

## Would it matter all that much if there were no selves?

(31 May 2006 Draft)

In commentaries both on Buddhism and on some western philosophers, such as Hume, three ideas—I'll call them *no-self ideas*—are sometimes discussed interchangeably, as if they were the same idea. The ideas are:

- §           that the self is an illusion,
- §           that is no such thing as the self, and
- §           that there is no such thing as a substantial and enduring self.

In the voluminous discussion of these ideas, it is often assumed, both by proponents and critics, that if one or another of them were true, it would be a big deal. Buddhist proponents of the idea that there is no self, for instance, often express the view that belief in the self, in the way in which almost everyone believes in it, is a major source of suffering; Buddhists also often express the view that the experiential realization that there is no self is life-changing in an extremely positive way. Western critics of the idea that there is no self sometimes express the view that there being no self would have profoundly disturbing implications for how one ought to behave, such as its then being the case that there would be no reason for future-oriented egoistic concern, including prudence. My thesis, in the present paper, is that the truth of any, or all, of these no-self ideas would not be a big deal—at least not for the reasons that it is usually thought that it would be a big deal. I shall suggest at the end that there is an idea that is closely related to the idea that there is no self that, if it were true, would be a big deal. I believe, but shall not argue, that this latter idea is true.

It is important to realize that the three no-self ideas listed really are three, and not just one. For instance, the idea that the self is an illusion, in the way in which this idea is usually understood, implies that many people experience the world as if they were a self even though they are not a self.<sup>1</sup> The other two no-self ideas do not imply anything about how anyone experiences the world. The idea that there is no such thing as a substantial and enduring self leaves open the possibility that there are enduring selves

that are not substantial, and substantial selves that are not enduring. Hence, it leaves open the possibility that there are selves. And so on. So, the three no-self ideas really are three. And, of course, since they are three they may have different implications for how one ought to behave.

In the west, one of the earliest indications of interest in the question of whether there is a substantial and enduring self occurs in a scene from a play, written in the fifth century B.C.E., by the comic playwright Epicharmus. In this scene a lender asks a debtor to pay up. The debtor replies by asking the lender whether he agrees that anything that undergoes change, such as a pile of pebbles to which one pebble has been added or removed, thereby becomes a different thing. The lender says that he agrees with that. “Well, then,” says the debtor, “aren’t people constantly undergoing changes?” “Yes,” replies the lender. “So,” says the debtor, “it follows that I’m not the same person as the one who was indebted to you and, so, I owe you nothing.” The lender then hits the debtor, knocking him to the ground. The debtor then protests loudly at being thus abused. The lender replies that the debtor’s complaint is misdirected since he—the lender—is not the same person as the one who struck him a moment before.<sup>2</sup>

An interesting—some might say, borderline amazing—thing about this scene is that it suggests that even in fifth century B.C.E. Greece, the puzzle of what it is about a thing that accounts for its persisting over time and through changes could be appreciated even by *theater audiences*. Another interesting thing about the scene is its content: both the debtor and the lender have a point. Everyone *is* constantly changing. In a very strict sense of *same person*, every time someone changes, even a little, he or she ceases to exist: so, in this very strict sense of *same person*, the debtor is *not* the same person as the one who borrowed the money, and the lender is *not* the same person as the one who hit the debtor. Obviously, this very strict sense of *same person* is not an everyday notion, but a philosophical term of art. Also obviously, it is not a very useful sense of *same person*—unless you owe someone money!

In everyday life, we want to be able to say such things as, “I saw you at the play last night,” and have what we say be true. If everyone is constantly changing and every

change in a person results in his or her ceasing to exist, no such remarks could ever be true. On the assumption that such remarks sometimes are true, there must be a sense of *same person* according to which someone can remain the same person *in spite of changing*. Saying what this sense is, or what these senses are, is *the philosophical problem of personal identity* (over time).<sup>3</sup>

However the problem of personal identity is resolved, it is important to remember that so far as our current discussion so far is concerned, the assumption that such remarks as, "I saw you at the play last night," sometimes are literally true is just that—an assumption. I have supported it only by pointing out that one sense of *same person*—what I called *the very strict sense of same person*—is not very useful. But this very strict sense of same person's not being very useful is different from its not being true. It could be true, but not useful. Or, what seems more likely, there may be no truth of the matter about whether any particular view of same person over time is true. Perhaps which view, if any, of same person over time is most plausible is only a question of utility, and there is nothing more to it than that. And, of course, a sense of same person that is more useful for one purpose may be less useful for another. So, relativism may be part of the picture.

Would that be troubling? In the case of potential objects other than self, or person (in the present paper, I shall use the terms, *self* and *person*, interchangeably, to express the same idea), probably a similar suspicion would be troubling, if at all, only theoretically. In the case of self, or person, it may also be troubling in ways that strike closer to home. For instance, if it is not true, but only at best useful, that each of us exists for more than a moment, then what becomes of the stories we tell ourselves about our personal pasts? Are they too only useful? And what becomes of future-oriented egoistic concern? Is it merely useful? To whom? In short, if our commitment to personal persistence in such stories and concerns is merely, say, a potentially useful manner of speaking, then it would seem that we ought to be able to let that commitment go by the boards and still make rational sense of something like a normal pattern of human values. But it's not clear that we could let that commitment go by the boards and

still do that.

In Western philosophy, the worry that the self, or person, is a fiction entered the tradition in the late seventeenth century in John Locke's famous suggestion that *person* is a forensic term. Locke made this suggestion immediately after giving a relational account of personal identity over time. These two proposals—that the self is a fiction and that as a fiction it should be understood relationally—tended to be lumped together in the minds of many of Locke's eighteenth century critics. Bishop Butler, for instance, said that if selves were to consist only in successive acts of consciousness, then it would be a mistake “to charge our present selves with anything we did, or to imagine our present selves interested in anything which befell us yesterday [or] will befall us to-morrow.” Under such circumstances, he continued, “our present self is not, in reality, the same with the self of yesterday, but another like self or person coming in its room, and mistaken for it; to which another self will succeed tomorrow.”<sup>4</sup> Thus, according to Butler, in Locke's view each of us would be persisting selves only in a fictitious sense. Butler thought that this consequence *refutes* Locke's view, but not that it *proves that* it is false. Rather, he said, “the bare unfolding this notion [that selves are fictitious] and laying it thus naked and open, seems the best confutation” of Locke's view.<sup>5</sup>

Most of Locke's original critics, including Bishop Butler, tried to defeat Locke's radical suggestion in order to save the traditional idea that the self is an immaterial, indivisible, and hence naturally immortal substance—a soul. Of course, at the time (and even now) the soul idea had great significance religiously. So, from a contemporary, secular point of view the debate between Locke and his early critics can seem like a battle between religion and science.

In Buddhism, the radical suggestion that the self is a fiction often gets expressed as the view that it is conventionally, but not ultimately, true that selves, or persons, exist. Many western philosophers with reductionistic proclivities would be comfortable with this much of Buddhism.<sup>6</sup> That is, they would be comfortable with reducing persons, without remainder, to sub-personal parts. In Derek Parfit's version of this view, for instance, the existence of a self, or person, just consists in the existence of a brain and body, and

interrelated physical and mental events.<sup>7</sup> Buddhists often go farther and hold that the sub-personal things and relations into which Western philosophers would like to reduce the self are themselves fictions. In other words, in the view of these Buddhists, just as it is merely conventionally true that selves exist, so too it is merely conventionally true, that brains, bodies, and interrelated physical and psychological events exist.

Few Western philosophers would be willing to go this far. Rather, most would insist on making sense of something pretty close to normal human values from a point of view according to which brains, bodies, and interrelated physical and psychological events exist not just conventionally, but actually. Hence, some of them are reluctant to admit that selves or persons are fictional, or if they admit this, are reluctant to admit that they are fictional merely because they strongly supervene on sub-personal parts and relations. In short, whereas the original critics of Locke's radical suggestion that the self is a fiction tried to defeat his view in order to save the traditional idea that the self is an immaterial, indivisible, and hence naturally immortal soul, contemporary critics of the view that selves, or persons, are fictions tend to be non-reductionistic materialists. As a consequence, the philosophical battle is no longer between religion and science, but over how best to understand notions such as *supervenience* and *realization*.<sup>8</sup>

It may seem that the Buddhist idea that everything to which we might reduce the self is at best only conventionally real is more radical than the contemporary western idea that the self is reducible to sub-personal parts that are real. But it may not be more radical. On the contemporary western view, there is something special metaphysically about the self: it seems to be a thing, but is not a thing, in a way that distinguishes it from the sub-personal parts into which it may be decomposed, which are things. On versions of Buddhism according to which *everything* goes into the same hopper, there is nothing special metaphysically about the self. This said, it seems to me that in both traditional and contemporary Buddhism, many theorists want to say that there is something special about the self—perhaps not metaphysically special, but special in the way that it is an illusion or in the role that as an illusion it plays in our relationships to ourselves and the world.

Among contemporary western philosophers, the question is whether and, if so, to what degree and in what way, selves, or persons, are expendable. Many contemporary analytic philosophers, including Mark Johnston and John McDowell, argue that selves are not expendable—that is, that one needs them in order to make sense either of human values or of epistemology, or both.<sup>9</sup> However, contemporary western philosophers tend to hedge their bets. For instance, in response to Parfit’s arguments that there is no “further fact” to the existence of selves, or persons, Mark Johnston, in an effort to distance himself from Cartesians, calls the selves to which he thinks one should be committed, *ordinary further facts*, as opposed to *superlative further facts*.<sup>10</sup>

By *reductionism* about the self (hereafter, simply *reductionism*), I shall mean, following Parfit’s formulation, the view that the existence of a self, or person, just consists in the existence of a brain and body, and the occurrence of a series of interrelated physical and mental events.<sup>11</sup> Reductionism, so-defined, is a metaphysical view and as such does not commit one to any particular normative theory or to any particular view about how on relational grounds personal identity over time should be understood. Specifically, it does not imply that personal identity is not what matters primarily in survival, or that what does matter, ala Parfit, is psychological connectedness and continuity, with any cause. These latter claims are in addition to reductionism

Even so, reductionism is a controversial thesis, particularly with respect to the rationality of future-oriented egoistic concern. The main challenge is what Parfit calls *the Extreme Claim*, that is, the claim that if reductionism were true, then people would have no reason to be especially concerned about their own futures. Parfit contrasts the Extreme Claim with what he calls *the Moderate Claim*, which is the claim that Relation R, which Parfit defines as psychological continuity and connectedness with any cause, gives one reasons for special concern.<sup>12</sup>

The intuition on which the Extreme Claim seems to rely, and what has made it seem plausible to many critics of reductionism, is that future stages of the series of events that on a reductionist view constitute the self are, in effect, *others*. As a consequence, it has often seemed, current stages of the series would have no *egoistic*

reason to be especially concerned about subsequent stages.<sup>13</sup> Parfit argues that one has moral, rather than self-interested, reasons to be concerned about future stages of oneself with whom one is not closely psychologically connected. Closely related to the question of what justifies future oriented egoistic concern, but much less discussed, is the question of what entitles one to anticipate *having* the experiences of one's continuers—of one's future selves, if you will.

In opposition to the Extreme Claim, in my view:

- § reductionists can appeal to continuer-interest, in place of self-interest, to justify the rationality of a future-oriented surrogate of *egoistic* concern;
- § reductionists can appeal to continuer-interest, instead of morality, to justify a continuer-interested surrogate of prudence; and,
- § the only practical concession that reductionists need to make to personhood conventions is to adopt an attitude of what I shall call *thin-ironic engagement*.<sup>14</sup>

In sum, what I hope to show is that with respect to some of the normative issues that have been most discussed in the analytic literature, reductionism is a paper tiger. By itself, it has little power to motivate revisionist views about what one should or should not value, or about how one should or should not behave. Its power, I shall suggest at the end, is that by forcing us to face the insubstantiality of the self, it encourages us to face the seeming arbitrariness of a core component of our future-oriented egoistic concern, namely, our anticipation of *having* the experiences of our continuers.

## I

### **self-interest vs. continuer-interest**

#### **as a justification of future oriented egoistic concern**

*The Extreme Claim* has been proposed as a problem for reductionists that does not arise for non-reductionists (henceforth, *selfists*). In my view, it does arise for at least some selfists, in pretty much the same form as it arises for reductionists. So, if the Extreme Claim is a problem for reductionists, it is also a problem for these selfists.

If one is a selfist about persons, then one believes that each of us persists into

future as the same persons we are now in virtue of some “further fact” that is over and above what a reductionist would acknowledge to exist. The most extravagant version of such a further fact is a Cartesian Ego. Hence, on a selfist view, it has seemed, future-stages of one’s self are not others, but oneself, and there is not an issue either about what justifies future oriented egoistic concern or about our entitlement to anticipate having the experiences of ourselves in the future.

But this seeming advantage of selfism does not accrue to all kinds of selfism. Metaphysical punctualists (or episodics), who believe that the self is real, but does not last for long, may also be selfists. Galen Strawson, for instance, recently argued for the view that selves are real, but last only for a few seconds. He claimed that there are a series of such selves associated with what we would call individual persons, a claim that he called *the pearl view*. Strawson is a materialist, and may or may not be a selfist. But whatever his view about reductionism, one could subscribe to his pearl view, for pretty much the same reasons that he gave to subscribe to it, and be a selfist. For instance, one could subscribe to his pearl view and hold that selves are like Cartesian egos in being immaterial and indivisible.<sup>15</sup> Because a punctualist (or an episodic) believes that the self does not last for long, then a punctualist who is also a selfist has pretty much the same problems as a reductionist in justifying future-oriented egoistic concern. Future pearls on the string, whatever their metaphysical status, are still “others.”

But surely, it may seem, among selfists, punctualism is a minority view. So even if a selfist who is a punctualist would have a problem justifying future-oriented egoistic concern, what about selfists who are not punctualists? What, for instance, about selfists who hold that the self is real and spans the entire lifetime of the person whose self it is. In my view, even such a selfist has a problem, similar to that faced by a punctualist, in justifying future-oriented egoistic concern. His problem is to explain why *me-now*, that is, the current temporal stage of oneself, should be egoistically concerned about *me-later*, a future temporal stage of oneself.

What a selfist who is not a punctualist would no doubt reply to this problem is that *me-now* should care about *me-later* because both are parts of me (or, alternatively,



because both are me). I shall call this reply, *the me-consideration*. Such selfists claim that the me-consideration adequately justifies future-oriented egoistic concern. However, the me-consideration's being an adequate justification depends at least on one's being justified in believing in the existence of selves, or of "the further fact," and whether anyone is so justified is open to question. And even if the further fact were acknowledged to exist, one still might question whether the me-consideration is an adequate justification of future oriented egoistic concern.

Suppose, for instance, that the further fact—say, one's soul-pellet—persists, but one's psychology does not. Do future stages of the resulting person, even if that person is oneself, justify special concern? And even if both the further fact and one's psychology were to persist "together," it would not *follow* that me-now *should* care in the special concern way about me-later. One reason it would not follow is that *is* does not imply *ought*; hence, me-then's *being* a future stage of me-now does not *imply* that me-now *should care* in the special concern way about me-then. Another reason it would not follow is that me-now may not *identify psychologically* with me-then in a way that supports special concern; for instance, me-now may not anticipate *having* the experiences that will be had by me-then.

It may seem that it would be pathological for me-now not to anticipate having the experiences of me-then—for instance, for you not to anticipate having the experiences of yourself in the future. But even if under ordinary circumstances such a failure to anticipate would be pathological, what about in extraordinary circumstances? Consider, for instance, teletransportation: You enter a transmitting station on Earth. Your body and brain is scanned and simultaneously decomposed, as the information scanned is sent to a receiving station on Mars, where one and only one exact replica of what you were on Earth is produced. Is that Martian replica the same person as you were on Earth? Some psychological continuity theorists would argue that he is the same person. Suppose they are right—that is, that the best way to extend prevailing criteria of personal identity is to the answer that, yes, he is the same person. Even so, someone entering the transmitting station on Earth could sensibly ask why he should care about extensions of

the prevailing criteria of personal identity to cover exotic cases, and hence why he should care about the fate of his replica on Mars. The answer, because on these criteria the Martian replica will be you, does not answer this question.

Something like this worry is ultimately what is the matter with John Perry's suggestion that going out of existence and being continued by a physical and psychological replica of oneself would be as good as being continued by oneself because what matters so far as one's continued existence is concerned is merely continuing one's projects and one's replicas could do that as well as oneself.<sup>16</sup> It is also what is the matter with Parfit's suggestion that a person stuck out on a "branch-line" on Earth ought not to be too concerned about his own impending death in a few days since his exact replica of a few days earlier is living safely on Mars. A sticking point with both views is that most of us would care egoistically—and, it would seem, rationally—about more than just there being someone in the future whose body and psychology is qualitatively similar to our own. And we would care more even if by some rational extension of prevailing criteria of personal identity that person in the future were to be us.

What more do we care about? For one thing, we also care about their being someone in the future whose experiences we anticipate having. And while some of us would be capable of anticipating having the experiences of our replicas, some of us would not, even if we thought correctly that we were rationally entitled to do so. Similar issues apply to fission scenarios, which is why some would prefer taking a conventional space ship to Mars to being teleported. It is also why many find so bizarre Parfit's claim that in his branch line case the human who stays on Earth should not regret his own imminent death for the reason that his replica on Mars will continue to live.<sup>17</sup>

What, then, of the claim that one's failure to rationally anticipate having the experiences and performing the actions of a replica who is justifiably regarded, on the basis of a rational extension of criteria of personal identity, as himself in the future would be pathological? In the case of exotic examples at least, it is hard to see how one could defend this claim without begging the question. If by every normal human standard, one

was not dysfunctional in any way, then one's failure to anticipate having the experiences of a Martian replica might just be a feature of the way in which some psychologically healthy people anticipate the future. In exotic cases, what would be rational and what would be psychologically healthy need not go hand in hand.

But just as one might in such exotic cases fail to go along with what by conventional circumstances would be rational and yet one might still be psychologically healthy, so also one might do so in more normal circumstances. Parfit's young Russian nobleman example is a case in point.<sup>18</sup> In addition, one could argue, as many Buddhists do, that our normal pattern of egoistic expectations is not healthy. There are worms underneath the lid of the psychopathology can. It may be that without having sullied the rationally coercive power of the me-consideration one can get the lid back on again, but it seems doubtful.

In short, the me-consideration by itself is not an adequate justification of future oriented egoistic concern because one can sometimes sensibly ask why me-now should have special concern for me-later. The reply, because me-later is me, is not always an adequate answer to this question. Nor is the reply, because me-now and me-later are both parts of me, always an adequate answer. The fact remains that even if me-now is me-later, one can still ask sensibly for a further justification of future oriented egoistic concern. This is especially apparent, for obvious reasons, if the question of which view, if any, of same person over time is most plausible is only a question of utility. But if the me-consideration is not a fully adequate justification of future-oriented egoistic concern, regardless of whether one is a reductionist or a selfist, and for pretty much the same reasons whether one is a reductionist or a selfist, then the problem of justifying egoistic concern is not brought about by reductionism, but only made more visible by it. The root cause of the problem of justifying egoistic concern would then be something else, such as the ubiquity of a certain sort of analytic perspective from which one can ask sensibly on behalf of one's present person stage why one should care about one's future stages.

It does not follow from what I've said that future-oriented concern cannot be justified. In my view, it can be justified. My point is only that selfists and reductionists

have different, but parallel, resources for answering the challenge posed by the request for a justification of future-oriented egoistic concern. Typically selfists answer it by appeal to what I have called the *me-consideration* and by assuming the rationality of self-interest. Reductionists may answer it by appeal to what might be called the *continuer-consideration* and by assuming the rationality of continuer-interest.<sup>19</sup> So far, it would seem, neither response has any advantage over the other. Selfists may seem to have an advantage in that the relevance of the me-consideration and the rationality of self-interest are widely acknowledged, whereas the relevance of the continuer-consideration and the rationality of continuer-interest are not. But that would be a weak reed on which to rest the justification of one's view.

Selfists may also seem to have an advantage in being able to make a temporally neutral appeal to self-interest: that is, in calculating one's self-interest, every stage of oneself counts the same. But in the contest between self-interest and continuer-interest, it is not clear why temporal neutrality should be an advantage, and in any case reductionists can make their own kind of temporally neutral appeal: they can say that all continuer-continuities that are to the same degree count the same, regardless of when they occur.

Reductionists, on the other hand, may have two advantages of their own: first, they do not have to suppose that anything exists (except as a linguistic convention) that there is no reason to suppose exists; and, second, at least for those who take a three-dimensional view of persons, certain hypothetical examples, especially fission examples, seem to support the view that personal identity is not primarily what matters in survival and, hence, the view that egoistic-concern, rather than being basic, is actually derivative.<sup>20</sup> If either of these two reasons is acceptable, then the reductionist's advantage is the more serious advantage.

Some philosophers, in addressing this sort of reductionist response to the challenge posed by the Extreme Claim seem to assume that continuer-interest must be based on relation R in the way that Parfit understood it, that is, as psychological connectedness and continuity with any cause. On this view, the proponent of continuer-

interest as the justification of future oriented egoistic concern would be saddled with unintuitive consequences, such as emerge from Parfit's branch-line case and to a lesser degree from teletransportation examples. So, it's important to remember that so far as reductionism is concerned, continuer-interest need not be understood as as psychological connectedness and continuity with any cause. It is open to a reductionist to understand continuer-interest in terms of bodily continuity or in terms of psychological connectedness and continuity with its normal cause, both of which would give the reductionist different resources on which to draw in defending the rationality of future oriented egoistic concern. The important point is that the reductionist who takes himself to see at least intellectually through the illusion of self, and therefore on that grounds to challenge the rationality of future-oriented *self*-interest, does not thereby challenge the rationality of any sort of future-oriented interest. In particular, he does not thereby challenge the rationality of future-oriented continuer-interest. On the face of it, it's rationally permissible for a reductionist to be continuer-interested about himself—his continuers—in the future.

## II

### **Ironic engagement**

According to Buddhism, commitment to the reality of the self in the way in which almost everyone is committed to the reality of the self is both rationally unjustified (since it's not ultimately true that selves exist) and a source of suffering. The remedy is to remove that sort of commitment to the reality of the self. Of course, one does not need to travel though Buddhism in order to get to this conclusion. Many contemporary philosophers seem to think that commitment to the reality of the self in the way in which almost everyone is committed to the reality of the self is not rationally justified. And, according to Parfit commitment to the reality of the self in the way in which almost everyone is committed to it is in his own case at least also a source of suffering. The suffering, he says, is caused by alienation—by a feeling a separation between oneself and others—and by a heightened fear of death.<sup>21</sup>

If one reaches—via any route—the conclusion that commitment to the reality of the

self in the way in which almost everyone is committed to it is both rationally unjustified and a source of suffering and then tries to remove that sort of commitment to the reality of the self, what, if any, part of that former commitment should one try to retain? In Parfit's view, perhaps none. In much of what he says, he seems to be an eliminativist with respect to the reality of the self in that he doesn't recommend that one leave any more of that former commitment in place than one has to leave in place. Some others recommend that we should replace our current commitment to the reality of the self with a similar commitment, but with this difference: we should be not fully engaged, but only *ironically engaged* with society's personhood conventions.

What does it mean to be only ironically engaged with our personhood conventions? Part of what it means, it seems, is that to whatever extent we continue to adopt our usual personhood conventions, we should do so in full awareness of the fact that they are only at best useful fictions, not true descriptions of the ways things really are. If one is a reductionist, it is hard to argue with this much of how we should regard our usual personhood conventions. But should we be more engaged with them than that? It seems to me that regardless of how we answer this question, there is an atavistic, but persistent, belief in the self that is resistant to intellectual arguments that expose it as a mistake. Since this belief is so resistant to intellectual dissolution, there is not much that one should do—since there is not much that one can do—to remove it. However, there is an additional concession that one should make to becoming entangled with our usual personhood conventions. It is the recognition that for practical purposes, such as straightening out with the airlines a confusion about when you originally ordered a ticket, it may be convenient *to talk as if you believed* in the reality of the self. Engaging in this sort of talk does not commit you in any way to actually believing in the reality of the self, except perhaps to believing in it as a convenient fiction.

Is there anything more to being optimally ironically engaged with our normal personhood conventions? Not, it seems, if one is an eliminativist about the self. But some philosophers claim that there is something more to being optimally ironically

engaged with our normal personhood conventions. Paul Williams, for instance, argues that for ethical reasons one has to acknowledge the conventional existence of “subjects” and of “individual persons.” One needs subjects because without them pains would be “free-floating” and it is “incoherent to treat pains as if they are free-floating.” One needs individual persons because we have to recognize human individuality in order to help one another. In Williams’ words:

Not only is it incoherent to treat pains as if they are free-floating, but—as anyone who has ever received training in counselling knows—to help others effectively requires *not* that we discount their individuality as the persons they are but actually [sic] to focus on that individuality most closely. The good counsellor—dare I say, the good bodhisattva—is someone who can actually discount to an unusually effective extent their own intervening concerns in order to focus on the other *in their uniqueness*. This requires a very vivid awareness of the other as an individual. It is not helped by denying uniqueness to either of us. The pain which we seek to remove is intrinsically embedded in the actual individual in front of us, who is different from other individuals and, of course, different from us. However the bodhisattva is going to develop the most effective way to work for the benefit of others, in order to be an effective helper he or she is going to have to recognize and start from the individual person, the fact that each person is an individual with unique circumstances, problems and potential. Anyone who actually works in the caring professions knows that.<sup>22</sup>

But Williams conclusion does not follow from his premises. One can have subjects and individuality without having selves or persons. Individual human bodies, including their mental states, can function as the unique owners of pain experiences. And, without acknowledging the existence of selves or persons, one can track the careers of individual human bodies. It may be that to avoid untoward consequences one has to acknowledge at least the conventional existence of something whose conventional existence some Buddhists wish to deny. But nothing that Williams says implies that to avoid untoward consequences one has to acknowledge the existence of selves or

persons.

To take another example, Mark Siderits recommends that rather than being an eliminativist about the self, one should concede that it is conventionally true that there are selves. Presumably, this concession, as he understands it, involves something more than simply making use in certain practical contexts, such as in dealing with the airlines, of the knowledge that in our language culture, almost everyone, almost all the time, not only believes in the reality of the self, but experiences the world as if they believe in the reality of the self. An eliminativist could cheerfully make this concession.

Siderits, on the other hand, says, “According to the Reductionist, the personhood convention prevails because it is more conducive to overall welfare than the readily available alternatives, such as punctualism and the *Weltgeist* convention. . . . [and] utility would be better served if there were some way to combine the virtues of the personhood convention (such as the avoidance of gross imprudence, and the gains in welfare achieved through individual initiative) with a strategy for avoiding existential suffering.”<sup>23</sup> The strategy that he recommends for this purpose is to adopt a certain attitude toward the conventional truth of the reality of the self that enables us to retain some sort of commitment to causal series that have the capacities for self-revision, self-control, and self-scrutiny. In his view, this commitment involves more than what is available to the eliminativist. But it is not clear that it does involve more. An eliminativist about selves, or persons, can cheerfully admit the existence of causal series that have the capacities for self-revision, self-control, and self-scrutiny. What more might be required?

According to Siderits, more that’s required is something that would allow one to answer what he calls “the alienation objection,” that is, the objection that having a life is not the sort of thing one can choose as a means to further some separate end. It would, for instance, be most peculiar for someone to claim as their reason for bestowing love and affection on their spouse and children that this is the best way open to them to contribute to overall welfare. To claim this would seem to show a singular lack of understanding of just what love and



affection are, and a person who said this might properly be described as alienated from their feelings of love and affection.<sup>24</sup>

A little later Siderits adds, “If Reductionism is true, it may also be true that welfare is maximized by our feeling genuinely personal regard for others, and our viewing ourselves as the authors of our own life-narratives. But the belief that Reductionism is true seems to irreparably alienate us from all such person-involving attitudes.”<sup>25</sup>

But is it reductionism per se that raises the specter of alienation? Nothing that Siderits says shows that it is. One can feel regard for “others” without supposing that the others for whom one feels regard are selves or persons. Hence, one can feel regard for others without acknowledging even the conventional existence of selves or persons in any sense that would be unavailable to an eliminativist about selves and persons. I would have thought, ala Parfit, that it is experiencing the world, including oneself, as if one believes in the reality of the self that leads to alienation. It is true that if one adopts a consistent Parfitian eliminativism, then one has to recast some of the ways in which one feels personal regard for others—say, recast personal-regard as continuer-regard. But what reason, in principle, is there to suppose that this will be a problem?<sup>26</sup>

If we become Reductionists about persons, we do have to scale back our beliefs. But suppose we scale them back. Should we then ironically engage with our former beliefs, in anything other than the practical way illustrated by my airlines example? I don’t see why any sort of fuller engagement is necessary. When we give up our belief in the reality of the self, we don’t give up our belief in the existence of a brain, a body, and a series of interrelated physical and psychological events. Nor do we abandon continuer-interest. What we used to think of as our future selves are still, albeit perhaps less robustly, our future continuers, and we may value them as our continuers. Our persistence is less than we thought before, but it is not nothing. We don’t value our future selves based on the me-consideration, but we value them—our continuers—based on continuer-interest, and so far as reductionism is concerned, we can understand continuer-interest more robustly than Parfit’s psychological connectedness and continuity with any cause.

Even so, to give up the me-consideration and value our continuers based merely on continuer-interest, some adjustments will be required. We will have to scale back. But in scaling back, I can't see that we're in any danger of plunging into an abyss. In fact, I can't see that we're not in any danger at all. It seems to me that the danger—primarily of alienation—is all on the other side of the equation. There is still a problem—that of getting ourselves to believe *at all levels of our own psychologies* in the reductionism to which we are intellectually committed. But meditation, not ironic engagement, seems to be the solution to that problem. In sum, to a reductionist, belief in the existence of the self is itself a kind of pretense. To whatever extent one can shed this belief, there is no need to keep on pretending.

Withdrawing one's commitment to the notion of self, and with it to the normative force of self-interest, and replacing these with the notions of continuers and continuer-interest involves a sort of scaling back similar to what occurs in other domains in which we also come to philosophy with naive commitments. In the case of free will, for instance, many people initially come to philosophy naively committed to a sort of libertarianism and leave as soft-determinists. They still believe in free will, but the free will they believe in is less robust. So too, for instance, in the case of the self and related commitments, such as to the rationality of prudence. What we are left with is not, as the Extreme Claim would have it, almost nothing, but with continuer-beliefs and commitments that are less robust, in a certain way, than the self-beliefs we brought to the table. Less robust, but still adequate for every practical purpose worth pursuing. Even so, most of us will not be able to shake an atavistic belief in the reality of the self. Nor will we want to avoid the practical advantages of employing self-talk. Both of these may require the sort of thin ironic engagement that is available to an eliminativist. But that seems to be all that's required.

There remains a final question. If there being no self is not a big deal, why have so many people, including so many philosophers, thought that it is a big deal? That, I think, is a genuinely interesting question. I would not have time to fully answer this question, even if I knew the answer, which I do not. However, it is obvious that an

important part of the answer is that ordinary people, as well as many philosophers, have supposed that if the self goes, other things of genuine value go it. As we have seen, some of these things that have been thought to be lost if we were to relinquish the self are not lost. But there may still be others that are lost. One of these, which is ingredient in Butler's criticism of Locke, has not gotten nearly as much attention as it deserves, has to do with ownership. In particular, it has to do with the possibility that if there were no selves or persons, then no one would be entitled to anticipate *having* experiences that will be had by "oneself" in the future. If this were a genuine implication of relinquishing belief in the self and if coming to terms with it were required by relinquishing belief in the self, most of us, I think, including myself, would find relinquishing belief in the self to be profoundly disturbing. And if, as I also think, coming to terms with this idea is required by genuinely relinquishing belief in the self, then it might be an important part what's such a big deal about ceasing to believe in the self.<sup>27</sup>

## ENDNOTES

---

1. Both Buddhists and many Western philosophers espouse this view. For instance, in a recent book, Thomas Metzinger announced that his "main thesis is that no such things as selves exist in the world: Nobody ever *was* or *had* a self. All that ever existed were conscious self-models that could not be recognized *as* models. The phenomenal self is not a thing, but a process—and the subjective experience of *being someone* emerges if a conscious information-processing system operates under a transparent self-model. You are such a system right now, as you read these sentences. Because you cannot recognize your self-model *as* a model, it is transparent: you look right through it. You don't see it. But you see *with* it. In other, more metaphorical, words, the central claim of this book is that as you read these lines you constantly *confuse* yourself with the content of the self-model currently activated by your brain." Thomas Metzinger, *Being No One: The Self-Model Theory of Subjectivity* (MIT Press, 2003), p. 1.

- 
2. This remarkable scene was known to Plato (*Theaetetus* 152e) and subsequently widely discussed in late antiquity as “the Growing Argument.” See David Sedley, “The Stoic Criterion of Identity,” *Phronesis*, 27 (1982) 255-75, p. 255.
  3. This paragraph and the previous two are from Raymond Martin and John Barresi, *The Rise and Fall of Soul and Self* (New York, Columbia University Press, 2006), p. 3
  4. *Ibid.*, pp. 331-2.
  5. *Ibid.*, pp. 332, 334.
  6. The only non-controversial examples of selves, or persons, that any of us has ever encountered are human organisms, which are, or at least certainly seem to be, entirely composed of sub-personal parts in various relations to each other. Selves, or persons, it would seem, if they exist at all, must supervene entirely on these parts. Thus, selves, or persons, have at most a derivative claim to exist: they exist solely because the things and relations out of which they are composed exist. If this is right, then however useful the notion of self, or person, in making practical sense of our lives, it would seem that a description of the underlying sub-personal parts and the relations among them, without mention of selves, or persons, would be complete ontologically in the sense that if one just mentioned them and omitted mention of selves, or persons, one would not have left anything out. Thus, selves, or persons, it would seem, either just are these sub-personal parts in relation to each other, or else they are part of a fictional way of talking about human life—about *our* lives.
  7. Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), p. 211. See also *Rise and Fall*, p. 211.
  8. For more on these notions and the relations between them, see J. Kim, *Mind in a Physical World* (Cambridge: Cambridge University Press, 1998) and Sydney Shoemaker, “Realization and Mental Causation,” in Carl Gillett and Barry Loewer, eds., *Physicalism and Its Discontents* (Cambridge: Cambridge University Press, 2001), pp. 74-98. See also Terry Horgan, “From Supervenience to Superdupervenience,” *Mind* (1993).
  9. Mark Johnston, “Human Concerns Without Superlative Selves,” pp. 149-79, and John

---

McDowell, "Reductionism and the First Person," pp. 230-50, in Jonathan Dancy, ed., *Reading Parfit* (Oxford: Blackwell, 1970).

10. *Reading Parfit*, pp. 154-56.

11. *Reasons and Persons*, p. 211.

12. *Ibid.*, p. 307.

13. Historically, as we have seen, the Extreme Claim surfaced in the eighteenth century in Bishop Butler's response to Locke's reductionistic account of personal identity. There were other selfists in the eighteenth century, both before and after Butler, who made a similar criticism of Locke. Subsequently, the Extreme Claim surfaced in Sidgwick's response to Hume. In our own times, it has been endorsed by Richard Swinburne, among others (see Parfit, *Reasons and Persons*, pp. 307-09). The reductionist whose work I have primarily in view is Derek Parfit. *Reasons and Persons* has been the point of departure for most recent discussions of the implications of reductionism for personal identity issues.

14. I am borrowing the expression, *ironic engagement*, from Mark Siderits, *Personal Identity and Buddhist Philosophy: Empty Persons* (Ashgate, 2003).

15. Many philosophers, following Plato, have supposed that if a self is immaterial and indivisible, then it must be naturally immortal, on the grounds that only something that could decompose could go out of existence and an indivisible substance could not decompose, but Kant questioned whether the only way for something to go out of existence is to decompose. Another possibility, Kant suggested, is that an incomposite thing might simply fade out (say, like a lightbulb that burns out). If Kant's suggestion is acceptable, then the same problem that Butler posed for Locke's reductionistic view might arise on a non-reductionist view. But, as I've suggested, even if Kant is wrong and immaterial beings are naturally immortal, the same problem might arise on a non-reductionist view due to our being able to distinguish different temporal stages of the same continuing entity and to ask why one's current temporal stage should have

---

egoistic concern for future temporal stages of oneself.

16. John Perry, "The Importance of Being Identical," pp. 67-90, in Amelie Rorty, ed., *The Identifies of Persons* (Berkeley: University of California Press, 1976). I have discussed Perry's view, in connection with the psychological issue of identification, in "Having the Experience: The Next Best Thing to Being There," in *Philosophical Studies* 70 (1993)63-79.

17. Parfit, *Reasons and Persons*, p.

18. *Ibid.*, p.

19. But, as Parfit pointed out in giving expression to the Moderate Claim,

20. So, for instance, it is possible to argue that in the case of many of us (whether or not we take ourselves to have seen through the illusion of self), our more fundamental value is continuer-interest and our valuing self-interest is derived from that. This might happen if in a situation of limited options, such as the situations in which humans almost always find themselves, continuer-interest and self-interest almost always coincide. In such circumstances, self-interested beings and continuer-interested beings would make the same choices. However, when it comes to certain hypothetical choice situations, they would choose differently, and if continuer-interest were the more fundamental value, then our hypothetical choosers should opt for the choices that express continuer-interest, rather than self-interest.

21. Parfit, *Reasons and Persons*, p.

22. Paul Williams, *Altruism and Reality: Studies in the Philosophy of the Bodhicaryāvatāra* (Richmond Surrey: Curzon, 1998), pp. 174-75.

23. Siderits, *Personal Identity and Buddhist Philosophy*, pp. 99-100.

24. *Ibid.*, pp. 104-5.

25. *Ibid.*, p. 105.

26. Siderits best explanation of his richer notion of *ironic engagement* is via an analogy. "Suppose I take pride in the city in which I was born and now live. I am, however, a

---

reductionist about cities (an “urbanist” for short): I know that the existence of a city just consists in the existence of certain buildings and infrastructure in a certain location, and certain people interacting in certain ways. I know that “city” is a mere convenient designator for these more particular entities when related to one another in certain characteristic ways. I know that cities are only conventionally and not ultimately real. Does this knowledge undermine my civic pride? . . . The suggestion is that it must have a terminally corrosive effect, since one cannot take pride in something one believes to be ultimately unreal. But this need not be true. I am, after all, a reductionist about cities, not an eliminativist. That is, I believe our use of the convenient designator “city” reflects the genuine utility achieved when these more particular elements are related to one another in these characteristically urban ways. Moreover, I may believe that this utility is greatly enhanced when the inhabitants of urban aggregates engage in various kinds of cooperative behavior, and that such behavior is more likely to occur if they feel a sense of attachment to their location. So I may conclude that it is better, all things considered, that city dwellers feel pride in their city; and since I am a city dweller, I should feel pride in my city. . . .

When, for instance, I enthusiastically describe the charms of the place to a visitor, my aim is not to enhance the experiences of all who dwell here. My aim is just to express my pride in the city. True, my pride came about because of my desire that the experiences of the inhabitants be improved. But that hardly makes that desire the motive behind all acts expressive of my pride. To think so is to commit the genetic fallacy. Indeed I can be perfectly clear how it is that I came to feel civic pride, yet still have the genuine article. The knowledge will induce a degree of ironic distance—enough to ward off the dangers of civic chauvinism. So as I wax poetic in singing the city’s praises to the visitor, I shall also comment wryly on the somewhat hyperbolic character of my account. Still I do wax poetic; I want to share my love of the city with others. I am ironically engaged.” (*Personal Identity and Buddhist Philosophy*, pp. 107-8). But if I understand Siderits correctly, a problem with his analogy is that he seems to be recommending a kind of pretense, and it doesn’t seem to me that in order to have civic

---

pride, one needs to go beyond reductionism about cities and engage in pretense. In other words, I don't see how urbanism interferes with civic pride. In the case of civic pride, it seems to me, an urbanist can be fully, not just ironically, engaged. Of course, one cannot believe in cities as something over and above "the existence of certain buildings and infrastructure in a certain location, and certain people interacting in certain ways," but I at least don't feel any pull to believe in cities in any such quasi-Hegelian way. In short, in the case of civic pride, there doesn't seem to be any problem to which appeal to ironic engagement might be the solution.

27. Earlier versions of this paper were read at Cambridge University and Colgate University. Thanks to people at one or the other of these events, especially Mark Siderits and Ulrich Meyer, and also to my colleague, Brad Rives, for helpful comments.