Raymond Martin
Philosophy/Union College
Schenectady, NY 12308
martinr@umail.union.edu
Office phone: 518-388-8011

What Really Matters

If our lives were truly horrible and the prospects of their getting better hopelessly dim, we might not want to continue. Fortunately, most people are not in such dire straights. Most of us, as we say, "want to live" (henceforth, to survive). Not only do we want to survive, but under normal, relatively favorable circumstances most of us, it seems, knows well enough what it would mean to survive. What it would mean is that one has a continuing opportunity, under circumstances that one would regard as at least minimally acceptable (that is, as better than death), to have experiences and to act in the world.

But when, even under normal, relatively favorable circumstances, we say that we want to survive, what is it that we *really* want? The obvious answer is that what we really want is simply to survive. Under such circumstances, over relatively short temporal intervals, and for practical purposes, there is nothing mysterious about this answer. When under such circumstances we want to survive, what we really want is simply to survive, and usually that is what we get.

Normally we do not change that much from moment to moment. So, for most of us, surviving for a relatively short time under normal, relatively favorable circumstances simply means continuing pretty much as we are right now. However, even under normal, relatively favorable circumstances, surviving for a long time–say sixty years–

almost always means something else. Over long periods, people tend to change quite a bit–their bodies age, their minds develop, and then perhaps decay, their tastes change, their values evolve, and so on. From the perspective of one's current values, most of these changes may be at least minimally acceptable, including that there would be changes in one's values. But some possible changes–say, severe Alzheimers or really repellant changes in one's values–may be so unacceptable that one would not regard as worth living a life that continued in that way. One wants to survive, one might say, but not like that!

So, when under normal, relatively favorable circumstances people say that they want to survive, what they tend to mean is that they want to continue to live a life that is worth living–that is, that is worth living from a point of view that embodies their current values, or from a point of view that has evolved from one that embodies their current values by some process that they currently endorse.

Can we sensibly ask someone who wants to have a continuing life that from the perspective of her current values is worth living, "In wanting to have that sort of life, what is it that you *really* want? Yes. We can sensibly ask this question if what we mean by it is, in wanting to survive, what do you want *in addition to* wanting a continuing life that is worth living? The answer may be something like, I want to be president, or I want to be a rock star. But what if we're not asking that question–not what, in wanting to survive, one wants *in addition* to wanting a continuing life that is worth living. What if we're simply asking her what, in wanting to survive, that is, in wanting a continuing life that is worth living, she *really* wants. Might our question still make sense? It might make

sense if we are asking for information about the sorts of circumstances under which she would regard her life as continuing to be worth living. It is at this point that philosophy enters the picture.

For the past thirty or so years analytic personal identity theorists have been preoccupied with hypothetical choice situations that are responsive to peculiar versions of this latter question. A typical choice situation of this sort goes something like this: Assume that you could choose either (i) to continue to live in a way that you would regard as worth living or (ii) to cease painlessly and be replaced by someone who resembles you closely (or, by two people who resemble you closely) in ways that matter to you and who would have a continuing life (lives) that is (are) worth living. Assume also that whichever of these options you chose, you will get it. Then, on the basis *solely* of a desire that in a normal choice situation would count as your simply wanting to survive, choose (i) or (ii).

Note that to ask a question that is philosophically interesting one cannot simply say, in the last part of this three part instruction, that your choice should be motivated solely by a desire to survive, that is, to have a continuing life that is worth living, where that implies continuing to exist as the person you are in a way that is worth living. If the instruction were phrased that way, then the outcome would be obvious, but uninteresting: choose (i). For philosophical purposes, it is critical that the second option–option (ii)–come into play.

In a normal choice situation, option (ii) would not be available, or not available as a genuine option that one might be tempted to choose. In the choice situation envisaged

above it is available as a genuine option. That's what makes this choice situation philosophically interesting. So, to make the answer to our question philosophically interesting, in the instruction about the basis on which one should choose, one has to say something like, "solely on the basis of a desire *that in a normal choice situation* would count as your simply wanting to survive, that is, wanting to continue to exist as the person you are in a way that is worth living."

A cost of making the choice situation more philosophically interesting in this way is that it becomes less clear. Replacing the expression, "by a desire to survive," with the expression, "by a desire that in a normal choice situation would count as your simply wanting to survive," introduces an ambiguity into the instruction. With this replacement it is no longer as clear on what basis you are being asked to make your choice.

Assuming that the basic idea behind the original question is nevertheless clear enough to proceed, the set-up is something like this. Imagine someone with only one motive–to have a continuing life for oneself that is worth living. Imagine that this person has never been in a choice situation that included option (ii). Then put her in a choice situation that includes both option (i) *and* option (ii). In such a choice situation, could the chooser by selecting option (ii) obtain enough of what she would have obtained by selecting option (i) that obtaining (ii) would be as satisfying, or almost as satisfying, from the point of view of her motive to have a continuing life that is worth living, as it would have been had she selected (i)?

That is the question that I want to ask. *It is a factual question about the chooser's actual values*. More specifically, it is a factual question about what's most fundamental–

what really matters–in that part of the chooser's actual values that has to do with her ordinary desire to survive. As such, the answer to it can be determined, if at all, only empirically. It is not a question about what *should* matter in survival and it is not a question about metaphysics. Questions about what *should* matter in survival and about metaphysics have been what has primarily preoccupied philosophers concerned with the puzzle cases in the personal identity literature. Our question, on the other hand, is about what *actually does* matter in survival. What makes our question philosophically interesting is that the answer to it could reveal that what actually matters to people in survival is not personal identity, but something else and, hence, that personal identity is not a fundamental value, but merely one that is derived from some more fundamental value.

To get a feel for what I have in mind by an empirical question about whether some human characteristic is fundamental or derived consider the following analogy. Suppose that the only behavior that someone had ever perceived as fatally risky, and also the only thing that he had ever feared, was riding on roller coasters. You ask this person what he is afraid of, and he replies truthfully, "riding roller coasters." It might, then seem plausible to suppose that *fundamentally* what this person is afraid of is riding roller coasters. However, another possibility is that fundamentally he is afraid of something else and only derivatively afraid of roller coasters. How could we discover which, if either, of these were true?

Suppose that we were to learn that in addition to being afraid of riding roller coasters this person would be afraid of any behavior that he were to perceive as fatally

risky. Then, by appeal to his fear of death, we could explain–in a unified way–not only his fear of riding roller coasters, but fears that he would have if he were faced with other apparently fatally risky situations. In addition, we might be able to explain certain aspects, otherwise inexplicable, of his fear of riding roller coasters, such as why some parts of the ride–the ones in which he takes himself to be in greater danger of dying–are scarier to him than other parts. By contrast, there would be no explanatory advantage to the thesis that fundamentally he is afraid of roller coasters.

This example is about fears, rather than wants, so it is not an exact analogy. Nevertheless, it illustrates how some mental state that is more fundamental can as an ingredient in and both explain and all but exhaust the significance of something less fundamental that is derived from it, and how all of this could be determined empirically. While there is a great deal more that might be said about the relationship between fundamental and derived, for now it is enough to note that when it comes to empirical questions about what actually matters in survival, what's fundamental in this sense can be used *to explain* what's valuable about what's derived–that is, why what's derived matters. It can explain this because what's fundamental is an ingredient of what's derived, it's value as an ingredient of what's derived all but exhausts the value of what's derived, and if one were to get just what's fundamental, and not also get what's derived, one would have gotten something as valuable, or almost as valuable, as what one would have gotten if one had gotten what is derived.

Could anything like what we imagined might happen in the roller coaster example happen in connection with empirical versions of the thesis that one's personal identity is

not fundamental, but derived? I think so. Suppose, for instance, that some fission examples were to reveal that in certain situations many of us, if we were to choose solely on the basis of a desire that in a normal choice situation would count as our simply wanting to survive, that is, wanting a continuing life for ourselves that is worth living, would actually prefer fission, even if we thought that our fissioning meant that we would cease, to options that we thought would allow us to persist under relatively favorable circumstances. Suppose that in certain situations we would have this preference even if our prospects without undergoing fission were quite bright–as bright or brighter than the actual prospects of any who are reading this paper–including you–have ever been.

In the case of those who in such circumstances would nevertheless choose fission, thinking that their getting this option would mean that they would cease, there would be a prima facie reason to believe that their persisting is not what matters primarily to them in their desire to survive and that the value that in actual choice situations they place on persisting is not fundamental, but derived. Depending on how the example were specified, it could turn out that in the case of those who in such circumstances would choose fission, there would be no explanatory advantage to the thesis that fundamentally persisting–personal identity–is what matters in survival. Are there any examples of this sort? Yes.

Fission Rejuvenation

Consider, for instance, a simplified version of an example involving a procedure I've elsewhere called *fission rejuvenation*.[1] In this example, someone, *A,* who is 20 years old

and in very good shape, both physically and psychologically, knows that even without doing anything unusual his prospects are good for a long and happy life. Nevertheless, he opts for an unusual procedure–fission rejuvenation. This involves his going to a hospital where he is put under a general anesthetic after which his brain is divided into functionally equivalent halves, each capable of sustaining his full psychology. Each half is then put into a body of its own that is qualitatively identical to his pre-fission body, which has been promptly destroyed. Hours later, one of A's fission-descendants, B, wakes up and begins a brief, painless recovery after which B leaves the hospital in excellent health, looking and feeling almost exactly like A looked and felt just prior to his undergoing the procedure. Except for such differences as are occasioned by B's knowing that the procedure took place and that another fission-descendant, C, also exists, the prospects for B's subsequent physical and psychological development (and ultimate decay) will be just like A's would have been had A not undergone the procedure. The other fission-descendant, C, has a different fate. Before C awakens he is given a drug that puts him into a deep, dreamless coma and preserves his body in its current state until he is awakened. As it happens, C is awakened fifty-five years hence. Throughout these years C is kept safely in the hospital.

During the operation a small device is implanted in B's brain that continuously scans all of his brain activity and immediately transmits complete information about what it finds to a similarly small device, designed to receive its signals, which has been implanted in C's brain. This latter device quickly encodes onto C's brain the information it receives from the transmitter, just as this information would have been encoded had it

been acquired as a normal product of changes originating in C. As a consequence every psychological change encoded in B's brain, which functions normally, is encoded almost instantly in C's brain in virtually the same way as it was encoded in B's brain. Thus, throughout the time C is in a dreamless coma he has a *dispositional* psychology exactly like B's of a few seconds in the past even though C is unconscious and does not age physically. Hence, except for these dispositional psychological changes, C remains as A was when A originally began to undergo fission rejuvenation.

Fifty-five years after this procedure is performed B dies from an independently caused heart attack. As B draws his last breath the device implanted in his brain sends a signal to the device in C's brain that causes C to wake up and begin a recovery period similar to the one that B underwent immediately after the procedure, after which C leaves the hospital, a *psychological* replica of *B* when B died, but a *physical* replica of A when A initially underwent the procedure. Once awake, C begins to age normally.

More needs to be said (some of which has been said) to develop this example so as to ward off various objections that might be raised.[2] I claim that it would be relatively easy to do this so that A's undergoing the fission rejuvenation procedure described in the example would unquestionably be a good deal for A. What it *means* to say that fission rejuvenation would be a good deal for A is that given his values, which we may suppose are both normal and healthy (though not the only normal, healthy values a person might have), and given his circumstances and reasonable expectations, were A faced with the option of either undergoing fission rejuvenation or of continuing normally, his better selfish choice (even though his prospects, without fission rejuvenation, are

quite bright) would be to undergo fission rejuvenation. What *would make* fission rejuvenation such a good deal for A is that by undergoing it he would secure a benefit as good (or, almost as good) for him as his doubling what would have been his remaining adult life span plus a benefit as good (or, almost as good) for him as his recovering physical youth in what otherwise would have been his old age. In other words, by undergoing fission rejuvenation A would be continued by B in pretty much the same ways and for the same length of time that without the procedure he would have persisted; and then instead of his dying when he would have died (that is, when B died), he would be continued by C. As a consequence, when A is seventy-three years old, he would be continued in a physically youthful body, with a normal life expectancy, similar to the one A had when he underwent the procedure, but without sacrificing B's psychological development or memories.

A crucial part of what would make fission rejuvenation such a good deal for A is that by undergoing the procedure A would create two fission-descendants of himself with each of whom he could fully and rationally identify. This implies that A could anticipate (or quasi-anticipate) having the experiences and performing the actions of each of his fission-descendants pretty much as he would otherwise would have anticipated having his own future experiences and performing his own future actions.

Extrinsic considerations, such as whether A thinks his fission-descendants will be able to maintain his significant personal relationships and, more generally whether he thinks they will be able to play the social roles that he now plays might also influence his ability to identify with his fission-descendants. As a consequence such extrinsic

considerations may profoundly affect how good a deal A thinks it would be for him to undergo fission rejuvenation. And were A to undergo fission rejuvenation, B's and C's extrinsic relationships might profoundly affect how good a deal it actually was for A to have undergone fission rejuvenation. But since, in this example, there is no reason to think that B, at least, would have to have any trouble slipping neatly into all of A's social roles (A and B could keep it a secret that the procedure was performed) I shall not have much to say about these extrinsic considerations. It is enough that were fission rejuvenation inexpensive, reliable, and painless it could be a good deal not only for A but for many young people with normal Western values who are in good physical and psychological shape and for whom, as I am supposing is the case with A, extrinsic considerations would not play a defeating role.[3]

So what?

What is the relevance of the fission rejuvenation example to the question of what *actually matters* fundamentally in survival? If I am right that many otherwise rational people who were given a chance under the circumstances specified to opt for fission rejuvenation would elect to undergo the procedure, then the example shows that in response to a hypothetical choice situation in which one’s prospects, were one to continue normally, are extremely bright and in which one believes that in opting for fission one is (or may well be) opting for one’s own cessation, many would opt for fission. Call this *the surprising result*. What are we to make of it?

According to some philosophers, not much. Collectively these philosophers have come up with four reasons to justify their dismissing the importance of the surprising

result. First, some of them–Kathleen Wilkes is the prime example–object that what people would choose in hypothetical situations in which they would be different than they actually are in real life can show nothing about their real life values. However, there is a basic problem with this objection. It is that such hypothetical examples can expose what we would choose if we were just like we actually are except for differences that would not affect our values but simply afford us options to choose that we do not now have. So long as whatever affects our choices in these examples, including our desires, motives, beliefs, intentions, values, and so on, are just what they actually are in real life and so long as we mentally process whatever affects us, except possibly for our processing it redundantly, just like we would process it in real life, then whether we would choose fission over other options, if we had the choice to do so, is evidentially relevant to determining what our real life values actually are. In short, the assumption that we are people in the circumstances depicted in such examples does not require us to suppose that our values or how we process and express them are any different, except perhaps for their being redundantly encoded in our brains, than they actually are. And redundant encoding is not a kind of difference that would block inferences from the choices we would make in such hypothetical circumstances back to our actual values.[4]

Second, some philosophers–perhaps Peter Unger is the prime example–claim that those whose expressed preferences yield the surprising result are in this instance choosing irrationally by choosing to thwart their own most deeply and rationally held values. Possibly this objection has merit in response how one should choose in response to many fission examples in the literature. It is hard to see how it has merit in

response to the fission rejuvenation example. Choosing fission rejuvenation is so eminently sensible that one would have to be in the grip of a dubious theory to deny that it could be an attractive option. Elsewhere I have responded to arguments– primarily by Unger and Ernest Sosa–that would imply that I am mistaken in thinking that choosing fission rejuvenation would be a sensible choice for someone like *A*.[5] Since these arguments and my replies to them are not easily summarized, I shall not repeat them here.

Third, some philosophers have denied that someone who fissions ceases and is replaced by somebody else. Most philosophers who take this line do so because they subscribe to a four-dimensional view of persons–here, David Lewis is the prime example.[6] These philosophers claim that the person-stage that makes the choice to fission is included in every person who includes one of the post-fission person-stages. Hence, in choosing to fission, one is not choosing to cease, but simply to split. Others–Daniel Kolak is the prime example–deny that if someone chooses fission he chooses to cease and be replaced by somebody else on the grounds that there is only one person, hence no one else by whom one might be replaced.[7]

Whether a four-dimensional view of persons is preferable to a three-dimensional view is debatable. And, of course, it is also debatable whether there is just one person. Whatever the outcomes of these debates, they do not threaten the conclusion that the fission rejuvenation example shows that many otherwise rational people would under certain circumstances choose what they take to be ceasing and being replaced by others rather than choosing to persist. And quite clearly this is a result that reveals

something important about these choosers' values.

The most that a four-dimensionalist or a one-person theorist is entitled to claim in response to such fission examples is that those who choose to extend their special concern to their fission-descendants, whether or not the choosers are aware of it actually extend their special concern only to themselves (or to person-stages that are temporal parts of themselves).[8] But this leaves open the possibility, which seems to actually obtain, that many who choose to extend their special concern to their fission-descendants do so thinking that, and intending, that they thereby are extending their special concern to others. Their thinking this may be mistaken, but what they think they are choosing nevertheless reveals something important about their values.

By analogy, a married man's choosing to try to seduce a woman under the mistaken assumption that she is not his wife, thereby reveals something about how much he values being faithful to his wife. If it turns out that the woman he tries to seduce is his wife in disguise, it would not thereby follow that in trying to seduce her, he was being faithful to his wife. On the contrary, since he believed at the time that the woman he was trying to seduce was someone other than his wife, and in having this belief he was not self-deceived (that is, he did not really believe, deep down, that the woman was his wife), then in trying to seduce her he was trying to do something that he knew would mean that he is unfaithful, and it's *that* fact, not the fact that the woman he tried to seduce was his wife, that reveals something important about his values.[9]

One might object that this analogy is misleading in that the man might not have wanted to seduce the woman (at least on the occasion in question) if he had realized

that she was his wife, but those in the fission rejuvenation example who choose an option that they think involves their ceasing and being continued by others would be just as enthusiastic, perhaps even more enthusiastic, to choose that same option were they to realize that it were one in which they would persist.

That may be true (I think it is true), but its being true does not negate that in response to the fission rejuvenation example those who choose option (ii), thinking that it involves their ceasing and being continued by others, are nevertheless happier with choosing what they take that option to involve then they are with choosing option (i), which they think involves their continuing normally. And they are happier, even when their prospects where they to opt to continue normally, are quite bright. That surely shows something important about their values. What it shows is that persisting is not nearly as important, or as fundamental, a value as has traditionally been supposed.

Suppose, however, that we were to assume that it is not the chooser's understanding of what she is choosing, but rather what she actually chooses, that is decisive in determining what her choice reveals about her values. Might it still be possible, on a four-dimensional view of persons, to frame an empirical question about what actually matters most fundamentally in survival? Yes. I shall explain how below. However, it would not be possible on a one-person theory.

Finally, Mark Johnston has argued, from the perspective of a three-dimensionalist about persons, for a "not much" response to what one can learn about a person's values from fission examples. He allows both that hypothetical examples of the fission rejuvenation sort may be relevant as evidence of what people do (or should)

value and that choices that yield the surprising result may be rational. His main point is that, even so, we should “quarantine” choices that yield the surprising result.[10] What he means by this is that since the choices that yield the surprising result are made in unusual circumstances they provide no basis for questioning whether in normal situations personal identity is what matters primarily in survival.

Johnston admits that in some fission cases it would be *reasonable* to care about each of one’s fission-descendants as if each were oneself. He says, for instance, that even though he *does not* believe that he would “determinately survive” a familiar sort of fission procedure in which his cerebral hemispheres were divided and both were transplanted into different receptacle bodies, but *does* believe that he would survive such a procedure if just one of his hemisphere were transplanted, he would not “make a significant sacrifice” to ensure that just one was transplanted. However, he claims, the most that it would be reasonable to conclude from this is that in cases like this “(neurally based) R, and not identity, is the relation in terms of which one should extend one’s special concern.” One would not be warranted, he claims, in concluding “that quite generally it is (neurally based) R that matters.”

Why wouldn’t one be warranted in concluding that quite generally it is (neurally based) R that matters? According to Johnston in such fission examples, at least two presuppositions of our special concern are violated: first, that it is always a determinate matter whether one is identical with some given future person; and, second, that at most one future person will continue one’s mental (and physical) life. When such presuppositions are violated, he says, future-directed concern neither determinately

applies nor determinately fails to apply. Hence, in such cases, it is reasonable to try to find a natural extension of future-directed concern. When a significant core of the relations that constitute identity in the determinate cases is present, “an appealing idea,” he says, is to extend one’s special concern to one’s fission-descendants. In his view, the holding of neurally based psychological continuity and connectedness, a relation that in many fission examples holds twice over, would count as being a significant core of the relations that constitute identity in the determinate cases. But, he insists, extending one’s special concern in such cases would be reasonable “*not because identity is never what matters*” but, rather, because “*caring in this way represents a reasonable extension of self-concern in a bizarre case*.”

Johnston says that inferring from the consideration of such fission examples that identity is not primarily what matters in survival is “much more radical” than “the indeterminate cases” require. He concludes that “there is a false apparatus of generalization” against identity-based concern at the heart of arguments based on such examples. The “relevant presupposition of self-concern is the holding of the determinate, ordinary fact of personal identity or difference,” he says. Since this presupposition is almost always met fission examples have “no effect beyond the imaginative fringe. What is not there in the fission case,” he says, “is almost always there. Identity is still almost always what matters.”

I agree with Johnston that it does not *follow* from the surprising result that identity does not matter primarily. But, it seems to me that he inaccurately characterizes the data that need to be explained. In his view, if we project ourselves into the role of the

protagonists in the sorts of fission examples that generate the surprising result, we are presented with a quandary–whether or not to extend special concern to our fission-descendants. He claims that we can solve this quandary reasonably either by extending our concern or by withholding it–that it is up to us which we choose to do, and that either choice is reasonable. True enough for those who find the prospect, say, of fission rejuvenation simply puzzling. But Johnson's characterization of the way in which many people would respond to such fission examples gets the phenomenology wrong. In projecting ourselves into the place of the protagonists in such examples, there is, for many of us, no quandary. That is, for many of us it is not a question of *whether* to extend our special concern to our fission-descendants. Rather, in considering the examples we *discover* that our special concern *already* extends to at least some of those who in the examples would be our fission-descendants, say, to B in the fission rejuvenation example.

Johnston's concession that even though he *does not* believe that he would survive a procedure like fission rejuvenation, but *does* believe that he would survive such a procedure if just one of his hemispheres were transplanted, he still would not "make a significant sacrifice" to ensure that just one was transplanted, suggests that in the case of some fission examples he too discovers that his special concern already extends to some of the fission-descendants. Otherwise why wouldn't he make a significant sacrifice to ensure that just one of his hemisphere were transplanted? Is whether he ceases such a light matter that he would willingly cease in deference to a rationally arbitrary extension of the conventions that govern special concern, for which

he feels no special pull?

Regardless of the answer to that question, contrary to Johnston's characterization, in the case of many of us who contemplate certain fission examples there is no deliberative problem that needs to be solved. The examples *reveal* something about what we *already* value. What they reveal is that our special concern *already* extends to at least one of our fission-descendants–that we *already* have a *strong preference* in the choice situations depicted in such examples to adopt options which conflict *prima facie* with what philosophers traditionally have supposed is what matters primarily in survival. That is why such examples, when they were first introduced, were *surprising*. If all that such examples revealed was that in cases like those depicted there is no particular pull to extend familiar patterns of concern to unfamiliar situations one way rather than another, that would have been less surprising, if it had been surprising at all. Rather, what is surprising, assuming that in such examples our fission-descendants really would be others, is how compelling it can be to extend our special concern across a divide that one believes separates self from other.

Partly because Johnston fails to acknowledge this central aspect of many peoples' responses to such examples he also fails to *explain* the surprising result. In his account, so far as it goes, extending one's patten of special concern, while reasonable, is essentially arbitrary–a brute fact. Hence, nothing explains the surprising result. In other words, he does not acknowledge that in response to such hypothetical examples, many discover something about what they already value, rather than create new values, so he feels no need to explain how they could have discovered something about their

existing values. Of course, even on the mistaken assumption that those of us who in response to such examples choose to extend special concern and thereby create new values, there would still be something to explain–why the very people who make this choice choose to extend, rather than withhold, their special concern. But if the choice to extend or withhold were truly arbitrary, there may be no psychologically interesting explanation of why specific people choose as they do. In any case, all of this, since it is based on a mistaken assumption, has little to do with how people actually respond to such examples.

On the assumption that fission-descendants really are others, one of the most interesting things about the surprising result, which even Johnston acknowledges, is that those who, in response to such examples, feel special concern for their fission-descendants are tracking not personal identity, but something that obtains *both* in normal cases of self-concern *and* in the hypothetical fission examples under consideration. Johnson says that this common element is "neurally based psychological continuity and connectedness." Suppose that he is right about that. Then a unified explanation of the survival-preferences of such choosers could be provided and, hence, just as in the earlier "fear of death" example, there would be an explanatory advantage to providing one. The unified explanation would be that what actually does (not should, but does) matter to many people in survival, whether in normal situations or hypothetical examples, is neurally based psychological continuity and connectedness. If that were true, it would be a good reason for holding both that in the cases of many people personal identity is not what actually matters primarily in survival and that in those

normal cases when it does matter, its mattering is not fundamental, but derived.

What about what actually matters to those who in response to hypothetical fission examples of the sort Johnston describes would extend their special concern to their fission-descendants? Is it, as he assumes, neurally based psychological continuity and connectedness? No doubt it is for some. But it may not be true of everyone and it may not even be the most fundamental truth about the values of those of whom it is true. To find out one would have to consider a variety of examples. Notoriously, in the case of some hypothetical examples, some people report that they feel special concern for certain descendants, such as teletransportation descendants, with whom they have no neurally based psychological continuity.[11]

Explanation Revisited

What would it mean for one's special concern to be derived, rather than fundamental? Answering this question requires a little stage setting.

Values, in the sense in which I have been using the term, are things that one wants. A person realizes her values if she gets what she wants. So, if a person wants to write a good book and she writes one, then she gets what she wants and thereby realizes that part of her values. Of course, a person might get some things that she wants–say, world peace–not by her doing anything in particular, but simply by something happening.

Egoistic values, on the other hand, are those values that a person embraces because she thinks that their realization will benefit herself primarily. For instance, a person's desire to be happy, not in order to benefit others, but to improve her own life,

expresses an egoistic value that she has–her own happiness. Her desire that others be happy, not in order to benefit herself, but to improve their lives, expresses a value that she has that is not egoistic–the happiness of others.

Personal identity based, egoistic values are those egoistic values that a person embraces at least partly because she thinks that their realization requires that she exist. In the what-matters-primarily-in-survival literature, it has been *future-oriented*, personal identity based, egoistic values, rather than *past-* or *present-oriented* values, that have loomed large in the debate.

Future-oriented, personal identity based, egoistic values are those *personal identity based, egoistic values* that a person embraces at least partly because she thinks that their realization requires that she exists at some point *in the future*.

Future-oriented self-concern, which is a species of *future-oriented, personal identity based, egoistic values*, is the concern that one has for someone who will exist, or who one imagines will exist, in the future because one thinks that that person is, or would be, oneself. For instance, I care especially and in a distinctive way about what will happen tomorrow to that person who is (or will be) me because that person is (or will be) me.

S*elf-referential concern*, on the other hand, is the concern that one has for persons or things because they stand in certain special relationships of ownership or appropriation to oneself.[12] For instance, I care especially and in a distinctive way about what will happen tomorrow to my children because they are *my* children and to my car because it is *my* property. In short, self-concern, in the sense in which I am

understanding it, is concern for what is (or was, or will be) oneself because it is oneself, while self-referential concern is concern for what is (or was, or will be) one's own because it is one's own. Typically, future-oriented, personal identity based, egoistic values get expressed in self-concern, not in self-referential concern.

Against the backdrop of this stage-setting, what does it mean for a person's future-oriented, personal identity based, egoistic values to be derived, rather than fundamental? What it means is that a person embraces non-egoistic values which are such that were she to realize these, she would also realize her future-oriented, personal identity based, egoistic values. I have argued above that the surprising result is evidence that for many people, personal identity is not what matters fundamentally in survival, that is, that it is a derived value. However, in making this argument I assumed that the chooser's understanding of what she is choosing is decisive in determining what her choices reveal about her values. But suppose that we were to drop this assumption and assume instead that what's decisive is what the chooser actually chooses, at least when she gets the options that she wants and when what these options actually involve is at least as desirable to her as what she thought they involved. Might it still be possible to frame an empirical question about what actually matters most fundamentally in survival, the answer to which is that something other than personal identity matters most fundamentally? I think so.

Consider, first, what, on a three dimensional view of persons, would suffice to get this result. Suppose, for instance:

$ that what matters to me primarily in survival is being related to someone in the

future in a certain way–say, by being psychologically continuous with and connected to (R-related to) someone;

$ that being R-related to *someone* in the future does not necessarily ensure that I will be R-related to *myself* in the future;

$ that I egoistically value being R-related to myself in the future; and

$ that in the circumstances in which I find myself, the only way to be R-related to someone in the future is to be R-related to myself in the future.

In such a situation, my future-oriented, personal identity based, egoistic value of being R-related to myself in the future would be derived, rather than fundamental. It would be derived since I want to be R-related to someone in the future, but not necessarily to myself, and in the circumstances can accomplish this by being R-related to myself and only by being R-related to myself.

On a four dimensional view of persons, people who exist now and will exist in the future cannot exist *wholly* at either time. So, we would have to describe the example just given differently. Rather than saying that one egoistically values being related to *oneself* in the future, we should say that one egoistically values being related to future person-stages that are (or, will be) temporal parts of oneself.

To see what difference this makes, suppose, as before, that what matters to me primarily in survival is being R-related to someone in the future, but that I also egoistically value being R-related to myself in the future. Then, on a four dimensional view, my personal identity based, egoistic value of being R-related to *myself* in the future would be derived if two conditions are satisfied: first, it were *not possible* for any

current person-stage of mine to be R-related to anyone whose future person-stages do not include my present person-stage as a temporal part; and, second, that the reason for this is that the unity-relation among person-stages has *by stipulation* been made to match that of the independently determined R-relation (as David Lewis put it, "the I-relation is the R-relation"[13]). In such circumstances, what would make the personal identity based, egoistic value of being R-related to myself in the future derived is that the unity relation–the I-relation–and the R-relation are the same and that the R-relation, as it were, leads the way, while the I-relation follows its lead. However, if this is how the personal identity based, egoistic value of being R-related to myself in the future is derived, then the thesis that it is derived is not an empirical thesis. For it to be an empirical thesis, it cannot have been arranged by stipulation that the I-relation and the R-relation are the same.

Suppose, then, that we do not stipulate that the I-relation and the R-relation are the same and that, as it happens, they sometimes diverge. In such a case, what it would mean for my future-oriented, personal identity based, egoistic value of being I-related to myself in the future to be derived is that there are no actual or possible cases in which my current person-stage is I-related to some future person-stage to which it is not also R-related.

The empirical project

I have argued elsewhere that the project of determining what *should* matter in survival has reached a dead end.[14] Even if I am right about that and even if empirically one could give a more unified explanation of the ways many people *actually* feel special

concern, say, for their fission-descendants, by assuming that personal identity is not what matters primarily in survival, why should philosophers bother with the project of what *actually* matters? Isn't that a job for psychologists, rather than philosophers? Yes. But, on this issue, philosophers can contribute importantly *as psychologists* in a way that it is unlikely that psychologists proper will ever themselves contribute. The reason is that it is only by considering hypothetical examples that plausible candidates for what matters primarily in survival other than personal identity can be isolated from personal identity–persisting–so as to generate evidence that personal identity is less important than we may have thought that it was. It seems unlikely that psychologists proper, at least any time soon, are going to consider such examples.

So, philosophers may be the appropriate people to conduct this empirical study. Even so, it is unlikely that the study itself would lead to the sort of results that one usually strives for in empirical studies of peoples' values. For how one responds to any given hypothetical example that sheds light on one's values may be profoundly influenced by how one responds to other examples and by how deeply immersed one is in the philosophical literature on personal identity and related issues. And to make matters worse (or better, depending on one's point of view), notoriously the puzzle cases have the potential not only to reveal what one values, but to change it. So, the data might be constantly in motion.[15]

What's more, even if we could discover what people in general value in survival, it may not be that interesting to discover it. What most of us who dwell on such questions most want to know, I suspect (certainly what I most want to know), is not what

people in general value in survival, insofar as this is revealed in their responses to the puzzle cases about personal identity, but what people who have thought deeply about the puzzle cases and related issues value, for only such people are likely to understand in appropriate depth what is involved in the choices depicted in the puzzle cases.

In light of these difficulties, why bother with an empirical inquiry into the importance of personal identity in survival? Almost certainly one shouldn't bother with it if all one is interested in is a general theory of the importance of personal identity in survival. One should bother with it only if one is interested in understanding as deeply as one can *one's own* personal identity related values, and perhaps also in understanding them in a way that sets them in motion. Such an inquiry may seem more like a meditative quest than a traditional inquiry in western philosophy or social science. I think that's right. It is an intellectual (as opposed to the experiential) side of a meditative quest. Whether that's a good thing or a bad thing depends on one's values– in this case, on what one really wants *in addition to* wanting a continuing life that is worth living.

ENDNOTES

---

1. See *Self-Concern: An Experiential Approach to What Matters in Survival* (New York: Cambridge University Press, 1998), pp. 53-72. An earlier version of this example may be found in "Fission Rejuvenation," *Philosophical Studies*, 80 (1995) 17-40, reprinted in R. Martin and J. Barresi, eds., *Personal Identity* (Oxford: Blackwell, 2003).

2. *Self-Concern*, especially pp. 57-72 and 85-92.

3. One might doubt whether fission rejuvenation results in A's cessation. I shall deal with two metaphysical motivations for this worry when I briefly consider the views of David Lewis and Daniel Kolak. Quite apart from the sorts of considerations these two

philosophers raise, someone might think that even though A and B, in the fission rejuvenation example, are physically separate and there is only one-way causal influence between them, the procedure in the example does not yield a genuine case of fission. The source of this worry is that it is not until after A's death that their brains develop independently of each other. So, someone may think, prior to A's death A and B may simply be one person with two bodies and two brains. To dispose of this worry, in *Self-Concern* I sketched two variations on fission rejuvenation that clearly are genuine cases of fission. In these variations, the two brains that emerge from the procedure are allowed to go out of synch just enough to quiet any doubts one might have about whether there has been genuine fission.

4. For more on this objection, see *Self-Concern*, pp. 13-15.
5. *Self-Concern*, pp. 57-69.
6. See especially Lewis' paper in Amelie Rorty, ed., *The Identities of Persons* (Berkeley: University of California Press, 1975, pp. 17-40; reprinted with a "postscript," in David Lewis, *Philosophical Papers*, Vol. 1, (New York: Oxford University Press, 1983), pp. 55-70, 73-77. See also Harold Noonan, *Personal Identity*, 2nd ed. (London: Routledge: 2003)
7. Daniel Kolak, *I Am You*: *The Metaphysical Foundations for Global Ethics*. (Dordrecht, The Netherlands: Springer, 2004).
8. See *Self-Concern*, pp. 69-71, for a discussion of Lewis' view.
9. There is no reason to suppose–certainly none has ever been given–that all of those who would choose fission rejuvenation would be self-deceived, that is, that all of them would know, at some level, that they were not choosing to cease and be replaced by others.
10. Mark Johnston, "Human Concerns Without Superlative Selves." In Jonathan Dancy, ed., *Reading Parfit* (Oxford: Blackwell), pp.149-179.
11. Subjectively, what seems to be critical to whether one feels special concern is whether one can anticipate having the future experiences (and performing the future actions) of one's fission-descendants pretty much as one would anticipate having one's own future experiences and performing one's own future actions. Perhaps this is even part of what it means to feel special concern. In any case, what makes the fission rejuvenation example especially useful is that it includes five features that make this sort of

anticipation easier than it is in many other sorts of fission examples. One of them is closely related to the one Johnston mentions: the causal mechanisms underlying the conscious experiences of whichever of the fission-descendants is conscious are (initially) a proper subset of the same ones that redundantly underlie the chooser's own pre-fission experiences. The others features are (i) that only one of the fission-descendants will be conscious at any given time, (ii) that the fission-descendant who is conscious initially–B–will cease to be conscious before the other–C–becomes conscious, (iii) that initially each of the fission-descendants will have the same psychology that he had just before the chooser lapsed into unconsciousness for the last time just prior to his fission, and (iv) that as long as the fission-descendants are *both* alive their dispositional psychologies develop in tandem. Most of the fission examples under discussion in the literature lack one or another of these five features. Their lacking one or another will block some people from anticipating having the future experiences (and performing the future actions) of their fission-descendants pretty much as they would anticipate having their own future experiences and performing their own future actions.

12. I am indebted here to Mark Johnston, who has made the distinction between self-concern and self-referential concern nicely in "Human Concerns without Superlative Selves," pp. 156-58, in Dancy, *Reading Parfit*, pp. 149-77.

13. David Lewis, op.cit.

14. *Self-Concern*, pp. 15-21.

15. Surely Nozick was at least partly right to call the puzzle cases a koan for philosophers, in *Philosophical Explanations* (Cambridge, Massachusetts: Harvard University Press, 1981, p. 47).