh, it was only a physical lapse—that might happen to anyone; I'm not ashamed of it. Only everything's been left in suspense, forever. [To ESTELLE] Come here, Estelle. Look at me. I want to feel someone looking at me while they're talking about me on earth. . . . I like green eyes.

INEZ: Green eyes! Just hark to him! And you, Estelle, do you like cowards?

ESTELLE: If you knew how little I care! Coward or hero, it's all one—provided he kisses well.

GARCIN: There they are, slumped in their chairs, sucking at their cigars. Bored they look. Half-asleep. They're thinking: "Garcin's a coward." But only vaguely, dreamily. One's got to think of something. "That chap Garcin was a coward." That's what they've decided, those dear friends of mine. In six months' time they'll be saying: "Cowardly as that skunk Garcin." You're lucky, you two; no one on earth is giving you another thought. But I—I'm long in dying.

GARCIN: [*putting his hands on (INEZ's) shoulders*]: Listen! Each man has an aim in life, a leading motive; that's so, isn't it? Well, I didn't give a damn for wealth, or for love. I aimed at being a real man. A tough, as they say. I staked everything on the same horse. . . . Can one possibly be a coward when one's deliberately courted danger at every turn? And can one judge a life by a single action?

INEZ: Why not? For thirty years you dreamt you were a hero, and condoned a thousand petty lapses—because a hero, of course, can do no wrong. An easy method obviously. Then a day came when you were up against it, the red light of real danger—and you took the train to Mexico.

GARCIN: I "dreamt," you say. It was no dream. When I chose that hardest path, I made my choice deliberately. A man is what he wills himself to be.

INEZ: Prove it. Prove it was no dream. It's what one does, and nothing else, that shows the stuff one's made of.

GARCIN: I died too soon. I wasn't allowed time to—to do my deeds.

INEZ: One always dies too soon—or too late. And yet one's whole life is complete at that moment, with a line drawn neatly under it, ready for the summing up. You are—your life, and nothing else.<sup>12</sup>

## GLOSSARY

**bad faith** Sartre's characterization of a person's refusal to accept himself or herself; this sometimes means not accepting the facts that are true about you. More often it means accepting the facts about you as conclusive about your identity, as in the statement "Oh, I couldn't do that, I'm too shy."

**continuity** (spatiotemporal continuity) The uninterrupted identifiability of an object over time in the same location or in a sequence of tangent locations.

**criterion** Test or standard.

**deconstruction** Initiated by Jacques Derrida, a current school of philosophical thought (especially popular among some feminist and African-American thinkers) that encourages critical reading for "cultural bias" and that rejects the idea of the "unified self."

**dualism** In general, the distinction between mind and body as separate substances, or very different kinds of states and events with radically different properties.

**empirical ego** All those characteristics of a person that can be discovered through experience and that distinguish each of us from other persons qualitatively; that which makes each of us a particular man or woman and gives us a particular "character." Compare *transcendental ego*.

**epiphenomenalism** The thesis that mental events are epiphenomena, that is, side effects of various physical processes in the brain and nervous system but of little importance themselves. The model is a one-way causal model: Body states cause changes in the mind, but mental states have no effect in themselves on the body.

**essence** That which is necessary for something to be what it is. The essence of a person is that without which we would not say one is *that* particular person (Fred rather than Mary, for example).

**existentialism** In Sartre's terms, the philosophy that teaches that "man's existence precedes his essence." That is, people have no given self-identity, they have to choose their identities and work for them through their actions. (Neglect and omission, however, are also actions. One can be a certain type of person just by not bothering to do the appropriate activities.)

**facticity** Sartre's term (borrowed from Heidegger) for the totality of facts that are true of a person at any given time.

**immediate** For certain and without need for argument.

**incorrigibility** Impossible to correct; cannot be mistaken. It has long been argued that our claims about our own mental states are incorrigible—we cannot be mistaken about them.

**inference-ticket** Ryle's term for referring to the proper function of a mental state: talk, as a description of a pattern of behavior and, therefore, as an "inference-ticket" that allows us to infer what a person will do in the future. (To say "George wants an olive" is to give us an inference-ticket regarding his future behavior around olives.)

**intentionality** In phenomenology, the thesis that every conscious act has an object. (The act is therefore called the "intentional act" and the object the "intentional object.") The importance of this concept is that it undercuts the metaphor of mental "contents" (as in a theater, an image explicitly used by Hume, for example). The concept was used by Husserl's teacher, Franz Brentano, who borrowed it from some medieval philosophers, before Husserl used it and made it famous.

**privacy** The seeming inaccessibility of mental states and events to anyone other than the person who "has" them.

**private language argument** Wittgenstein's argument that even if there were such "private

objects" as mental states and events, it would be impossible for us to talk about them and impossible for us to identify them, even in our own case.

**privileged access** The technical term used by philosophers to refer to the curious fact that a person usually (if not always) can immediately know, simply by paying attention, what is going on in his own mind, while other people can find out what is going on—if they can at all—only by watching the person's behavior, listening to what he or she says, or asking (and hoping they get a truthful answer). It is important to distinguish privileged access from incorrigibility. The first means that a person knows directly what is "in his mind" without having to observe this behavior; the second means that he knows for certain and beyond the possibility of error.

**resemblance** Having the same features. All people resemble each other (or at least most do) in having one and only one head; you resemble yourself five years ago in (perhaps) having the same texture hair, the same color eyes, the same fear of spiders, and the same skill at chess.

**self-consciousness** Being aware of oneself, whether "as others see you" (looking in a mirror or "watching yourself play a role" at a party) or just "looking into yourself" (as when you reflect on your goals in life or wonder, in a moment of philosophical perversity, whether you really exist or not.) Self-consciousness requires having some concept of your "self." Accordingly, it is logically tied to questions of self-identity.

**self-identity** The way you characterize yourself, either in general (as a human being, as a man or as a woman, as a creature before God or as one among many animals) or in particular (as the person who can run the fastest mile, as an all-"C" student, or as the worst-dressed slob in your class). Self-identity, on this characterization, requires self-consciousness. The self-identity of a person, in other words, is not merely the same as the identity of a "thing," for example, the identity of a human body.

**transcendence** Sartre's term for a person's plan, ambitions, intentions, and hopes for the future (Do not confuse this use of the word with those introduced in Chapters 3 and 5.)

**transcendental ego** The bare, logical fact of one's own self-consciousness: Descartes "I think"; the self "behind" all of our experiences; the mental activity that unifies our various thoughts and sensations. (The term comes from Kant's *Critique of Pure Reason*.)

**unconscious** Freud's famous way of referring to the fact that there are ideas, desires, memories, and experiences in our minds to which we do not have privileged access, which we may be wrong about (and, therefore, about which our claims are not incorrigible), and which may be more evident to other people than to oneself. He also distinguishes a *preconscious* ("the antechamber of consciousness"). Preconscious ideas can be made conscious simply by being attended to. (For example, you do know what the capital of California is, but you weren't conscious of it before I mentioned it; it was preconscious.) Truly unconscious ideas, however, cannot be made conscious, even when one tries to do so.

## BIBLIOGRAPHY AND FURTHER READING

A recent anthology on the question of self-identity is J. Perry, *Personal Identity* (New York: Lieber-Atherton, 1975). An extended study of this traditional problem is Sidney Shoemaker, *Self-Knowledge and Self-Identity* (Ithaca, NY: Cornell University Press, 1963). Jean-Paul Sartre's existentialist view of the self is developed in his essay *Transcendence of the Ego* (New York: Noonday, 1957). Hermann Hesse's complex theory of the self is best developed in