

# Why There Still Are No People

JIM STONE

*University of New Orleans*

This paper argues that there are no people. If identity isn't what matters in survival, psychological connectedness isn't what matters either. Further, fissioning cases do *not* support the claim that connectedness is what matters. I consider Peter Unger's view that what matters is a continuous physical realization of a core psychology. I conclude that if identity isn't what matters in survival, nothing matters. This conclusion is deployed to argue that there are no people. Objections to Eliminativism are considered, especially that morality cannot survive the loss of persons.

This paper will argue that there are no people.<sup>1</sup> Let me summarize the argument (part I is covered last). In part II of what follows, I argue that if identity isn't what matters in survival, psychological connectedness isn't what matters either. Psychological connectedness, according to Derek Parfit, is the 'holding of particular direct psychological connections,' for example, when a belief, a desire, or some other psychological feature continues to be had (Parfit 1986, 206); psychological connectedness consists in two other relations—resemblance plus a cause that produces it. For our purposes, to say of a relation that it is 'what matters in survival' is to say that it carries the burden of responsibility, remorse, and regret for past misdeeds; and that it is what makes rational the anticipation of, and the special hopes and fears we have about, our own future experiences. A consequence of II is that if identity isn't what matters in survival, either something other than psychological connectedness is what matters or nothing matters.

However fissioning cases appear to show that connectedness to a future person who isn't me is about as good as survival. Part III explains why fissioning cases do *not* support the claim that connectedness is what matters. Part IV considers Peter Unger's view that what matters in survival isn't connectedness but a physically continuous realization of my core psychological properties (Unger, 1990). IV concludes that if identity isn't what matters in survival, nothing matters.

Part V argues that there are no people if nothing matters in survival. It follows that either identity is what matters or there are no people. It will be

---

<sup>1</sup> This essay is meant to buttress my paper, 'Parfit and The Buddha: Why There are No People' (Stone, 1988). 'People' and 'persons' are used as synonyms.

helpful to set out here some of the ontological implications. On the Reductionist account, a person just consists in the existence of a brain and body, and the occurrence of a series of interrelated physical and mental states. The fact of a person's identity through time consists in the holding of certain more particular facts which can be fully described in a wholly impersonal way. As Parfit writes: "Since personal identity over time just consists in the holding of certain other relations, what matters is these other relations" (Parfit 1984, 478). The argument above concludes that either identity is what matters or there are no persons; but if identity is nothing more than these other relations, it is *not* what matters. It follows that identity is a deep further fact—if there are people. Persons must be ontologically extra to bodies, brains, and psychophysical events.

Then what are persons? The Reductionist account is the plausible naturalistic account of persons, I submit. I take the physical alternatives that have been considered (e.g. we are subatomic particles (Chisholm, 1978)) to be non-starters. If persons are extra to bodies, brains, and psychophysical events, they are immaterial souls. Naturalism denies that entities exist which lie beyond the scope of scientific explanation (see Danto, 1967). A consequence of Naturalism, therefore, is that there are no persons. As Eliminativism is a consequence of scientific naturalism (not to mention Empiricism), we have as forceful an argument for Eliminativism as can be expected in this area of philosophy; and the 'No souls, no people' conclusion is important whether or not Naturalism is accepted.

Indeed, this might be taken to provide a knock-down argument against Naturalism. We must choose between Eliminativism and Realism; as the former is zany, there are souls. Of course the objection loses its force if Eliminativism turns out *not* to be so zany, especially as Dualism is notoriously problematic. Consequently V and VI consider objections to Eliminativism. However it will prove useful to begin with another question: "If Reductionism is correct, and connectedness, not identity, is what matters in survival, which criterion of personal identity should we choose?"

## I

Let's identify the Psychological Criterion (PC) with the claim that personal identity consists of non-branching psychological continuity. Psychological continuity consists of overlapping relations of psychological connectedness.<sup>2</sup> Let's identify the Biological Criterion (BC) with the claim that human persons are identical to human animals, so that person X at t1 is identical to

---

<sup>2</sup> "Psychological continuity is the holding of overlapping chains of *strong* connectedness" (Parfit 1986, 206); chains are strong when they contain enough direct connections for personal identity. Parfit writes: "Of these two relations connectedness is more important both in theory and practice" (206).

person Y at  $t_2$  iff X and Y are one and the same human animal (or 'man,' to use Locke's term).<sup>3</sup> As human persons are self-aware men, and self-aware men existed before they were self-aware and can survive the loss of self-awareness, persons aren't essentially persons (Stone, 1987; Olson, 1997).

If psychological connectedness is what matters in survival, PC might seem superior; in effect it maintains that identity consists in what matters—at least when that doesn't branch. Indeed, as we're used to thinking that identity matters, a criterion of identity which apparently tracks what matters appears to track identity. However Parfit observes:

Now that we have seen that identity is not what matters, we need not try to revise or extend our criterion of identity, so that it coincides more often with what matters. On any natural understanding of personal identity, such a coincidence could be only partial, as the case of division shows. And revising our criterion may misleadingly suggest that identity is what matters (Parfit 1984, 272).

As both criteria coincide only partly with what matters, and as it's easy enough to express the facts once identity and connectedness are separated (so that the 'biologist' can say if I fall into a PVS, that I survive but lose what matters, while in Teletransportation I have what matters but I don't survive),<sup>4</sup> PC's superiority in tracking what matters isn't a decisive advantage. Indeed, once identity is separated from what matters, the fact that PC apparently tracks what matters no longer suggests that it tracks identity. The principal reason philosophers have accepted PC is undercut.

On the other hand, BC's advantages are striking. Plainly, it has the advantage of ontological simplicity (Stone, 1987; Olson, 1997). Persons are identified with self-aware men. Given PC there is no man (or brain) to which I'm identical. Indeed, I can jump from man to man, as in Teletransportation. What then is my relation to the human animal whose hands are typing these words? Are there two physical objects, the person and the animal, coincident in space and time? Which is thinking this thought (Olson, 1997)?

Worse, PC yields strongly counter-intuitive consequences. I might have become sentient at a slightly different time, been born a few minutes earlier; as a child I might have sat at a different place at the dinner table; in the first grade I might have been assigned a different seat. Indeed, it might have been the case that throughout my entire life all of my token mental-events were just different enough, either in content or in temporal location, to be numerically distinct from the tokens I actually had. Nonetheless my life could have been as close to my actual life as these differences allow: the same parents,

---

<sup>3</sup> Olson, 1997, originates the name 'Biological Criterion.'

<sup>4</sup> If 'in Teletransportation I have what matters but I don't survive,' to what do the two occurrences of 'I' refer? To me, the man about to step into the Teletransporter. I don't survive, because the fellow materialized on Mars isn't me. I have what matters because my relationship to him warrants anticipating his experiences, and so on.

the same name, the same job—I might write this paper, but each sentence would be written a few minutes earlier or later. This might have happened to *me*. BC seconds the conclusion; as this person (Mr. A) is the same man as me, he is me.

According to PC, my life could have been different, certainly. If I'd studied economics in graduate school, I would probably be living somewhere else, I would be married to someone else, and certainly I would be having different experiences. As the economist would be psychologically continuous with the person I was before I went to college (Mr. B), according to PC the economist would be Mr. B. Of course I'm also continuous with Mr. B; so, by the transitivity of identity, the economist would have been me. In short, my life could have branched and still been mine—so long as the person I would have been and I share a psychological past part. A strongly counter-intuitive consequence of PC, however, is that Mr. A would *not* have been me (Stone, 1987). For no stage of A's life is psychologically connected to any stage of mine. There are no causal relations whatsoever between events in his life and mine; for instance, neither of us remembers anything of the other's life.

The counter-example's force can be increased. Suppose the first years of my life proceed as they do in the counter-example. Token mental-events are just different enough to be distinct from the tokens I actually had, and this continues until there is plainly a persisting self-aware person, five-years-old (Mr. C). A consequence of PC is that C is distinct from me his whole life through, no matter how his life goes. For if C dies at five, as he and I have no common psychological part, C must be distinct from me. But if C is distinct from me if he dies at five, then he is distinct from me if he goes on living; for surviving cannot make C a numerically different person from the person he would be if he died at five.

The story now takes a new turn. When I go to first grade I'm assigned the seat I was assigned in my actual life, so that my visual and auditory perceptions are the ones I actually had. At home I switch to the place at the table in which I actually sat, and so on. As my memory is poor, I gradually forget the experiences I had before. As my perceptions are caused by the same physical objects, affecting the same neurons in the same way at the same time, they are the very token perceptions I've actually had; the same thing is true of my token memories of these perceptions. By the time I'm six, my life consists of just the token mental-events that actually comprised it; and so it goes right down to this moment, as I write these words in just this way, having these token thoughts and visual perceptions. A catastrophic consequence of PC is that this person, embodied in the same man as me, who has literally lived my life for the last fifty-five years, would be somebody else. This very thought, realized by these neurological processes in this brain, could have belonged to somebody else. Or, if we insist on individuating token mental-states by their

subject, the token neural-states my brain has undergone in the last fifty-five years could have produced, in response to the same environment, a completely different set of mental states. This toothache's identity depends on which mental states this brain realized more than a half-century ago.

These difficulties sink PC in its contest with BC—*unless* the latter leads to consequences at least as counter-intuitive. The most serious difficulty for BC is what Eric Olson calls the 'Transplant Intuition' (Olson, 1997). Suppose your cerebrum is destroyed; your body persists in a vegetative state, sustained by your lower brain, until O.J. Simpson's cerebrum (not his entire brain, which would bring along the animal (Olson 1997, 45)) is implanted in it. The resulting person is psychologically continuous with O.J., not you. He has O.J.'s apparent memories, his personality, and he acts on O.J.'s intentions. But that person is *you*, the person at the beginning of the story, according to BC. This is strongly counter-intuitive, surely.

But does BC really lead to this consequence? Suppose that my brain is transplanted to the brainless body of my identical twin. Most philosophers would agree that the psychologically connected person who results is *me*. BC can accommodate this intuition (Olson 1997, 45), for it's reasonable to view brain transplant as the limiting case of amputation and transplant, that is, as body transplant. A man goes where his brain goes. If my brain is removed and thrown away, the animal is destroyed even if the brainless body is kept alive artificially—just as if we destroyed all of me but my liver and kept it alive somehow.

It might be tempting to identify me with my brain; I would resist this. On the face of things I'm an animal, not an animal's organ. I have hands and feet as proper parts. Roughly it's sufficient for an animal to persist that there are non-branching spatio-temporally continuous maximal aggregates of cells functioning collectively according to a genetic developmental blueprint; consequently I once existed without a brain—when I was a blastula, say. Once the human animal is embrained, however, he goes where his brain goes—not because he's identical to his brain but because the embrained animal can survive the amputation of everything but his brain. That material thing T persists if pared down to proper part P doesn't compel us to identify T with P. Absent a good reason not to do so, Reductionists ought to identify persons with the most intuitive candidate.

What about the identity conditions for human brains? Certainly it seems reasonable that if one excises most of the brain's matter at a go and substitutes a chunk of new brain-cells, the result is a numerically distinct brain. Consequently the result of the cerebrum transplant is that my brain is destroyed, for the cerebrum comprises about two-thirds of the brain. Therefore the animal is destroyed—so I cease to exist. On the other hand if the majority of the brain's matter remains intact and functional, the brain survives and so

does the animal. A consequence of BC, therefore, is that in the cerebrum transplant case I do *not* survive—unless, of course, my cerebrum is also transplanted—but O.J. does. On the other hand, if my cerebrum is transplanted to O.J.'s body, then I survive, too. In short, it's reasonable to treat cerebrum transplant as if it is brain transplant; for the brain cannot survive the sudden replacement of most of its matter, but it can survive the loss of some of its matter if most of its matter remains intact and functional.<sup>5</sup>

To sum up: once what matters in survival is separated from identity, the principal argument in PC's favor is undercut. Further, PC, not BC, leads to strongly counter-intuitive consequences. The Reductionist must therefore accept BC (or, less intuitively, identify human persons with brains or cerebra; my argument will go through anyway, so I set these options aside). This much is plain: As Reductionism entails that identity isn't what matters, the Reductionist is ill-positioned to object to our deploying BC in the argument that follows.

## II

Consider this case. I suffer from headaches. I'm examined by a neurologist who tells me that I have a brain tumor which must be removed immediately. This can be done by a single radiation treatment dispensed by a Radiation Machine. The treatment is painless, however it will erase all the information encoded in my cerebrum—my memories, my personality, everything I know—though the rest of my brain won't be affected. This neurally encoded information supervenes upon the molecular structure of my cerebrum. The treatment undoes that structure without damaging my brain, leaving me with the mind of a newborn infant—educable, but blank. Fortunately the Radiation Machine also records the precise molecular structure of my brain. The Machine's computer will transform that information into a set of instructions. Three seconds after it erases my personality, the neurologist will press button A and the Machine, by following these instructions, will imprint that same pattern on my brain again, restoring all the information it erased. I will lose consciousness for three seconds. When I wake up, I will be psychologically connected to myself as I was three seconds ago. I will resemble myself on account of a reliable cause and, though it isn't the normal one, this is insufficient reason to conclude that I will lose what matters in survival. I can anticipate my vacation next week just as much as if I had declined the procedure.

---

<sup>5</sup> The case where my cerebrum is removed and destroyed, and the lower brain sustains the body in a PVS, involves neither the replacement nor the transplantation of brain matter. The sufficient condition that I mentioned above for a brain's being destroyed isn't satisfied. So far nothing precludes our saying that the brain survives cerebrum removal *simpliciter*.

The neurologist confides that the Machine is also being used in research aimed at creating the physical basis for a complete personality. On the Machine's computer's hard drive are instructions (a, b, c, d...) for rearranging cerebrum molecules to encode the information for a personality which the computer generated randomly one year ago. Indeed, there is a button B on the console—push B and the Machine will rearrange the molecules of my brain according to this other set of instructions. By an extraordinary coincidence, that randomly generated personality is mine, down to the last detail. She knows this because the examination she gave me included a reading of my brain's molecular structure. Pushing B will rearrange the molecules of my brain in exactly the way that pushing A will. The instructions will have been recorded twice on the machine's hard drive, the first time one year ago, the second three seconds ago. The machine will follow a, b, c, d... whichever button gets pressed. The neurologist assures me that of course she has no intention of pushing B.

An hour later I'm lying on a table waiting for the procedure to begin, when she staggers by on her way to the control console in the next room. She's drunk as a lord, and she's muttering: "Let's see...I'm supposed to push B. Or ish it A? Or ish it B?" Plainly B is as likely to be pushed as A. How should I view this situation? It makes no sense to be frightened that she will push B. If she pushes A, the Machine will follow instructions a, b, c, d..., rearranging the molecules of my cerebrum in a certain way; if she pushes B, the Machine will follow a, b, c, d..., rearranging the same molecules in the same way. It would be irrational to think "If in three seconds she pushes A and the molecules in my cerebrum are arranged according to a, b, c, d..., producing a personality just like the one I had before, that is as good as survival; but if she pushes B and the molecules are arranged according to a, b, c, d..., producing a personality just like the one I had before, that is as bad as death." The difference between pushing A and B cannot plausibly make that difference. If pushing A gets me what matters in survival, so does pushing B. I can anticipate my vacation just as rationally whichever button she presses.

Of course there will be some differences. If the neurologist pushes A, my memories of visiting Afghanistan in 1972 will be veridical; if she pushes B, they won't be memories at all—for I won't have them because I visited Afghanistan. Indeed, if she pushes B I won't remember my past life at all. Still I will have the same token 'memories' whichever button she pushes, and the same information about my past. As that information will be just as reliable if she pushes B as if she pushes A—either way, what I have will be just as good as memory—this difference provides no reason to view the pushing of B as less in my interest than the pushing of A.

Of course if she presses B, when I leave the hospital I will not be as I am *because* I was as I was. If the computer had generated a different personality,

I might be speaking Chinese and ‘remembering’ operating a fishing boat off Taiwan. It’s extraordinary luck that the computer generated my personality, certainly. But all that is well in the past. As I lie on the table awaiting the procedure, there is no possibility whatsoever that I will leave the hospital speaking Chinese. The machine is going to rearrange the molecules of my brain according to a, b, c, d..., producing the same personality, come what may. The situation might well produce a sense of vertigo—if the computer had selected differently, I could have lost almost everything I value about myself. But it didn’t select differently, thank heaven! To the extent that luck was involved, it was definitely good luck. Nothing of value is now at risk.

I might feel a thrill of fear at the thought: “Maybe the person who leaves the hospital if B is pressed won’t be *me!*” As psychological connectedness will be severed, perhaps he will be somebody else in my body who is just like me. We’re supposing, however, that identity doesn’t matter in survival; so connectedness cannot matter because it makes for identity—nor can pressing A. As identity isn’t what matters, my doubts about whether I will be identical to that person are beside the point. In addition, we’ve accepted BC which, once we’ve separated what matters from identity, is plainly superior to PC. The person who leaves the hospital will indeed be me, for he is the same man as me. As I observed above, the Reductionist is ill-positioned to object that we’ve deployed the wrong criterion of personal identity.

What about remorse and repentance? Suppose we find Adolph Hitler still alive; we charge him with genocide. We learn that in 1980 Hitler underwent a treatment like the one in my example; medical records report that B was pressed accidentally. Consequently he doesn’t remember committing his crimes, though he knows that his ‘memories’ are completely accurate. Hitler is consumed with remorse; he begs our forgiveness. It would be wrong-headed to tell him that, as connectedness is what matters in survival, not identity, his relation to the perpetrator of his crimes cannot warrant remorse—although remorse would certainly be in order if A had been pressed. Suppose we’re back in 1980, watching the procedure. It would be irrational to say: “If in three seconds the neurologist pushes A and the molecules of Hitler’s cerebrum are arranged according to e, f, g, h..., producing a personality just like the one he had before, then Hitler’s future remorse will be rational; but if she pushes B and the molecules are arranged according to e, f, g, h..., producing a personality just like the one he had before, then Hitler’s remorse will be as senseless as his feeling repentant for the crimes of somebody else who merely resembles him.” Imagine later trying to console Hitler on the ground that, after all, the neurologist pressed B, not A. If repentance and contrition are warranted whichever button gets pressed, so is moral responsibility. The difference between her rearranging the molecules of Hitler’s cerebrum according to e, f, g, h... by pushing B rather than A is too slender to absolve Hitler of



responsibility. It's unlucky for Hitler, but fortunate for the court, that the computer generated his old personality.

To sum up: If identity isn't what matters in survival, I would be irrational to fear that B, not A, will be pressed. If B is pressed, the person who leaves the hospital will be responsible for my past misdeeds; and it's rational for me to anticipate his experiences. Since I have what matters without it, psychological connectedness isn't what matters in survival. If identity isn't what matters, either something other than connectedness matters or nothing matters.

### III

The chief support for the claim that psychological connectedness, not identity, is what matters in survival flows from cases in which a man divides. Consider Teletransportation first. Parfit writes:

The Scanner here on Earth will destroy my brain and body, while recording the exact states of all of my cells.... Travelling at the speed of light, the message will take three minutes to reach the Replicator on Mars. This will then create, out of new matter, a brain and body exactly like mine. It will be in this body that I shall wake up (Parfit 1984, 199).

Like most people, I'm leery of Teletransportation. I'm afraid I'll be murdered by the Scanner; somebody else will be created on Mars. That's because there is a complete and sudden change of my matter, marked spatial discontinuity, plus a three-minute gap when I'm not located anywhere. Nor am I inclined to think that I ought to anticipate my Martian replica's experiences, fear his torture, and so on. In the absence of a clear intuition that he is me, the fact of psychological connectedness doesn't motivate the intuition that I have what matters in survival.

Contrast Division:

My body is fatally injured, as are the brains of my two brothers. My brain is divided, and each half is successfully transplanted into the body of one of my brothers. Each of the resulting people believes that he is me, seems to remember living my life, has my character, and is in every other way psychologically continuous with me. And he has a body that is very like mine (Parfit 1984, 254-255).

Here most people feel that they *do* have what matters. As my relationship to just one of the resulting people would contain what matters if I stood in it to him alone, how can its duplication be as bad as death? What's baffling about Division is that I'm convinced initially that some future person is me, yet on inspection there's no easy candidate. There is no ground for saying I'm just one of the two offshoots; and I'm not both of them, for then they would be identical to each other. Arguably here is a case where I have what matters in survival but no future person is me.

Yet why do we feel that we have what matters in Division but not in Teletransportation? Note that the features of Teletransportation that make me doubt my survival—the complete and sudden change of matter and the drastic collapse of spatio-temporal continuity—are absent in Division. Instead we have psychological connectedness plus physical continuity of enough of the central nervous system to support conscious life. As the sufficient condition for identity in ordinary cases is apparently satisfied, naturally we're more inclined to think we survive.

Indeed, might I not survive Division with a divided mind and body? Parfit rejects this option because it involves 'grotesque distortions in our concept of a person,' (Parfit 1984, 476) but (earlier) he owns that if the questions "Will I survive?" and "Will there be some person who is alive who is the same person as me?" are treated as equivalent, 'then the least unsatisfactory description of William's case is...that I survive with a divided body and a divided mind' (Parfit 1971, 205). Each offshoot is a person in his own right, but a person might be composed of two other persons, as the Pope's crown is composed of three other crowns (Parfit 1971, 203). Indeed, Parfit asserts that in Division 'my brain continues to exist, and, because it is divided...it supports not just one but two lives' (Parfit 1984, 476). The hemispheres, now brains in their own right, are still parts of the original brain. If the animal follows his brain, he also persists and supports two lives. We might say that the animal (A) survives as a scattered object, the two new animals (B and C) resulting from Division are A's parts.

Parfit asks: "Suppose the resulting people fight a duel. Are there three people fighting, one on each side and one on both?" (Parfit 1984, 257) Yes. It's reasonable to ascribe to A, who now supports two lives, the acts of his parts; so A can be standing and sitting just as a road can be curvy and straight. "And suppose one of the bullets kills. Are there two acts, one murder and one suicide?" (257) Suppose B shoots C with the intention of thereby destroying A; then so does A. Both B and A murder C. "How many people are left alive? One or two?" (257). The presumption that a man (or a brain) can survive a hemisphere's destruction is defeasible in extraordinary cases: if both A and B survive, coincident men have the same experiences. Nothing jeopardizes B. This gets us a second murder—B murders A—plus a suicide. I maintain only this: the idea that I survive with a divided mind and body isn't *so* bizarre that it cannot sustain the intuition that I persist.

The upshot is that Division does *not* support the claim that psychological connectedness is what matters in survival. That I'm connected to some future person is a constant in both Teletransportation and Division. As I don't feel that in Teletransportation I have what matters, connectedness doesn't explain why I *do* feel I have it in Division. That's explained by the intuition that I somehow survive Division (with a divided mind and body, perhaps), which in

turn is motivated by the apparent satisfaction of the sufficient condition for identity in ordinary cases. As “The intuition that in Division I have what matters depends upon the intuition that I survive” is at least as plausible as its converse, Division doesn’t support the view that what matters in survival is something other than identity.

#### IV

If neither identity nor connectedness matters, yet something still matters, what might it be? According to Peter Unger, what matters in survival is that some future person will have my core psychological properties on account of the ‘physical continuity of whatever physical entity, or whatever succession of such entities, then realizes those particular mental capacities,’ that is, my brain; the capacities must persist, even if I’m unconscious (Unger 1990, 109). ‘Core psychological properties’ are meager and widely shared: the capacities for consciousness, memory, and low-level reasoning. A future amnesiac moron could have my core psychological properties. Whether A or B is pressed, the resulting person will have my core psychology because he has my brain, which continues to realize my core psychology during the three seconds I’m unconscious. That’s why it doesn’t matter which button gets pressed. If the computer had selected differently, and a person embodied in this man had left the hospital speaking Tamil, with ‘memories’ of Jaffna, I would have as good a reason to anticipate his experiences as I do my own. Indeed, they *would* be my own, for identity just consists of non-branching brain-based continuity of core psychological properties.

Suppose the Radiation Machine works differently, however. In radiating my brain the Machine will damage (only) my cerebrum so that it loses its mental capacities. In the original case the machine leaves my brain like a baby’s, educable but blank. In this new case, however, the cerebrum’s capacity for memory, consciousness and reason are destroyed utterly; the ‘live meat’ that remains has the mental capacity of my liver. The neurons’ molecules are so arranged that the neurons’ capacity to function is destroyed, though they remain alive and structurally identifiable as neurons. Fortunately, in three seconds A will be pressed and the machine will rearrange my cerebrum’s molecules as they were before, restoring my old personality—otherwise I would linger in a PVS, sustained by my lower brain. That person will be psychologically connected to me, so (a fortiori) he will have core psychological properties just like mine. As my brain won’t realize those capacities during the three seconds I’m unconscious, however, they will *not* persist in a ‘physically continuous realizer.’

For Unger ‘what matters in survival’ has several different meanings—the one that concerns him is prudential concern, which he determines by ‘the

avoidance of future great pain test.’ In thought experiments concerning my alleged survival, I have just two options.

I may elect to experience...very considerable pain and...thus ensure that the being who emerges will...feel no pain. Alternatively, I may elect to endure no pain before the process begins and...thus ensure that the being who emerges from the process will, after awakening, undergo really excruciating tortures for quite a long time. (Unger 1990, 28-29)

What results does the test yield when applied to the new thought experiment? *Pace* Unger, I would endure considerable pain before the radiation treatment in order to ensure that the being who emerges will not be tortured. This is because I would view the outcome as my own continued life. I’m psychologically connected to a single future person who is embodied in the same man as I am. That my cerebrum was so damaged that it briefly lost its mental powers seems unimportant given the outcome. My lower brain kept on functioning normally, no matter was removed from my brain, so my brain survives. I have about as much reason for prudential concern as I ordinarily do, so I have what matters in survival without a physically continuous realizer of my core psychology.

Is the outcome mischaracterized? There’s a new cerebrum, arguably, configured from the matter of the one that ceased to exist three seconds earlier. A new cerebrum makes for a new brain which makes for a new man. However the cerebrum at t4 is made of the same molecules as the cerebrum at t1, and it does seem plain that there’s a continuing material thing (C) composed of those molecules that preceded the three second span and continues past it. If C isn’t the cerebrum, then both the t1 and t4 cerebra are coincident with C, and realize the same thoughts as C. It’s simpler to identify them with C; the cerebrum is the brain-part that in fact has those powers, but it preceded them (when it formed in the embryo) and it can lose and regain them—even as a corpse’s hand is still a hand and at the resurrection will be the same hand. This is an intuitive way to talk, it’s simpler ontologically, and we have no reason not to adopt it.

Perhaps what really matters in survival, however, is a *causally* continuous physical realization of my core psychological properties: my cerebrum, the hard drive of the radiation machine, then my cerebrum again. At every moment in the process some physical thing or other has the ability to realize my core mental capacities (if not in its own matter, then in that of something else), and the capacity of each physical thing in the chain is explained by that of an earlier member. Arguably causal continuity is the continuity that matters.

Now add this wrinkle. Will the neurologist press A or B? If identity isn’t what matters in survival, I have nothing to fear. The molecules of my cerebrum will be arranged according to a, b, c, d... whichever button she presses, producing a personality just like my old one. The difference between so

arranging them by pressing B rather than A cannot plausibly make the difference between an outcome as good as survival and one as bad as sudden death. But if B is pressed, there's no causally continuous chain of physical realizers of my core psychological properties. Yet I have what matters in survival, even by the 'avoidance of future great pain' test. For the only consideration that might dissuade me from suffering intense pain now in order to avoid excruciating torture later is the thought that the person resulting from pushing B may not be me. If identity isn't what matters in survival, I would certainly make the sacrifice.

To sum up: If identity isn't what matters in survival, either a relation other than connectedness matters or nothing matters. Given the failure of Unger's thesis, it's hard to see what the 'something other' could be. Not that some future person will have core psychological properties like mine: decapitation would be as good as continued life. Not that some future person's personality is *exactly* like mine; I have no reason to anticipate the experiences of a coincidental replica. The thought experiments show that the existence of a causal chain from me to that future person, accounting for the resemblance or just for his having my core psychology, isn't what matters—not even when it flows through a physically continuous realizer. I draw the warranted conclusion that if identity isn't what matters in survival, nothing matters.

## V

As John Locke observed: "[Person] is a forensic term, appropriating actions and their merit." (Locke, *Essay Concerning Human Understanding*, book II, ch. 27). Persons are conceived as responsibility bearers, capable of rational hopes, fears, and remorse. If nothing matters in survival, there are no persons.<sup>6</sup> Identity is what matters or nothing matters; therefore identity is what matters in survival—if there are persons. As I argued earlier, it follows that persons are ontologically extra to bodies, brains, and psychophysical events. The most plausible remaining account of persons is that they are souls. A consequence of Naturalism, therefore, is that there are no persons. But I am a

---

<sup>6</sup> The Reductionist might object that there will still be 'thinking intelligent beings that can consider themselves as themselves in different times and places.' However if this Lockean definition of 'person' is to be satisfied there must be at least one intelligent being who exists at t1 and t2, and who recognizes at t1 that he is identical to the being that exists at t2 (or vice versa). But if at t1 I recognize that the intelligent being at t2 is identical to me, then if I know he will have experience E, then it is rational for me to anticipate experiencing E. If Reductionism is true, it is never rational for me to anticipate an experience I know a future person will have. Therefore I never recognize that the being at t2 is identical to me. Further, if at t2 I know that the being at t1 is identical to me, then if I know that he performed a destructive and unnecessary action then it is rational for me to regret having performed it. But if Reductionism is true, it is never rational for me to regret performing an act I know a past being performed. Therefore I never recognize that the intelligent being at t1 is identical to me. Consequently if Reductionism is true, the Lockean definition is never satisfied (see Stone 1988, 530).

person if I am anything: a consequence of Naturalism is that I don't exist. There is nothing in nature for me to be. Of course I might identify myself with something that comes and goes in a moment, the transient subject of this very thought, perhaps (Strawson 1994, 133-134). But having come so far, it is ontologically simpler, and more consistent with what is revealed in introspection, to say that there is, after all, only the thought.

Doesn't this provide a knock-down argument against Naturalism? Obviously there are people; Naturalism precludes them, so there are souls! But perhaps the victory is Pyrrhic, for now the existence of people is no more obvious than the existence of souls: the objection's first premiss is false. I cannot do justice here to the arguments of the Friends of Souls; nor will I rehearse all the difficulties for Dualism. Instead consider how far we go together. Realists and Eliminativists agree that persons must be souls. We agree that Reductionism is metaphysically incoherent because it tries to ground persons in an ontological base too impersonal to support them. Also, we share a solution to the Division puzzle. If I'm a soul, I go where it goes (Swinburne in Shoemaker and Swinburne, 1984). If it goes with one hemisphere, that life is mine. If it ceases to exist, I lose what matters. Souls are conceived as simples, so they can't divide. If I'm *not* a soul, I never existed to begin with, so nothing happens to me when I fission. Generally we can transform metaphysical persistence-puzzles into mere epistemic difficulties by positing an underlying Ship of Theseus, say, which goes either to the ship we assemble out of Ship's discarded planks or to Ship's spatio-temporal continuer. The idea that persons underlie the flux and diversity of human lives, and go we-know-not-where in puzzle cases, strikes me as plausible as the idea that there is an underlying Ship of Theseus.

As there is no middle-ground between Realism and Eliminativism, the most persuasive argument for Realism is that Eliminativism (but not Realism) is incredible. Since I've foresworn attacking Realism, let me try to make Eliminativism less incredible. Some philosophers find Eliminativism incoherent on its face. According to Parfit, 'an outright denial of the existence of persons is of course absurd' for entirely grammatical reasons (Parfit 1987, 312). He writes: "There are persons or subjects in this language-dependent way." (312) But the Eliminativist's denial of persons involves no grammatical absurdity. Certainly my assertion "I don't exist" sounds ungrammatical: *who* doesn't exist? However 'I' can be construed as 'the thinker of this very thought'; and "The thinker of this very thought doesn't exist" Russell's nicely into the non-absurd: "It is not the case that there is one and only one thinker of this very thought." Eliminativism's absurdity is unlikely to be found in grammar.

The Eliminativist can still talk as *if* there are people, however, even though much of what she says will be literally false. When I utter a sentence

like 'I believe there are no persons,' the proposition I assert (that there is just one subject of this thought and it believes that there are no persons) is false, but uttering it expresses indirectly the truth that the belief that there are no persons is causing the utterance. Talking as if there are subjects is useful, for it often tracks features of reality that it would be difficult to express otherwise. Once again it is skillful to 'think with the learned and speak with the vulgar.'

Galen Strawson believes it's a metaphysical truth, not a grammatical illusion, that experience is necessarily for a subject of experience.

Experience necessarily involves experiential what-it-is-likeness *for* someone or something. Whatever the nature of this experiencing something, its real existence cannot be denied (Strawson 1994, 129).

Experience requires subjectivity, but does subjectivity require a subject? David Rosenthal argues that there is something it's like to have a sensory experience just in case it is targeted by a higher-order state (Rosenthal, 1997). (Rosenthal takes HOSs to be thoughts, but Locke, Armstrong, and Lycan identify HOSs with inner perceptions.) For Rosenthal, subjectivity gets cashed-in in terms of consciousness, which is explained in terms of HOSs. An Eliminativist can plausibly argue that the claim "What-it-is-likeness is necessarily what-it-is-likeness-*for* someone" is really tracking another fact: that what-it-is-likeness necessarily involves the sensory state's being conscious, which requires only that it is targeted by a HOS. This is the truth that motivates Strawson's claim. Of course it's natural to express it in terms of subjects, for we generally talk as though mental states belong to subjects. In a universe consisting only of mental states, however, some of which are HOSs, everything required for what-it-is-likeness would be present.

Alternatively, Thomas Nagel writes that there is something it's like for the subject (a bat, say) to have a mental state only if there's something it's like from her 'point of view' (Nagel, 1979). Subjectivity gets cashed-in in terms of point of view. The bat's perceptual experiences represent the world from a certain perspective, and they represent it in virtue of certain aspects. *That* is the bat's point of view, I've argued; consequently there's something it's like to be a bat even if bats lack HOSs (Stone, 2001). Otherwise it's doubtful that bats have experiences, which is not what Nagel wants to say. Subtract the bat from the experiential stream and her point of view remains.

We needn't decide between these rivals. The whole point is that Eliminativism has ample resources to explain subjectivity in terms of the intentionality of mental states and/or their relations to one another. Arguably the facts in virtue of which we say that 'an experience has a qualitative aspect only if there is something it's like for someone to have it' can be expressed without

mentioning subjects. So Strawson's appeal to 'what it is like *for...*' isn't a decisive argument for the thesis that an experience requires an experiencer.

Of course Eliminativism still flies in the face of the most fundamental datum of common sense. We exist. But common sense is often a casualty of science, while Eliminativism, if I'm right, has behind it the whole force of Naturalism. Why, then, do we all believe that we exist? The answer given by Eliminativists is that the sense of self is a delusion borne of human psychology. The delusion can be overcome, at least briefly, by attending to the mind-body process, which interrupts the slurring together of successive mental states to which David Hume credited the idea of self. Attention can also dissolve the powerful attachment to mind and body, and the resulting craving for permanence, of which the sense of self is an artifact—as the Buddha taught. To spend a long while attending to the mind-body process is to be struck pretty forcefully with the insight that no one is in it. There is a sense of awakening from a dream. Even the HOSs that comprise introspection are fleeting and impersonal, arising due to the causation of other transient states. A consequence of Hume's view is that mindfulness meditation will produce just this insight. As the 'delusion of self' is dissolved by practices that undercut what, according to plausible psychological explanations, are its cause, the appeal to common sense is inconclusive, too.

## VI

Why, then, is it so hard to accept the consequences of an impersonal universe? There is, of course, the bafflement of our self-concern. In short, what becomes of *me*? Of course, given what will happen to the 'self-cherishing I' pretty soon anyway, the idea that it never existed has its consolations. As human life is largely a bafflement of our self-concern, these increase the more the idea is accepted. We're now fishing in the right waters, at any rate. What makes Eliminativism really hard to swallow is that morality is destroyed. Right and wrong cannot exist in an utterly impersonal universe. Who was harmed by the Holocaust, after all? None of it happened to anybody! It would be crazy to accept such consequences, surely.

An immediate response is that Utilitarianism survives, for it quantifies only over mental states, not subjects. An insight of classical Utilitarianism is that pleasant mental states are intrinsically better than painful ones—not better for anybody, just better *simpliciter*. An act is right if it makes the world a happier place. If 'acts exist, but the person who acts does not,' there are right acts and wrong acts. Also there are good and bad intentions, for nothing in Eliminativism precludes actions motivated by a desire for a better world. A more serious difficulty is that justice doesn't supervene on utility. There are fundamental human rights, contracts and obligations, special duties to our families and countrymen, which we violate if we always do what's



most useful. Rights bearers are required, people wronged by promises broken, persisting subjects whom death would deprive of a future. How can deontological ethics survive the loss of subjects?

Perhaps Buddhism contains an answer. The Buddha knew that the universe is impersonal through and through; he wasn't deluded at all. Yet he attacked the caste system because he recognized its injustice. How was he able to do this? The delusion of self is an artifact of craving, attachment, and inattention; it dissolves as they do. In their place a powerful compassion arises—at least for those 'well established in virtue.'<sup>7</sup> Suppose my paranoid friend believes himself to be the object of a conspiracy. If I'm compassionate I will enter into that delusion, until it becomes sufficiently real that he knows that I know what it's like for him. A natural activity of compassion is to enter into the delusion that causes pain; it takes up the point of view of delusion, investing its intentional object with a vivacity that makes it emotionally real. As that object's emotional reality flows from the activity of compassion, not from the perception or cognition that the object is real, it is consistent with recognizing the delusion as delusion. Compassion can create its own object.

An awakening Buddha finds herself surrounded by transient minds and bodies possessed of the delusion that they contain persons, which involves a cycle of pleasure and pain. Compassion enters into the delusion, investing it with emotional reality, thereby creating persons whom compassion loves. The Buddha spent most of his life bringing to people he knew didn't exist the liberating news that they never were. He did this from compassion for them. Compassion enabled him to recognize the caste system's unfairness to people he knew didn't exist.

According to Buddhism, wisdom and compassion support each other 'like two hands that wash one another.' The realization that my brother is a bag of bones finds its complete expression in compassion for my brother. Conversely, to care for somebody all the way down to her metaphysical ground is to recognize her emptiness. From the perspective of compassion, the world is full of suffering beings crying out for protection and for justice. From the perspective of wisdom, while the natural order contains terrific pain, no one suffers it. Without compassion, wisdom would be other-worldly, less than wise. Without wisdom the compassionate heart would be broken. From the perspective of compassion, knowing the universe is impersonal is a reason to be kindly and fair. A Buddhist saying goes: "Ultimately there is neither right nor wrong, but right is right and wrong is wrong."

Kant believed that morality must be founded on respect for persons. I suggest that morality is better founded on compassion. The sentiment of respect

---

<sup>7</sup> Buddhist practice is devoted largely to developing virtues like patience, kindness, and compassion, which provide the indispensable setting for enlightenment.

overly narrows the moral community—it excludes infants, fetuses, and non-human animals. Compassion populates the universe with moral objects. In Mahayana Buddhism, the Bodhisattva vows to be reborn for the sake of all sentient beings—'yea, until the last blade of grass is liberated!' Of the two sentiments, compassion is less fragile ontologically: it motivates a satisfying morality even in an impersonal universe like ours.<sup>8</sup>

---

<sup>8</sup> Thanks for comments to John Cottingham, Eric Olson, Galen Strawson, and to anonymous referees from *Philosophy and Phenomenological Research*. Thanks to Dean Stretton for helpful correspondence. Special thanks to Judith Crane.

## Bibliography

- Block, Ned; Flanagan, Owen; Guzeldere, Guven, eds. (1997), *The Nature of Consciousness* (Cambridge, Mass.: Bradford, MIT Press).
- Chisholm, Roderick M. (1978), 'Is There a Mind-Body Problem' *The Philosophical Exchange*, 2, no. 4, 25-34. In van Inwagen and Zimmerman (1998), 291-296.
- Danto, A. C. (1967): 'Naturalism', in P. Edwards, ed., *The Encyclopedia of Philosophy* (New York: Macmillan), pp. 448-450.
- Nagel, Thomas (1979), 'What is it like to be a bat?', in Thomas Nagel (1979), *Mortal Questions* (Cambridge: Cambridge University Press), pp. 165-179. Originally published in *The Philosophical Review*, LXXXIII, 4 (October 1974), pp. 435-450.
- Olson, Eric T. (1997), *The Human Animal* (New York: Oxford University Press).
- Parfit, Derek (1971), 'Personal Identity,' *Philosophical Review*, vol. 80, no. 1 (January, 1971). In Perry (1975), 199-224.
- Parfit, Derek (1984), *Reasons and Persons* (Oxford: Oxford University Press).
- Parfit, Derek (1987) 'Divided Minds and the Nature of Persons' in Blakemore, Colin and Greenfield, Susan, eds. (1987) *Mindwaves* (Oxford: Blackwell). Reprinted in van Inwagen and Zimmerman (1997), 310-317.
- Perry, John ed. (1975), *Personal Identity* (Berkeley: University of California Press).
- Rosenthal, David M. (1997), 'A Theory of Consciousness,' in Block (1997), 729-753.
- Shoemaker, Sydney and Swinburne, Richard (1984), *Personal Identity* (Oxford: Blackwell).
- Stone, Jim, (1987), 'Why Potentiality Matters,' *Canadian Journal of Philosophy*, vol. 17, no. 4 (December 1987), 815-830.
- Stone, Jim (1988), 'Parfit and the Buddha: Why There Are No People,' *Philosophy and Phenomenological Research*, vol. 48, no. 3 (March, 1988), 519-532.
- Stone, Jim (2001), "What is it Like to Have an Unconscious Mental State?" *Philosophical Studies*, 104 (2001), 179-202.
- Strawson, Galen (1994), *Mental Reality* (Cambridge: MIT Press).
- Unger, Peter (1990), *Identity, Consciousness and Value* (New York: Oxford University Press).
- van Inwagen, Peter and Zimmerman, Dean, eds. (1998), *Metaphysics: The Big Questions* (Oxford: Blackwell Publishers).