

Would It Matter All That Much if There Were No Selves?

Raymond Martin

Pointing at the Moon: *Buddhism, Logic, Analytic Philosophy* – © 2009 Oxford University Press, Inc. (pp. 115 – 134)

9

Would It Matter All That Much if There Were No Selves?

Raymond Martin

Reductionism about a kind of thing is the view that things of that kind just consist in things of a more basic kind. Buddhism and Western philosophy have each appealed to reductionism to argue that, strictly speaking, there is no such thing as a self, or person. Mark Siderits put this point nicely:

The Buddhist view of non-self... says that the existence of a person just consists in the occurrence of a complex series of impermanent, impersonal *skandhas*. But Buddhists are not the only ones to hold a reductionist view of person. On some interpretations, both Locke and Hume held such a view. More recently Derek Parfit has given a sophisticated defense of reductionism about persons, which he explains as the denial that the continued existence of a person involves any “further fact” over and above the facts about a causal series of psychophysical elements.¹

Parfit’s view, in a little more detail, is that the existence of a self, or person (in what follows, I shall use *self* and *person* interchangeably), just consists in the existence of a brain and body and the occurrence of a series of interrelated physical and mental events.² There is more to Parfit’s view. But, in what follows, just this portion is what I shall mean by *reductionism about the self*.

How important would it be *practically* if reductionism about the self were both true and believed to be true? Buddhist and Western

philosophers tend to agree that it would be a big deal, but for different reasons. Buddhists have held, and still do hold, first, that belief in the self—at least in the way in which almost everyone believes in it—is a major source of suffering, and, second, that the realization that there is no self is life changing in an extremely positive way: “Buddhists say . . . that becoming enlightened, coming to know the truth of reductionism, relieves existential suffering. They also say that it makes one more concerned about the welfare of others.”³ In the West, both proponents and critics of the idea that there is a substantial and enduring self have focused on what many have taken to be more pessimistic implications, especially the worry that, if there were no self or even if there were no substantial and enduring self, then there would be no reason for future-oriented self-concern, including prudence. Bishop Butler, for instance, who believed that the self is a simple immaterial substance, tried to counter John Locke’s relational theory of personal identity with the criticism that, if selves were to consist only in parts that do not endure, then it would be a mistake “to imagine our present selves interested in anything which befell us yesterday [or] will befall us tomorrow.” Under such circumstances, he said, “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.”⁴ Many Western philosophers have echoed Butler’s worry.⁵

My main goal, in the present chapter, is to put this worry about future-oriented self-concern to rest. To do this, I shall argue that whatever difficulties there may be in justifying future-oriented self-concern, they are no greater for skeptics about the self than they are for believers. A secondary goal will be to show that some recent Buddhist commentators have exaggerated the extent to which those who deny the existence of a substantial and enduring self need, for practical purposes, to pretend that one actually exists. I shall argue that not much pretense is required.

East Meets West

It is surprising how similar are the concerns of ancient Buddhist philosophers about reduction of the self to those that preoccupy contemporary analytic philosophers. In Buddhist philosophy, concern over whether the self or person is real and, if real, real in what sense, was present from the beginning. The Pudgalavādins, who appeared within a few centuries of the death of the Buddha and included several of the early schools of Buddhism, maintained that persons (*pudgala*) are both distinct from the five aggregates (material form, feeling, ideation, mental forces, and consciousness) and real. Other

Buddhist philosophers who were reductionists, such as Vasubandhu (flourished c. 360 c.e.), argued in opposition to the Pudgalavādins that persons or selves, while real, are nothing but the aggregates.

In the idiom of contemporary analytic personal identity theory, the Pudgalavādins, while not arguing for anything like an immaterial substance, were nevertheless arguing that the self is what Derek Parfit would call “a further fact,” while Vasubandhu, like Parfit, was arguing that the self, while real in a sense, is not a further fact. Vasubandhu seems also to have subscribed to two other views for which Parfit is famous: that it is possible to describe reality completely in impersonal terms (the impersonal description thesis); and that it is an *empty question* whether something that is agreed to be constituted out of other, more basic things “really exists.”⁶

Mādhyamika Buddhist philosophers, such as Nāgārjuna (c. 150–250 c.e.), Āryadeva (c. 180–250 c.e.), and Candrakīrti (c. 600–650 c.e.), denied both that selves are reducible to the aggregates and that they are distinct from the aggregates. They thus rejected realism about the self altogether. By claiming that what is ordinarily taken to be belief in or reference to the self is actually an act of appropriating (*upādāna*) one’s experiences, emotions, and body, they focused instead on the self’s seeming ownership of these things. According to Candrakīrti, for instance, the proper explanation of our sense that we “own” our experiences, emotions, and bodies is that our everyday conception of self consists in “an appropriative act of laying claim to the elements of the psychophysical aggregates, an act that does not require there to be any ‘entity’ or ‘object’ that is the self.”⁷

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 Greek comic playwright Epicharmus. In this scene, a lender asks a debtor to pay up, and 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. When the debtor protests loudly at being thus abused, the lender replies that his complaint is misdirected since he—the lender—is not the same person as the one who struck him a moment before.⁸

In spite of such sophistication about the self, there is not much evidence that Epicharmus or other Greek or Roman philosophers seriously entertained

the idea that selves or persons do *not* really exist—that is, that they are fictions. In the West, that idea came onto center stage in the late seventeenth century, via Locke’s famous remark that *person* is a forensic term, which Locke made immediately after giving a relational account of personal identity over time. These two proposals—that the self is a fiction and that, whether a fiction or not, the self over time should be understood relationally—tended to be lumped together in the minds of many of Locke’s eighteenth-century critics. Butler, for instance, thought that it followed from Locke’s relational view that each of us would be a persisting self only in a fictitious sense. He thought that this consequence refuted Locke’s view, but not that it *proved* it wrong. Rather, he thought, it enabled people to *rationality intuit* that it is wrong: “the bare unfolding of this notion [that selves are fictitious] and laying it thus naked and open, seems the best confutation.”⁹

In Mādhyamika Buddhism, the radical suggestion that the self is a fiction was 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 Mādhyamika Buddhism.¹⁰ However, in the view of Mādhyamikas, 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 reductionists about the self—indeed, few Western philosophers of any sort—are willing to go that far. Instead, 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 say that selves or persons are fictional or, if they do say this, are reluctant to say that selves are fictional *merely because* they strongly supervene on subpersonal parts and relations. In sum, so far as the West is concerned, whereas the original critics of Locke’s radical suggestion that selves or persons are fictions tried to defeat his view in order to save the traditional idea that the self is an immaterial soul, in the twenty-first century critics of the view that selves or persons are fictions tend to be nonreductive materialists. As a consequence, in the West, the contemporary philosophical battle is no longer, as it once was, between religion and science, but over how best to understand notions such as *supervenience* and *realization*.¹¹

It may seem that the Mādhyamika Buddhist idea that every (composite) thing 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 subpersonal parts that are real. But, in one respect, the Buddhist view may be less radical. On most contemporary Western reductionistic views, there is

something metaphysically special about the self that distinguishes it from the subpersonal parts to which it may be reduced. The self seems to be a thing, but is not a thing; the subpersonal parts are things. On versions of Buddhism according to which *everything* goes into the same hopper, there is nothing metaphysically special about the self. Yet, most Buddhists have wanted to say that there is something special about the self—perhaps not metaphysically special, but special in its being an illusion, or in the way that it is an illusion, or in the role that it plays as an illusion in our relationships to ourselves and the world.

Contemporary Western analytic philosophers—as we shall see, like some of their contemporary Buddhist counterparts—are divided about whether and, if so, to what degree and in what way selves or persons are expendable. Some, including Mark Johnston and John McDowell, argue that selves are not expendable since one needs them in order to make sense either of human values or of epistemology, or both.¹² However, 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, Johnston, in an effort to shore up the reality of the self while simultaneously distancing himself from Cartesians, has called the selves to which he thinks one should be committed *ordinary further facts*, as opposed to *superlative further facts*.¹³ Johnston’s view, interestingly, seems to be virtually identical to that of the Pudgalavādins.

Reductionism and the Extreme Claim

Reductionism about the self is a metaphysical view that does not commit one to any particular normative theory or to any 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, nor that what does matter, à la Parfit, is psychological connectedness and continuity. These latter claims are in addition to reductionism about the self. Even so, in the West today, reductionism about the self, particularly with respect to its implications for the rationality of future-oriented egoistic concern, is a controversial thesis. An important challenge to it is a contemporary version of Butler’s claim that, if reductionism about the self were true, then people would have no reason to be especially concerned about their own futures. Parfit calls this challenge the *extreme claim* and contrasts it 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, does allow one reason for special concern.¹⁴

The intuition on which the extreme claim seems to rely, and what has made it seem plausible to many critics of reductionism about the self, 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 is said, the current stages of the series would have no *egoistic* reason to be especially concerned about subsequent stages. In a partial concession to this view, Parfit argues that one has moral, rather than self-interested, reasons to be concerned about future stages of oneself.

It is incumbent upon those who claim that reductionism about the self would not be a big deal practically to explain what's wrong with the extreme claim. In my view, the main thing wrong with it is that reductionists about the self can appeal to continuer-interest, instead of self-interest or morality, to justify the rationality of future-oriented surrogates of egoistic concern, including prudence. This is the main thing for which I will argue in the remainder of this section. If this is right, then some of the reasons that have been given to support the view that reductionism about the self motivates revisionist views about what one should or should not value, or about how one should or should not behave—and hence that reductionism would be a big deal practically—do not succeed. Of course, reductionism about the self still might be a big deal practically, for other reasons. I don't think it is, but I shall not argue for that in the present chapter. Instead, in the final section, I shall conclude by explaining why the only practical concession that reductionists about the self need to make to personhood conventions may be to adopt an attitude that I shall call *thin ironic engagement*.¹⁵

The extreme claim is supposed to be a problem for reductionists about the self that does not arise for nonreductionists about the self (henceforth, *selfists*). My first objective is to show that the same problem arises for at least some selfists, in pretty much the same form that 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 who persists as the same person we are now does so 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. Less extravagant versions include the further facts postulated in the views of the Pudgalavādins and of Mark Johnston. In any case, on a selfist view, since future stages of oneself are clearly not others but oneself, it has seemed to many that there is no issue either about what justifies future-oriented egoistic concern or about our entitlement to own, or to anticipate having, the experiences of ourselves in the future.

But this seeming advantage of selfism does not accrue to all versions of selfism. Metaphysical punctualists (or episodics), who believe that the self is real but that it does not last for long, may also be selfists. Galen Strawson, for instance, has argued for the view that selves are real but last only for a few seconds. He claims that there is a series of such selves associated with what we would call an individual person, a claim that he calls the *pearl view*.¹⁶ 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. Since a punctualist (or an episodic) believes that the self does not last for long, 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 himself, should be egoistically concerned about *me-later*, a future temporal stage of himself.

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 in the existence of the further fact, and whether anyone is so justified is open to question. But even if the further fact is acknowledged to exist, one still might question whether the *me-consideration* is an adequate justification of future-oriented egoistic concern.

Suppose, for instance, that one's further fact persists, but one's psychology does not persist. Would the fact that there will be future stages of such a person, even if that person continues to be oneself, justify special concern? It is not obvious that it would. 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-later's *being* a future stage of me-now does not *imply* that me-now *should* care in the special-concern way about me-later. Another reason it would not follow is that me-now may not *identify psychologically* with me-later in a way that supports special concern; for instance, me-now may not anticipate having the experiences that will be had by me-later.

It might be objected that it would be pathological for me-now not to anticipate having the experiences of me-later—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, it is question begging to assume that it would necessarily be pathological in extraordinary circumstances. Consider, for instance, teleportation. A man enters a transmitting station on Earth. His body and brain are 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 he was on Earth is produced. Is that Martian replica the same person as he was on Earth? Some psychological continuity theorists would argue that he is the same person. Suppose they are right—that is, right that the best way to extend prevailing criteria of personal identity is to 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, that because on these extended criteria the Martian replica will be himself, 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. Perry says it would be as good 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.¹⁷ Something similar is also what is the matter with Parfit's suggestion that a person 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 will be living safely on Mars. A sticking point with both views is that most of us would care egoistically—and, it would seem, would care rationally—about more than just there being someone in the future whose body and psychology are qualitatively similar to our own; and one would care more even if one were to learn that, by some rational extension of prevailing criteria of personal identity, that person in the future is oneself.

What more might one care about? For one thing, one might also care about there being someone in the future whose experiences one can 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 anticipate having them.

What, then, of the reply that one's failure to anticipate having the experiences and performing the actions of a replica who—on the basis of a rational extension of prevailing criteria of personal identity—is justifiably regarded as oneself in the future would be pathological? In the case of exotic examples, such as teleportation, it is hard to see how one could defend this claim without begging the question. If, by every normal human standard, one is 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 normal circumstances, what would be rational and what would be psychologically healthy may go hand in hand, but in exotic circumstances they may not.

But, just as one might in exotic circumstances fail to go along with what in conventional circumstances would be rational and still be psychologically healthy, so also one might even in normal circumstances fail to go along and still be psychologically healthy. Parfit's young Russian nobleman example, in which a person tries to identify with a future stage of himself whose values he finds abhorrent, is a case in point.¹⁸ In addition, one could argue, as many Buddhists have argued, that our normal pattern of egoistic expectations is not healthy. In the context of discussion of the philosophy of the self, the import of these reflections is that appealing to the psychopathology accusation as a way of defending the rationally coercive power of the me-consideration is problematic. Perhaps one could solve all of the problems mentioned without sullying the rationally coercive power of the me-consideration, but this seems doubtful.

In sum, the me-consideration by itself is not an adequate justification of future-oriented egoistic concern because one can sensibly ask why me-now should have special concern for me-later. The reply "because me-later is me" is not an adequate answer to this question. Nor is the reply "because me-now and me-later are both parts of me." In both cases, one can still ask sensibly for a further justification of future-oriented egoistic concern. This is especially apparent if the issue of which view, if any, of the same person over time is most plausible hinges on considerations 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 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.¹⁹ 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, to hold that, 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 temporal neutrality 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, except as a linguistic convention, anything exists 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, that egoistic concern, rather than being basic, is actually derivative.²⁰ If either of these two reasons is accepted, then the reductionist has 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 those that emerge from Parfit's branch-line case and, to a lesser degree, from teleportation examples. So, one needs to remember that, so far as reductionism is concerned, continuer-interest need not be understood in

terms of 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 herself to see through—at least, intellectually—the illusion of self and therefore on these grounds might be thought plausibly to challenge the rationality of future-oriented *self*-interest does not thereby challenge the rationality of any sort of future-oriented interest. In particular, she 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 herself—that is, about her continuers—in the future.

Ironic Engagement

According to Mādhyamika Buddhists, commitment to the reality of the self in the way in which almost everyone is committed to it 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 through Buddhism to get to this conclusion. Many contemporary non-Buddhist Western philosophers seem to think that commitment to the reality of the self in the way in which almost everyone is committed to it is not rationally justified. According to Parfit, for instance, that sort of commitment to the reality of the self is not only unjustified but, in his own case at least, is a source of suffering. His suffering, he says, is caused by his feeling alienated from others and by his fear of death.²¹

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, 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 of sorts about the self in the sense that he doesn't recommend that one leave any more in place of one's former commitment to the self than one has to leave in place. In contrast to Parfit, some contemporary Buddhist commentators recommend that we replace our former (or current) commitment to the self with a similar commitment, but with this difference: we should not be fully engaged, but only *ironically engaged*, with society's personhood conventions.

What does it mean to be ironically engaged with society's personhood conventions? Part of what it means, it seems, is that, to whatever extent we continue to adopt society's personhood conventions, we do so in full awareness of the fact that they are at best only useful fictions, and hence are not true or at least not deeply true descriptions of the way 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. And there is an additional concession that one should make to becoming entangled in society's 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 believe* 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 contemporary Buddhist 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."²² One needs individual persons because we have to recognize human individuality in order to help one another. In Williams's 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 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.²³

But Williams's 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 might 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 dealing with the airlines, of the knowledge that, in our language culture, almost everyone, almost all of 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 that, according to the reductionist:

[T]he 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.²⁴

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, the more that's required is something that would allow one to answer what he calls the "alienation objection":

[This 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.²⁵

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."²⁶

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, à la 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 self-regard as continuer-regard. But what reason, in principle, is there to suppose that this will be a problem?²⁷

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, as reductionists, 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 we still regard, albeit perhaps less robustly, as our future continuers and we may value them as such. We don't value our future selves based on the me-consideration, but we value

them—our continuers—based on continuer-interest, and we are free to understand continuer-interest more robustly than Parfit's psychological connectedness and continuity with any cause. Our persistence may be less than we thought, but it is not nothing.

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 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 the problem of getting ourselves to believe *at all levels of our own psychologies* in the reductionism to which we are committed intellectually. 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 to the normative force of self-interest, and replacing these with the notions of a continuer 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 their exposure to philosophy as soft determinists. They still believe in free will, but the free will in which they believe is less robust. So too in the case of the self and related commitments, such as a commitment 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 continuer commitments that are less robust, in a certain way, than the self-beliefs and self-commitments 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 practically, why have so many people, including so many philosophers, thought that it is a big deal? That, I think, is an interesting question. I would not have time to fully answer it here, even if I knew the answer, which I do not. However, it seems that part of the answer is that ordinary people, as well as many philosophers, have supposed that if the self, or belief in the self, goes, other things of genuine value go with it. As we have seen, some of these things that have been thought to be lost if we relinquish belief in the self are not lost. But there may

still be others that are lost. One of these, which has to do with ownership and is an ingredient in Butler's criticism of Locke, has not gotten nearly as much attention in the West as it deserves. It is the possibility that, if there were no selves or persons, then no one would *own*, or be entitled to anticipate *having*, experiences that will be had by "themselves" 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, myself included, would find relinquishing belief in the self to be profoundly challenging.

As we have seen in Candrakīrti's Mādhyamika theory, our everyday conception of self consists in an appropriative act of laying claim to the sorts of elements in the psychophysical aggregate to which Western philosophers, such as Parfit, reduce the self. This may give Candrakīrti an advantage over Western reductionists in explaining the rationality of prudence, which would include such things as the significance of the distinction between the *anticipation* of one's own future pain and the *concern* one feels for the future pain of another. For in Candrakīrti's view, the function of self-talk is not to talk about objects in the world—about selves—or even about the subpersonal parts to which selves may be reduced, but to appropriate experiences, emotions, and bodies. If such appropriation is regarded as rational, as it might well be on many Western views, then concern for whatever is appropriated probably would also be regarded as rational.²⁸ Candrakīrti, however, uses his account of the appropriative function of self-talk not to make a case for the rationality of special concern for ourselves in the future, but to speak to the idea, which some find suggested by the Buddha, that the concept of self, even correctly understood, is an ill from which we must be cured and, as such, has no place in a properly constituted mental life.²⁹ If this is right, it would help to explain something that is left obscure in many accounts of Buddhism: how one can subscribe to a no-self view, and thereby align oneself to that extent with the truth, yet nevertheless remain unenlightened. It helps to explain this by pointing out that how one understands self-talk is only one part of aligning with the truth. Another is eliminating the practice of appropriating experiences, emotions, and bodies as one's own. The suggestion that this latter element of aligning with the truth is crucial gives content to the Buddha's saying that "clinging" or "attaching" is the fuel that feeds the fire that is the idea of me and mine. It also explains why properly understanding the nature of the self can go only so far in bringing about enlightenment.

Jonardon Ganeri, in apparent endorsement of Candrakīrti's view, concludes that "learning to think of oneself as a whirlpool of self-appropriating

actions” is a remarkable and potentially transformative achievement.³⁰ He sums up the view that is embedded in this achievement as follows:

The utterance of “I” serves an appropriative function, to claim possession of, to take something as one’s own. The appropriation in question is to be thought of as an *activity* of laying claim to, not the making of an *assertion* of ownership. Grammatical form notwithstanding, the avowal or self-ascription of a mental state, “I have a pain,” is not a two-place relation between me and my pain; nor is it like a club’s having members, or a tree’s having roots. . . . When I say “I am in pain,” I do not *assert* ownership of a particular painful experience; rather, I *lay claim to* the experience within a stream.

Ganeri continues:

This is a performativist account of the language of self, in which “I” statements are performative utterances, and not assertions, and the function of the term “I” is not to refer. This account has the virtue of elucidating the relation between “I” and the psycho-physical stream, and it clarifies the sense in which the facts of ownership are the “further facts” left out of account by a reductionist theory of self.³¹

But, even if this view about the self is true, it would be one thing to understand it and another to live it. How could one live it? What would be the consequences? Ganeri cites Nāgārjuna and Candrakīrti as answering that it is possible to abandon all activity of self-appropriation of the psychophysical, thereby completely transforming “oneself,” and that this would usher in enlightenment and end rebirth. Sounds good, perhaps. But, good or bad, it unquestionably sounds like it would be a big deal practically.

NOTES

Earlier versions of this chapter were read at Cambridge University and at Colgate University. Thanks to people at both of those events, especially Mark Siderits, Ulrich Meyer, and Jay Garfield, for helpful comments. Thanks also to Brad Rives.

1. Siderits 2003: 69.

2. Parfit 1984: 211.

3. Siderits 2003: 69.

4. Butler 1852: 331–332.

5. Including, for instance, Henry Sidgwick and Richard Swinburne. For details, see Parfit 1984: 307–309 and Martin and Barresi 2006: 215–216. For discussion of

the closely related question of whether “nihilism” is depressing, see Olson 2007: 180–210.

6. For discussion of this aspect of Vasubandhu’s views, see Ganeri 2007: 166–167.

7. For discussion of what is taken to be Candrakīrti’s irrealism about the self, see *ibid.*: 196–203. For the distinction between irrealism and fictionalism, see Garfield 2006.

8. An interesting thing about this scene, which was known to Plato (*Theaetetus* 152e) and subsequently widely discussed in late antiquity as the “growing argument” (see Sedley 1982), is that 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).

9. Butler 1852: 332 and 334.

10. In Parfit’s view, for instance, it is partly because selves or persons consist solely of brains and bodies, and interrelated physical and mental events, that they are a kind of linguistic fiction. But the subpersonal parts into which persons are reduced are real. See Parfit 1984: 211 and Parfit 1995: 13–45, the latter of which is reprinted in Martin and Barresi 2003: 292–317.

11. For more on these notions and the relations among them, see Horgan 1993, Kim 1998, Gillett and Loewer 2001, and Shoemaker 2001.

12. Johnston 1997 and McDowell 1997.

13. Johnston 1997: 154–156.

14. Parfit 1984: 311.

15. I am borrowing the term “ironic engagement” from Siderits 2003.

16. Strawson 1997; reprinted, with a “Post Script,” in Martin and Barresi 2003: 335–377.

17. Perry 1976. For a discussion of Perry’s view in connection with the psychological issue of identification, see Martin 1993.

18. Parfit 1984: 327–329.

19. As Parfit pointed out in giving expression to the moderate claim (*ibid.*: 311–312).

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. These relationships could obtain in a situation of limited options, such as the situations in which humans almost always find themselves, in which continuer-interest and self-interest 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 1984: 279–280.

22. An interesting discussion of whether the idea of free-floating experiences really is incoherent may be found in Dainton 2008: 244–249.

23. Williams 1998a: 174–175.

24. Siderits 2003: 99–100.

25. *Ibid.*: 104–105.

26. *Ibid.*: 105.

27. 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. (Siderits 2003: 107–108)

But, if I understand Siderits correctly, a problem with his analogy is that he seems to be recommending a kind of pretense in which one does *not* need to engage to have civic pride. 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.

28. See, for instance, Johnston 1997, in Martin and Barresi 2003: 269: “In barest outline, the defense of self-referential concern would be that we find it utterly natural, and that, at least so far, critical and informed reflection on such concern has not made it out to be unreasonable.”

29. Ganeri 2007: 159, 185.

30. *Ibid.*: 204.

31. *Ibid.*: 202.