

Journal of Cognition and Neuroethics

Philosophical Profiles

Jacob Ross

Associate Professor of Philosophy at USC

IN BRIEF

Jacob Ross is Associate Professor of Philosophy at the University of Southern California. If one had to capture his wide interests under one umbrella term, perhaps “normativity” would be closest. He has written about ethics, epistemology, practical reason and aesthetics, publishing in such diverse and venerable journals as *Ethics*, *The Philosophical Review*, *Mind*, *Ratio*, and the *Journal of Philosophical Logic*, as well as a number of anthologies. He is a bit of a philosophical “bomb-thrower,” defending novel positions that challenge orthodoxy with ingenious and convincing argumentation. In this (albeit not in all things) he resembles the great Derek Parfit, on whose work I asked that he expound, who advised him on his dissertation, and with whom he maintained a philosophical exchange until the latter’s death. In our interview, Professor Ross expounds on Parfit’s most influential arguments and positions and explains, among other things, in what way Parfit is superior to Immanuel Kant and John Rawls, why he himself is definitely *not* a Buddhist, and why Parfit might feel a little betrayed by some of his publications.

DETAILS

Simon Cushing conducted the following interview with Jacob Ross on 15 July, 2021.

CITATION

Ross, Jacob. 2021. “Interview by Simon Cushing.” *Journal of Cognition and Neuroethics* (Philosophical Profiles): 1–21.

Jacob Ross

a philosophical profile

SC: Hello Jacob Ross of the University of Southern California. Thank you very much for agreeing to talk to me and for agreeing to talk to me about Derek Parfit despite it being a while since you have written on him. I asked you to talk about Derek Parfit because of a piece you wrote some time ago, introducing him in *12 Modern Philosophers*, ed. Christopher Belshaw and Gary Kemp (Blackwell, 2009). It was one of the few pieces I found that provided an accessible introduction, which is why I asked you to do this interview. I'd like to talk a little bit about your work too, but the bulk of the interview is going to be about Parfit. But, let's start with you. How did you first become interested in philosophy? What brought you to philosophy?

JR: I would say I became interested in philosophy from quite a young age. By 9th grade, my first year of high school, I was already thinking I'd likely become a philosopher. By 10th grade, I was certain I wanted to be a philosopher. I'm from London, Ontario, Canada and while I was in high school, I was already attending lectures, although I wasn't understanding them very well, at the University of Western Ontario. I went to London Central Secondary School and I had a philosophy column in our school newspaper, which was called The Central Perspective. It was called "Jacob's Ontology Corner."

And, you were beaten up regularly after school, I presume?

(Laughs) My interest in philosophy goes way back. I came from an academic family. So, for me the question growing up was simply, "Which branch of academia will you go into?" At a young age, I was interested in theoretical physics and cosmology and so on, but I did not understand the advanced math to make sense of the physics, so I drifted into philosophy as a way I could be asking big questions without knowing the advanced mathematics.

What were the early questions that made you decide in 10th grade, which is way too young, that philosophy was where your future lay?

I believe, if I recall correctly, I was initially interested in metaphysical questions; questions about free will and determinism were the first ones that captured my imagination. I was also really interested in early analytic philosophy: Wittgenstein, Russell, logical atomism. I remember being enamored by the *Tractatus*.

You were a weird kid!

I was so enamored by it that I attempted to imitate it. I tried writing in a tractarian style. I filled notebooks with numbered little sections with complex numerals beginning each entry.

When did you first come across Derek Parfit?

I came across Derek Parfit, his work anyway, when I was in grad school. At Rutgers, where I went to grad school, we wrote an area exam, and I chose ethics to be the area in which I wrote that area exam, expecting to write my dissertation in ethics, which I ended up doing. Larry Temkin was the advisor of my dissertation, and he was also the person responsible for the area exam. He had been a student of Derek Parfit, so he put a bunch of Parfit's work on the required readings for the area exam. I remember, at the time, I read pretty much everything we were supposed to read for the area exam. I didn't know anything about Derek Parfit at that time, and I just thought there was an inordinate amount by Derek Parfit on the preparation material, so I kind of put it off until last. The last thing I read right before the exam was what he recommended on Parfit. At that point, I realized it was a mistake for me to put it off until the end. It certainly was among my favorite things I read in preparation for it. Then, when I began working on my dissertation, he gave me more Parfit to read. "The things that are most relevant to your project, you have to read this stuff from Parfit and this stuff from Nagel." I remember at the time being very underwhelmed by the Nagel he assigned, but extremely impressed, thrilled actually, enthralled by the Parfit. So my first introduction to him was just being assigned things to read by him by Larry Temkin.

What particularly enthralled you about it? Was it the questions or his way of doing philosophy? Was it the way he did it? Or the things that he said, or the topics that he found interesting, or the connections that he tried to make?

It was the former. It was more his style of philosophy than either the topics or the particular ideas. I was impressed by the density, and I'm using density in the positive sense, of argument. You have all these very brief sections. Each section is only a couple of pages long and each section has a jaw-droppingly impressive argument. I remember one sort of caricature of philosophy is that philosophical writing consists of a series of "I will argue" followed by a series of "I have argued". Very often when you're reading philosophy, little of what is written is actual argument, and when you actually isolate the argument, it's hard to understand, it's not very impressive, it's disappointing, etc. So, let's take for example someone who is considered one of the greatest moral philosophers of modern times, and probably most people would regard him as the greatest modern moral philosopher, namely Immanuel Kant. It's hard to find good arguments in Kant. It's very hard to make sense of what arguments there are and if you try to reconstruct them, what you get is a very weak argument. Whereas Parfit is the opposite. It's like pure argument, there's a ton of arguments and each argument is extremely powerful.

The thing about Parfit is, like you said, most of philosophy is "I will argue" and "I have argued", but what they also do, is "here's why I'm arguing this and what implications this will have for the world in general". I think Parfit is missing a lot of that. So, particularly in Part One of *Reasons and Persons*, a lot of what he's doing you're like, "why is he doing this?" Would you agree?

Somewhat. I don't think I would fully agree with that. I think that in some cases what he's doing is a little bit hard to understand because he's responding to other philosophers who've discussed this, but he doesn't make that clear. I think that there is a mistake

“It's hard to find good arguments in Kant. It's very hard to make sense of what arguments there are and if you try to reconstruct them, what you get is a very weak argument. Whereas Parfit is the opposite. It's like pure argument, there's a ton of arguments and each argument is extremely powerful.”

Philosophical Profiles

that I think some philosophers make, which is they want their work to be eternal, and because they want their work to be eternal, they don't want it to be seen as simply responding to other people's work, so they will avoid referencing that. But, really the only way to understand it is to understand how it relates to other people's work. So, I think a lot of what's going on in the early parts of *Reasons and Persons* is that, which I think can make it a little bit hard to follow. But, I think the topics are important enough that even if you don't understand the connections to other things in the literature, I still find the arguments fascinating.

The story I heard about *Reasons and Persons* is that Parfit was at, that college in Oxford where you don't have to teach [All Souls], where his job was just to think deeply. Eventually they said, "OK, enough thinking deeply, let's see some publications!" I got the impression this was fairly quickly put together because he was told "You haven't produced a book and we pay you to think. If you died it would all be lost, so let's have a book out of it!" I get the impression he's very pleased with his photograph on the cover, but I'm not sure his photography skills match his philosophical skills.

Many people would disagree with the latter claim, but I'll mention something that I found curious which is Parfit very much wanted to use that photograph for the cover of the book. The publisher didn't want it because the publisher thought, there's no connection between this photograph and the book. Apparently, Parfit added some stuff to the book just so there would be an excuse for using this cover. I am no photo critic, so I cannot evaluate the quality of Derek's photos, but the impression that I get from stories that I've heard, is that he at least put as much effort and enthusiasm into, and was as obsessive about his photography as he was about his philosophy.

Did you meet him when he was at Rutgers?

Yes, I met him and he was one of my advisors for my dissertation. So, he was co-teaching a class at Rutgers. He was only there for a small fraction, maybe a half a semester or less, something like a quarter of a year. So he was only there briefly, but I did meet him on many occasions. We had a lot of conversations. At that time, he was working on *On What Matters*, so most of our discussions were discussions of his manuscripts for *On What Matters*.

Wow. That must've been great for a graduate student to be asked for input on something like that. Must've been very flattering.

Yeah. He didn't just ask for input; we had long e-mail exchanges, where I'd read a chapter of his, I'd send him comments and for each of my comments, he'd break them down and give responses to each of them. We had long back-and-forths on a number of his chapters.

Frank Arntzenius was at USC when I was there and then he went to Rutgers. So, presumably he was there when you went to grad school at Rutgers. He has since gone to Oxford. Same as Parfit and retired early, I believe. But, he said he found

“ Parfit very much wanted to use that photograph for the cover of *Reasons and Persons*. The publisher didn't want it because the publisher thought, there's no connection between this photograph and the book. Apparently, Parfit added some stuff to the book just so there would be an excuse for using this cover. ”

Parfit a bit much just because he was all philosophy. Frank is one of those guys who can handle the math. I took a philosophy of quantum mechanics from him and it was a huge mistake... Frank is a ferociously intelligent person, but he said he found Parfit a little bit exhausting, because it was just philosophy all the time. Maybe if he had been interested in photography, that wouldn't have been true!

I can somewhat sympathize with that. So, I should say, I took a course from Frank Arntzenius. Brilliant guy, one of the greatest teachers I've ever had. In the case of Derek, he had interests outside of philosophy. I've mentioned photography. He was interested in the arts more generally. He was interested in music. So, he had other interests besides philosophy. I think that he wasn't much for small talk, so when he was talking with other philosophers, he wanted the conversations to be so that the focus was mainly on philosophy, but it wasn't exclusively on philosophy.

If you were button-holed in a pub and somebody said to you, "Why should I care about Derek Parfit?" what would you say are his major contributions to philosophy? Perhaps this has happened to you. Perhaps this happens to you all the time.

I guess I would break down his contributions into a number of areas. First, there are topics within philosophy in which he made hugely important contributions. And I think that population ethics and personal identity are probably the two biggest areas. So, I think he's probably the most important philosopher to have discussed those. I think he's made the most important contributions of any philosopher, probably, to population ethics and I think probably also to personal identity. At least with respect to its moral and practical/prudential implications. Not so much the metaphysics of personal identity, but the practical implications of issues about personal identity. There are certain topics that I think are very important within philosophy in which he's made very important contributions. I also think that, even though he wasn't teaching a lot of classes at All Souls, he was doing a lot of advising and I think he created sort of generations of philosophers, very important philosophers, whom he benefited very much in an advisory role. So, that's one of his contributions. Finally, I think that, as I mentioned, one of the things I find most impressive about him is his style of philosophy, his approach to philosophy, the style of his argumentation. I feel that has had an influence far beyond those people who were lucky enough to have him as an advisor.

If you go through a hiring process and you get work submitted by tons and tons of graduate students, you do notice that there are fashions in philosophical writing. I noticed one recently, where everybody was doing this: they would introduce an example by using the name. So, every example has a name. A catchy name like "Drunk Farmer" or whatever, and that's certainly something you see in Parfit. All of his examples are punchy and they have a name, which is certainly refreshing if you've had to read someone like John Stuart Mill or John Rawls before that. As much as I love both those philosophers, the examples are few and far between.

I think that while Parfit's influence on philosophical writing has been primarily positive, I don't think it's been entirely positive. I think that Parfit liked to write as crisply as possible, with as little fluff as possible, as efficiently as possible and I think that some

“ Even though he wasn't teaching a lot of classes at All Souls, he was doing a lot of advising and I think he created sort of generations of philosophers, very important philosophers, whom he benefited very much in an advisory role. So, that's one of his contributions. ”

Philosophical Profiles

people who try to copy this style end up writing in ways that are less clear than if they were to write longer sentences and add more clarifications.

You say that his major contributions are in population ethics and personal identity. Could you say a little bit about personal identity for someone who doesn't know what the topic is about? Could you give a quick summary of the topic, and Parfit's contribution, and why it was considered so earth-shaking at the time.

Personal identity refers to the idea of being the same person over time. So, I should explain what is meant by "the same person." I take it anyone who is listening to this interview was once a child. You will obviously have changed in lots and lots of ways since you were a child. You will presumably be physically larger. You will have more knowledge. Your hair may be grayer and thinner, etc. So you're not the same person as you were as a child, in the sense of being qualitatively the same. You're qualitatively very different in a whole bunch of ways. But, you are the same individual. It is you who were once a child. And you who will one day be older than you are now. So, over your entire lifetime you remain one individual despite all these qualitative changes.

Now, personal identity, we normally think is very important in a number of ways. One respect in which we normally think that personal identity is important is with respect to self-interest. One of the main themes of Parfit's *Reasons and Persons* is to challenge the self-interest theory of rationality and reasons for action. Parfit holds that traditionally, for thousands of years, people thought the rational thing to do, or what we have most reason to do is whatever will be most in our self interest in the long run.

Do you think that Parfit got to the topic of personal identity because he was interested in rationality first? Because his earliest major work, where he first made a splash, was "Personal Identity," in *The Philosophical Review* in 1971. But, of course, the topic of personal identity is not what begins *Reasons and Persons*. You get to that in Part 3. Do you have an idea of the order he came to it? Because the way you're introducing it, it looks like: "rationality tells us to be self-interested which raises the question, what is the self?"

I don't know. I can't say authoritatively. If I were to guess, I would guess that these interests in the nature of rationality and reasons were always in the background of his interest in personal identity. I think a lot of these issues and questions emerged from his reading of Sidgwick. He regarded Sidgwick's *Methods of Ethics* as the most important work in ethics, and I think it was working through some of the ideas and arguments there that inspired a lot of his ideas. I think it may have been the emphasis that Sidgwick placed on self-interest. Sidgwick famously thought there was this conflict between morality and rationality, since morality taught us to maximize the welfare of everyone whereas rationality tells us to maximize our own self-interest. I think it's in part because of this emphasis on self-interest in Sidgwick that it became an important topic within *Reasons and Persons*. I wouldn't be surprised if that was always in the background of his interest in personal identity, although I can't say for sure.

“ I think that while Parfit's influence on philosophical writing has been primarily positive, I don't think it's been entirely positive. I think that Parfit liked to write as crisply as possible, with as little fluff as possible, as efficiently as possible and I think that some people who try to copy this style end up writing in ways that are less clear than if they were to write longer sentences and add more clarifications.”

So, according to Parfit, for thousands of years, the self-interest theory has been the standard theory of practical reasons or rationality. I'm not sure if that's true today. I think most philosophers today don't think that all of our reasons for action come from self-interest or that what's rational for us to do is whatever maximizes our own self-interest. But, most people think that self-interest is at least one source of reasons for action. We have at least some reason to promote how well our life goes in the future. Insofar as that's true, personal identity becomes relevant. There is a whole literature on personal identity; people coming up with different accounts of when we're the same person over time. These accounts will have different kinds of implications. One theory of personal identity, for example, will be physical. On one theory, what it is to be the same person over time is for there to be something like the continuity of the body, or perhaps the continuity of the most central part of the body, namely the brain. That's one view.

Another view says that what matters is psychological continuity. What matters is a continuous sequence of mental states that stand in the right kinds of causal relation to another regardless of whether there is a continuously existing physical substrate such as a human body or a human brain. These kinds of alternative theories will have different implications in, for example, a teletransportation case. Take a case like, some people might be familiar with this from *Star Trek*. You're on board the Starship Enterprise. There is a device that is going to scan your body and figure out where every particle in your body is. Then, it's going to annihilate every particle on the platform on the Starship Enterprise, and it's going to create an exact duplicate on the planet's surface. Molecule for molecule identical that's going to have all the same memories, all the same beliefs, desires, intentions, etc., and is going to carry on like nothing had happened. That would be a case of teletransportation as it exists in the philosophical literature. I'll mention that's not how it works on *Star Trek*. In *Star Trek*, the actual particles move. But, forget that.

Is that canon? I didn't know that.

I believe so. But, let's imagine the particles are all annihilated and an exact duplicate is created on the planet's surface. Now, you'll have a case of philosophical teletransportation. On some theories of personal identity in that kind of case, you will continue to exist and you will exist on the planet's surface. That would be true on something like a psychological continuity theory of personal identity. Whereas, on a physical continuity theory of personal identity, perhaps you would not continue to exist because your body and brain has been destroyed and the body and brain on the planet, is a new body and brain.

It would just be a replica who was under the mistaken impression that it's you because it is a replica that has been implanted with the memories of you. Like in *Blade Runner*, there's the character of Rachael who thinks that she has grown up from a child, but she is in fact a replicant and she's been given the memories of somebody who died. That's what the Kirk on the planet's surface would be an entirely new one and actually Captain Kirk has been destroyed. That's according to the physical theory.

“ He regarded Sidgwick's *Methods of Ethics* as the most important work in ethics, and I think it was working through some of the ideas and arguments there that inspired a lot of his ideas... I think it's in part because of this emphasis on self-interest in Sidgwick that it became an important topic within *Reasons and Persons*. I wouldn't be surprised if that was always in the background of his interest in personal identity ”

Philosophical Profiles

Exactly right. The individual on the planet's surface would think, "five minutes ago I was onboard the Starship Enterprise," but that would be a false belief because as a matter of fact someone else was on the Starship Enterprise five minutes ago.

Let's suppose you are on board the Starship Enterprise and Klingons are chasing after you. You're endangered of being shot up by the Klingons and you have the choice, do you stay and fight the Klingons or do you enter the teletransporter where this will happen and an exact duplicate of you, or perhaps you, will emerge on the planet's surface? If you think that self-interest matters, then you would think, "in order for me to make this decision, I need to know what the correct theory of personal identity is." Because if the psychological continuity theory is correct, then you should transport yourself to the planet's surface because then you will survive and continue to exist, and have many more years of presumably positive life experiences. If you stay behind, there's a good chance you'll be caught by the Klingons.

By contrast, if you hold the physical theory, you would think it would be suicide to get on the teletransporter. It will simply mean that you will cease to exist. Somebody else might have a good life, but that won't be you. Better to take your chances fighting the Klingons than to engage in something that would be suicide.

These views will have diverging implications if we assume that self-interest matters. Each one of us has special reason to care about he or she will fare in the future. I'm mentioning self-interest as one kind of case where personal identity matters, but there are other kinds of cases. For example, let's take obligations. Let us suppose that you promise to return someone's money that you borrowed from them. Now, you enter the teletransporter. Teletransportation occurs. There is somebody on the planet's surface holding onto that money. Do they have an obligation to return that money. Well, if they, themselves promised to return the money, then they have a moral obligation, but if somebody else promised the money would be returned, it's not obvious they have a moral obligation.

I'll mention one other example which is desert. Many people think that if you do a very bad thing, then you may be deserving of punishment. Let's suppose Kirk kills Spock and then the teletransportation occurs.

Good God, man! Not even in fun!

Then, we have a Kirk-like being who appears on the planet's surface. Would it be appropriate or fitting to punish this individual for the murder of Spock? Well, if it's the same individual, then yes. This person killed Spock, therefore he should be punished. But if it's not the same individual, not so obvious. So here's a bunch of ethical domains. I've mentioned three of these, where at least to common sense ethics, personal identity is central. What Parfit does is, he creates cases in which many people will give up the intuition that personal identity matters. He creates cases where many people will hold, "you know what, there's no personal identity in this case. The person who exists after an operation is not the same person who existed before the operation." Still, the person

who exists before the operation has just as much reason to care about the person who exists after the operation as they would if it was the same person. So, what he argues is, our idea that personal identity matters to self-interest is a mistake. Something else matters to self-interest. Maybe we shouldn't even call it self-interest.

There is something we've traditionally thought of as self-interest. We've traditionally thought of an interest we should have in our future selves, but when we explore these hypothetical cases we realize what really matters is not the fact that the future being is our future self, but rather the fact that this future being stands in some other kind of relation to our present self. A relation other than identity.

Much of what he's arguing for in the personal identity discussion is that what really matters from a practical point of view--I've been focusing on self-interest, but similar things could be said about other ways in which personal identity is thought to matter--is not itself self-interest, but something else.

That was great. Hence, whenever people talk about Parfit they can't help but use the phrase "what matters," because, as you said, when you can no longer say "self-interest," what do you say? There's talk about "what matters" and what's initially confusing, if you give students his 1971 article, is that he does two separate things in it. The first one, he defends a Lockean view of personal identity, which is the psychological continuity one, against famous criticisms of Lockean contemporaries or people who came shortly after him. He patches it up, this psychological continuity theory. But then the more important and more radical thing he does, is he says, "But we shouldn't care about personal identity anyway." It looks as if on one stage he's being a Lockean, but then he pulls the rug out from under you and says, "But you shouldn't care anyway."

Some of the things that he talks about, I think are particularly mind-blowing to even casual readers, even philosophy undergraduates who are not happy that they have to take a philosophy class. For example, where he talks about split brains. Would you like to say something about that?

Sure, that's fun. I think it's true that, the teletransportation cases I've discussed are obviously science-fiction cases, we don't have that technology, we may never have that technology. The split brains cases: he discusses what are called "fission" cases. A case where someone has the two hemispheres of their brains divided in two, one half put into one person's body and the other half put into someone else's body. These kinds of science-fiction scenario play a big role in many of his arguments, at least concerning personal identity. Obviously, some students, certainly many non-philosophers are turned off by this appeal to science-fiction scenarios or realistic scenarios. I think that there are a couple of ways of fleshing out this concern.

One way of fleshing out this concern would be to say something like, these cases are so remote or so fantastical, or etc., that we don't know what to say about them. We

“ There is something we've traditionally thought of as self-interest. We've traditionally thought of an interest we should have in our future selves, but when we explore these hypothetical cases we realize what really matters is not the fact that the future being is our future self, but rather the fact that this future being stands in some other kind of relation to our present self. A relation other than identity. ”

Philosophical Profiles

have no idea what to say about them. And, I think that that would be a mistake. What makes his arguments powerful is that most people think they know exactly what to say about these cases. If he was to come up with hypothetical possibilities where we had no intuitions about these cases, then they would obviously be weak arguments. But, he comes up with hypothetical cases where we have extremely clear intuitions about them, where almost everyone will agree on what judgment we should make about these cases.

I think there's a big difference though between the teleporting examples, which are pure science-fiction and as you said, there's dispute as to whether or not this is even in theory possible. I think I read somewhere that it would take more energy than there is in the universe to do teletransportation. But, the split-brain cases were actually supposed to be influenced by real cases where extreme epileptics had their corpus callosum severed. This isn't done anymore as far as I know, but these poor individuals were studied just incessantly, because they appeared to show two seats of consciousness. There's the famous case, of course, where you can fit people with these glasses so that what one hemisphere is seeing, the other one can't see, and vice versa. So you show different pictures to the two different hemispheres, and they appear to be able to think independently of each other. I think he thought that this would work for him in breaking down this intuition we have that we are this unified, persisting, point of view. I think that those examples are, if anything, less science-fictiony and therefore more kind of mind-blowing because students are also learning that this happened. Then, they're ready to have lessons drawn about it.

I agree. I was going to say that, one source of concern about hypothetical examples might just be that we don't know what to say about the hypothetical cases. I don't think that's true in the case of Parfit, or most of Parfit's cases. Another concern you might have is well, these issues arise only in these bizarre hypothetical cases, therefore they have no bearing on the real-world cases, which is what everybody cares about. I would resist that objection in a couple of ways.

First, let us suppose that as a matter of fact, what matters is not self-interest, but some other relation, "R," that's other than self-interest. Let's suppose that as a matter of fact, R is always co-extensive with the identity relation, so we always stand in relation-R towards and only towards our future selves.

Where relation-R is that the future self remembers being us and that our intentions are carried out by our future self. Something like that. It's some sum of those kind of things.

Relation-R sounds very mysterious, but it's a more sophisticated version of the basic Lockean idea that there's this connection of consciousness.

True. That's Parfit's version of relation R. I'm trying to leave open, how exactly we define this. Let's suppose we found that what we should care about are future person stages who stand in some kind of relation, call that "R," to our present person stage. That might be understood in terms of psychological continuity or something else. Let's suppose, as a matter of fact, because there are no teletransportation cases, because there are

no fission cases, etc., we always stand in that relation precisely to our own future self. To the stages of our own future self. In that case, the theory that says that self-interest is what matters, would have the same practical implications as the theory that says that what matters is relation R. So, in some sense, this discussion might not have any practical significance; still it would have a philosophical significance, even in those cases, because insofar as we are philosophically minded ethicists, we're concerned to know not just what to do, but why we should do it.

We wouldn't be satisfied if someone just gave us a set of instructions: "here are all the things you should do over the course of your life." We want to know *why* we should do it. The explanation of why we should do the things that we should do is something that matters to us. So, if the reason we should be promoting the welfare of certain future stages is not the fact that these future stages are ours, but rather the fact that these are future stages to which we stand in a certain kind of relation, that's a different explanation for why we should care about these future person stages and why we should behave in ways that affect their interests. So I think even if the practical implications of an alternative view would only differ in hypothetical science-fiction scenarios, the explanation for why we should do what we should do will differ even in the real world.

I'll mention one other thing: Parfit will take a hypothetical scenario like a teletransportation case or like a fission case to give an argument against a particular theory, let's say the self-interest theory. Once we begin to see that the self-interest theory may not be as obvious as we thought it was, once we begin to question the idea that we have a general reason to care about our future selves, because they are our future selves, then this may open the way, or at least open our minds to other alternatives which we might have initially excluded because we thought they were incompatible with what we thought was self-evidently true. But now we no longer think this theory is self-evidently true. Now, we have a more open-minded attitude towards alternatives and these alternatives might actually have practical implications in the real world.

I'll give you an example. Once we think it's no longer self-evident that the rational thing to do is always to maximize our total wealth over the course of our lives, and once we think, "hey what might matter is things like psychological continuity," then it allows for us to think, "well wait a minute, we have different degrees of psychological continuity to different future person stages." So, if psychological continuity is what matters, then maybe degrees of psychological continuity can matter, so maybe we could have more reason, for example, to care about our near future than about our far future if we're more psychologically connected to our near future than to our far future. It's well known that, as a matter of fact, people care more about their near future than about their far future. Somebody will take one cookie now rather than two cookies ten years from now. There's various possible explanations for this. One view is: that's irrational. The traditional view is that we do this and that's a failing of ours. We should give equal weight to all parts of our lives, but insofar as we are weak-willed and irrational, we may care more about the near future.

“We wouldn't be satisfied if someone just gave us a set of instructions: "here are all the things you should do over the course of your life." We want to know *why* we should do it. The explanation of why we should do the things that we should do is something that matters to us.”

Philosophical Profiles

A second explanation would be, well that might be rational but it's because of epistemic fact. Because we can never predict the far future, it's epistemically rational, in some sense, for us to ignore or discount, or some such thing.

A third possibility is, it could be rational even in cases where there is no such uncertainty concerning the further future. Even in the case where we know we're in the position to choose between one cookie in the near future and two cookies ten years from now, we might have most reason to take the one cookie in the near future because what matters is something like psychological continuity and we're going to stand in that relation more to our near future selves than our distant future selves. I guess these would be my three kinds of response to that's just science fiction; that could never happen in real life.

First point, the fact that it's science-fiction, or that it may be scientifically impossible, or metaphysically impossible doesn't mean we don't have clear intuitions about that case. And if we have clear intuitions about that case and we want to hold onto clear intuitions about that case, and if our preceding theories imply something that contradicts those intuitions, then we've got to change our preceding theories.

Second point, changing our preceding theories will affect our self-understanding even if it has no practical implications.

Third point, by rejecting views that appear to us to be self-evident, we may open our minds to further possibilities that may have practical implications even in the real world.

I was listening to the radio the other day and a woman came on who was a financial planner. She said that she'd had a terrible experience in 2008. She'd been starting her own business and the housing crisis started, and it crashed, and she had to come back and live with her parents, age 30 or something. But she worked her way out of it and now she was a motivational speaker on financial planning. She said, "here's the way I think of it: I think of the 70-year-old me, as this woman, I'm going to call her something" - she gave her a name like Zola or something - "every time I start to do something, I imagine Zola getting mad at me for doing something that was detrimental." I thought, Parfit would have so much to say about this, but it's interesting that she viewed her future self as a different person, and yet as a different person to whom she had obligations. There are all kinds of interesting implications of Parfit's view, of course.

Parfit said, himself, that his view is pretty much Buddhist. Of course, he said the same thing about Hume, who I don't think would regard himself as Buddhist, although I don't know. There's probably a burgeoning school of people who say that Hume is influenced by Buddhists. Has Parfit made you a Buddhist?

No. The quick answer to your question is no.

“ The fact that it's science-fiction, or that it may be scientifically impossible, or metaphysically impossible doesn't mean we don't have clear intuitions about that case. And if we have clear intuitions about that case and we want to hold onto clear intuitions about that case, and if our preceding theories imply something that contradicts those intuitions, then we've got to change our preceding theories. ”

Let's distinguish between a couple of claims about Buddhism. One of which, I think is simply false. And one of which, I think is both false and repugnant.

The false view which I don't even think Parfit holds, is that there is no personal identity. So, I take it, the Buddhist view was to reject personal identity. I think they do so on the basis of what I take to be a rather poor argument. Although it's an argument that Parfit quotes. He has an appendix, I believe in *Reasons and Persons*, where he gives the argument. The argument seems to begin with, like, a cart. The argument is well, there is no cart. There is just the wheels and you list the various parts. That's all there is. There's just the parts of the cart. There isn't, in addition, the cart. I think the mistake is using the words "in addition." There is no cart, in addition to all those parts, but there is a cart that consists of all those parts. Then, they will hold a similar view is true of persons. Well, in the case of the person, there's all the physical parts. There's all the flesh and sinews and bones. And then there's the mental parts: there's the experiences, beliefs, desires, etc. But, there isn't over and above that, a person. Therefore, there's no person. So, in the case of the Buddha, I think he was responding to a view that in addition to the body and all the physical parts and all the mental parts, there was an Atman. A capital S self, so I think that the Buddha was rejecting the idea that there's some special self that is separable from all of the parts of the person. But, I think that the argument overreached and he came to the conclusion there is no person, there is just the parts. And so we do not really continue to exist. I don't think that even Parfit holds that view. I think Parfit does believe there are persons, which are individuals that continue to exist over time. I'll mention there was a school of Buddhist thought. They were called the Pudgalavādans. Pudgala means something like person. The Pudgalavādans held that there really are persons that endure over time, but perhaps, there's no Atman. There's no special, separable soul, but there are persons which consist in these parts standing in the right kinds of relation. There was a Buddhist tradition which is compatible with Parfit's view, with the view that I hold, that I think many people today hold, which is that there are objects that endure over time and persons are among those things. I think the Buddha made a mistake in rejecting the idea of continually existing persons. I think it might be the same mistake that you can find in certain pre-Socratic philosophers who say you can never step in the same river twice, in fact you can't even step in the same river once. Going from the qualitative change in the river to the qualitative change in the person stepping in the river to the conclusion that there's no continually existing thing over time. That's the claim that Buddhists make about personal identity, which I think is simply false, although not repugnant. In fact, it might even be a positive view. Parfit suggested that the conclusion that he had reached about personal identity and self-interest, they were liberating.

The glass tube comment. I never quite got that image. Maybe you could explain.

I take it, the idea is, he used to think that all that he should care about is his own future self, but now he realizes that this obsessive concern with one's own future self was misguided. That what matters about his own future self is certain kinds of causal relations which he can stand to them. But you can stand in certain important kinds of causal relations to other people as well, so it sort of broadened his perspective. I take

“ I think that the Buddha was rejecting the idea that there's some special self that is separable from all of the parts of the person. But, I think that the argument overreached and he came to the conclusion there is no person, there is just the parts. And so we do not really continue to exist. I don't think that even Parfit holds that view. ”

Philosophical Profiles

it something similar might be true for the Buddhist. Compassion is one of the highest virtues in Buddhism. The idea that we should be alleviating the suffering of others. I wouldn't be surprised if ceasing to believe in personal identity might make people more compassionate. If you think there's no special relation in which I stand to my future self, this might diminish the difference between the attitude you have toward your own future self and the attitude you have towards other people's future selves.

One implication of not caring about your future self might be just complete apathy and nihilism - not caring about anything. That would be, presumably an unfortunate result. But, an alternative, possible result is caring more about others, becoming a more compassionate self. I think the Buddhist view that the self doesn't exist, there's no such thing as enduring persons, while it's a false view, it may well be a positive view.

My main reason for rejecting Buddhism is not this false view about personal identity, but rather with the first of the Four Noble Truths. So, the first of the Four Noble Truths is, "everything is suffering." The second Noble Truth is "craving is the cause of suffering." The third Noble Truth is "there is a way to the cessation of suffering." The final Truth is "The Eightfold Path is the way to the cessation of suffering." I think the mistake they make is the very first one. Everything is suffering. Because if you held that, you would have to hold the following: if there were a button that you could press that would lead to the annihilation of the universe, and if it would be successful, then you should press the button. I have pressed Buddhists with this. I said, "let's suppose there was a button, and if you press this button, the entire universe and everyone in it will be annihilated. Should you press the button?" They will say, "oh no, you shouldn't because if you press the button it simply wouldn't work, and you'd be reincarnated, and blah blah blah." They reject the premise.

The problem I always had was if you were going to say that we're trapped in the cycle of being and being reincarnated, don't you have to believe in a soul theory of personal identity or something like that? Because otherwise, what is it that's persisting?

They have a causal theory of reincarnation. The idea is that even in your ordinary life, there's no person that continues to exist, there's simply causal relations between different temporal slices. In a case of reincarnation, there's a similarity between the causal relations. Take the last slice of what we would ordinarily call your life. This can stand in a similar causal relation to what we would ordinarily call the first slice of the next life as the various slices within your life stand to one another. So, it's not really you who will exist in the next life, but it's as much you as what we would ordinarily call your future self is.

But, if you're not going to believe in persisting people, because presumably one response a Buddhist could make, or one response somebody could make who believed that existence was suffering, that you still shouldn't press the button is because there's some intrinsic value to existing even if it's suffering. Like it's better

“ I have pressed Buddhists with this. I said, “let's suppose there was a button, and if you press this button, the entire universe and everyone in it will be annihilated. Should you press the button?” They will say, “oh no, you shouldn't because if you press the button it simply wouldn't work, and you'd be reincarnated, and blah blah blah.” They reject the premise. ”

to exist just to exist and to suffer is better than not to exist at all. That would not be acceptable to any hedonist, but presumably you could say something like that.

Certainly, that wouldn't be the Buddhist view. The Buddhist view is not we should carry on existing because existence is desirable even if it consists of suffering. Moreover, it's hard for me to see how that could be a plausible view. I could imagine someone thinking, it's worthwhile continuing to exist, even if all of your experiences will be painful.

Schopenhauer and Samuel Beckett didn't commit suicide did they?

That's true, but I could imagine someone thinking, one should continue to exist even if all of one's future experiences will be painful, because existence has some further positive property. Like there's a certain dignity in continuing to exist. You'd have to at least believe in some kind of a positive property. I believe that that would be against the spirit of the idea that everything is suffering. Basically, I regard the Buddhist view as a kind of negative monism. In a sense it's the converse of the following. There was a view that you find in the Christian tradition, which says, everything is made by God. Everything that God makes is good, therefore everything that exists must be good, therefore there's no positive evil- what we're calling evil is merely the absence of goodness. Which is itself a kind of absence of reality or absence of being. That would be a kind of positive monism. The converse view is a kind of negative monism. There's nothing of positive value. There's only things of negative value. There's only things that consist of suffering that have negative value. Even things like happiness and pleasure do not have intrinsic value, they are simply reductions in this negative thing of suffering. I believe that the Buddhists, or perhaps, contemporaries of the Buddhists had an analogy of someone whose skin was all covered in welts and if they were to immerse themselves in a bath, they might get a temporary relief from the pain. All pleasure is like that. It's just a temporary relief from or reduction of the suffering that's essential to existence.

“In a sense you might say that Parfit's work on population ethics was more pioneering even than his work on personal identity.”

What about population ethics? You said if anything Parfit has just made a larger contribution to this because [on the issue of personal identity] Parfit was part of the tradition of which Locke was the most influential member, but with population ethics, there's no Locke before him, really is there?

I agree. In a sense you might say that Parfit's work on population ethics was more pioneering even than his work on personal identity.

I'll mention a few things. One contribution that he makes is to provide what I think many regard as perhaps the strongest objection, or the strongest counter example, to utilitarianism, as a theory of value. Utilitarianism is the view that what we want to do is maximize the welfare of all beings. So how good an outcome is is simply the sum of how well off all the individuals are. Many people reject utilitarianism as a theory of rightness and wrongness, as a theory of what one is morally obligated to do. There's lots of ways in which one might see utilitarianism as problematic as a theory of what we're morally required to do. But a person might do that while accepting a utilitarian theory of value, while holding that the best outcome is the outcome that has the best overall

Philosophical Profiles

sum of happiness for all individuals. One thing that Parfit did was he gave what many regard as a knockdown objection to that, which is called The Repugnant Conclusion. He imagines two scenarios. In one scenario, there is a very large population of individuals. You might imagine there's 10 billion individuals. They each have an absolutely wonderful life. Take the best life any human being has ever had: each of them has a life that's at least as good as that, let's say. You've got a large population, everyone has an absolutely wonderful life. That's one possibility.

Here's a second possibility. You have a world where everyone has a life that is barely worth living. How should we imagine a life that's barely worth living? There are a couple of ways we can imagine this. One way we could imagine this is the bad things in life almost perfectly outweigh the good things in life. For every positive experience they have, they have a negative experience that is almost exactly equal. Although, I would prefer to think of it as a life which simply has very little positive experience of any kind. We could imagine, for example, someone who is perhaps comatose for the vast majority of their life and for a very brief moment, they have an extremely mildly pleasurable experience. That's one way to think about it. Or, perhaps they have the kind of experience that you'd normally attribute to a non-human animal, perhaps an oyster. There's various ways of imagining lives that are barely worth living. There's very, very little good in these lives. One way to imagine them is, if they were to stub their toe one more time, their life would no longer be living. If it was even slightly worse, they'd have no reason to be glad that they'd lived at all. Whereas the first individuals, if we were to assign a number to how good their lives were, we might say each individual has a welfare level of, let's say, a thousand or a million or some such thing. In the second case, if we were to assign a number to how much value was in one of these lives, it would be a tiny number. Just slightly above zero 0.001. But now we make the population large enough. So rather than having 10 billion people as there were in the first scenario, now we have a googol of people, or a googolplex of people, or one hundred to the one hundredth power of people, or whatever it might be, some enormous number of individuals. If the number of individuals is great enough, and if the value of each of these lives, the amount of welfare or the amount of happiness is positive, greater than zero, well if you multiply a quantity greater than zero by a large enough number, you will get a product that exceeds any given number. Whatever the finite value was of the first world, we can construct the second world with a large enough population that's going to be assigned a greater value by the utilitarian way of calculating utility.

But, most people think if they have the choice between a world of 10 billion people who all had absolutely wonderful lives, the best lives that anyone has ever lived, or an enormous population of people, each of whose lives is barely worth living, most people would think it is better to choose the first than choose the second. Most people regard this as a reason for rejecting the utilitarian way of understanding the value of an outcome.

One thing that Parfit does is he gives this wonderful thing where part of the conclusion is what many people regard as a devastating objection to perhaps the simplest and most straightforward and, before then, kind of widely accept view of how you calculate how good an outcome is overall involving an entire population of people. But, he doesn't

stop there. The next thing he does is, he says “OK if we’re going to reject the total view. We’re going to reject the view that says how good an outcome is is simply equal to the total of how well each individual fairs, then what alternative view should one accept?” He considers a bunch of alternatives. I won’t go through all the alternatives he considers. I’ll just mention one.

The most obvious alternative, and I know from experience when you talk to people about this and when you give them part of the conclusion, the alternative that most people propose is the average view. The average view says that what we want to maximize is not the total welfare of all individuals, but rather the average welfare of all individuals. The reason the first scenario where you have 10 billion people who all have wonderful lives is better than the second scenario where you have an enormous number of people who all have lives worth living, is that the average level of welfare is better in the first scenario. Parfit shows that this view, like the total view, likewise has unacceptable implications. In fact, it has implications that are even worse.

Here would be an example of a problematic implication of the average view. Let us suppose you have a world where you have Dante’s ten levels of Hell, and let us suppose that each level of Hell has greater and greater levels of torture. If you’re in the first level of Hell, you have mild torture, if you’re in the second level of Hell, you have slightly worse torture, and by the time you’re in the ninth level of Hell, you have absolutely excruciating torture. But the eight is almost as bad. The eight is absolutely excruciating torture, just not quite as excruciating as the ninth level of Hell. Let’s suppose you had a world where everyone was in the ninth level of Hell. Not through any fault of their own. Let’s imagine they’re morally wonderful, totally innocent people, but somehow Satan has captured all their souls and put them all into the ninth level of Hell.

This is not painting God in a good light.

Somebody approaches Satan and says, “you know, this is very unfair. All these poor individuals that are completely innocent, who are all being tortured in this ninth level of Hell. There’s got to be something you could do. There’s got to be some way you could improve the situation. Would you consider perhaps releasing them from the ninth level of Hell and allowing them to go somewhere that isn’t so bad?” Satan says “no I will not do that, but I will offer you an alternative. I will create some more individuals and I’ll put them in the eighth level of Hell and they will all experience nothing but extraordinary torture for the entirety of their lives. Will that help?”

If you hold the average view, you’re going to have to accept that offer and say, “yes please, that would definitely improve the situation,” because now the average level of welfare has increased, but obviously that would not improve this situation.

You have these two fairly commonly accepted views of how to measure the overall value of an outcome. One says the value of the outcome is proportional to the total welfare of all the individuals. Another one says it’s the average welfare. He gives very strong objections to each of these. Then he considers pretty much every other view that

““ The reason the first scenario where you have 10 billion people who all have wonderful lives is better than the second scenario where you have an enormous number of people who all have lives worth living, is that the average level of welfare is better in the first scenario. Parfit shows that this view, like the total view, likewise has unacceptable implications. In fact, it has implications that are even worse. ””

Philosophical Profiles

anyone might propose in light of the failures of these two. For every one of these, he gives a very powerful objection.

He doesn't formerly present this as an impossibility proof, but it was at least the beginnings of an impossibility proof. What he showed was basically there's no theory of population ethics. There's no theory of how to evaluate the overall value of an outcome involving various individuals of various welfare levels that doesn't have extremely counterintuitive implications in certain scenarios. What he proves is we've got to bite a very big bullet no matter what view we adopt.

I've had this theory of philosophy that the great philosophers, or at least the enduring contributions to philosophy are negative. This is why Hume, of course, is my favorite philosopher, because there are philosophers who are system-builders and then there are philosophers who are system-destroyers, or at least people who make it their job to come up with problems. I think systems never last, but problems tend to endure. Would you say Parfit was a system-builder or a problem maker?

I think he was both. I think that he was perhaps more successful as someone who raises problems than as someone who proposed solutions or built systems. I think that perhaps the biggest contributions of *Reasons and Persons* could be seen as negative. I don't quite like the word negative, because he doesn't just show the views are false: he also raised questions that other people hadn't raised before and thus prepared the way for a whole bunch of very fruitful, positive discussions.

And besides, if a view is false, it's not a negative thing to point that out.

I agree with that, but I also say his contribution wasn't just saying here are some existing views that are false. I think he also sort of laid out the conceptual space in which to think about views, and allowed for a whole bunch of views to be discussed and explored that wouldn't have been discussed and explored had it not been for his work. That goes beyond some of the purely negative, but it's not quite the level of system-building. I think, for example, in his more recent work, in *On What Matters*, there is some system-work attempted. In *On What Matters*, he discusses various traditions in moral philosophy that had been thought to be in opposition. So, he discusses Kantian views of moral philosophy, and consequentialist views of moral philosophy, and contractualist views of moral philosophy, and he suggests that the very best versions of each actually come to coincide. He suggests that there is what you might call like a triple theory, which is the best version of each, which all agree in their practical implications. So, I would call this a kind of system-building, although I don't think the contributions he makes here are as great as the contributions he made in *Reasons and Persons*, for example.

Yeah, *On What Matters* grew out of a famous unpublished work of his which is called *Climbing the Mountain*, which comes from a quote from Bertrand Russell's father, wasn't it? Or at least Bertrand Russell's father attributed to Mill because he'd been speaking with Mill and Mill said, "you should never be discouraged when people disagree with you because everybody shares this idea that morality is important.

“What he showed was basically there's no theory of population ethics. There's no theory of how to evaluate the overall value of an outcome involving various individuals of various welfare levels that doesn't have extremely counterintuitive implications in certain scenarios. What he proves is we've got to bite a very big bullet no matter what view we adopt.”

But, if you have different moral intuitions you're at least on the same hill, but on the other side." Isn't that right? So, the "climbing the mountain" is everybody's coming up the same mountain. That speaks to this reconciliation project, which perhaps wasn't very successful.

***On What Matters* isn't as good a title. He should've stayed with *Climbing the Mountain*.**

Who do you think is the most important Parfit critic? Who do you think has done most important sort of poking holes in Parfit, or challenged him the most? In his lifetime, Parfit was arguing against Bernard Williams the most wasn't he? Or at least, they had an exchange. Certainly, they disagreed about most things.

They certainly disagreed. I know that Parfit greatly admired Bernard Williams. I don't know who I would regard as the greatest critics of Parfit. The impression I get is that generally the people who discuss him most are also people who are at least partially sympathetic. I think the people who most delve into Parfit's writings and ideas and arguments, are people who are sympathetic to his approach, sympathetic to his arguments and are maybe proposing friendly amendments.

Patches.

They might go beyond patches. I think the people who have the most negative views about Parfit simply dismiss him and don't interact too much with his ideas.

You yourself seem drawn to big, bold, big-picture philosophizing. For example, I'm looking at a couple of your most cited things. "Rejecting Ethical Deflationism" [Ethics 116, July 2006] for example. I'd like you to talk about that for a little and then maybe "From Teleosemantics to Normative Ethics." Would you regard those as bold, big-picture philosophizing and if so, do you think that's partly the Parfit influence?

In the case of the first one, "Rejecting Ethical Deflationism," that was my first paper. I wrote that when I was in grad school. That came out of my work on my dissertation where Parfit was my advisor, so there's an influence there, definitely. I know that Parfit was fairly sympathetic to the views that I presented in that paper.

In the case of the second paper you mentioned, "From Teleosemantics to Normative Ethics," I think that he would be less sympathetic. He would regard me as a great disappointment as a former student of his for having written this. The reason is that the latter paper is arguing for a kind of naturalism. In "From Teleosemantics to Normative Ethics," I'm suggesting that we can understand wrongness in terms of what makes certain attitudes fitting. So, the idea is that we can understand wrongness in terms of fitting disapproval. The wrong actions are precisely the actions that it's fitting to disapprove of, at least with first approximation. I then argue, we can understand the fittingness of attitudes in terms of the representational content, of these attitudes. The idea is attitudes represent their objects as being a certain way. For example, fear: when

“ In the case of the second paper you mentioned, "From Teleosemantics to Normative Ethics," I think that he would be less sympathetic. He would regard me as a great disappointment as a former student of his for having written this. ”

Philosophical Profiles

we're afraid of something, our fear for the object may represent it as being dangerous. Fear will therefore be fitting, if and only if, the object is dangerous. My third claim, is we use teleosemantics to understand the representational content of certain attitudes. Teleosemantics is an attempt to understand the meaning or content of mental states in terms of their biological functions, which in turn are understood in evolutionary terms. The idea is, if you have an evolutionary story of the function of certain kinds of attitudes - let us take for the example, of fear - the evolutionary story can help you to understand what the representational content is of fear. In this case, it may imply fear represents its object as dangerous, and that will give you the fittingness conditions, that will tell you when it's fitting to fear something. I say something similar is true of disapproval.

If we have a good evolutionary story of the evolutionary function of disapproval, given the proper teleosemantic theory of content, this can tell us what the representational content is of disapproving of something. This, in turn, will tell us when it's fitting for us to disapprove of something, and this in turn, will tell us what is wrong. The idea is that we can ground normativity, and in particular, we can ground moral wrongness in something purely naturalistic namely the evolutionary story behind moral disapproval. Parfit was very much an anti-naturalist. I do not suspect...

So it would be fitting for him to disapprove of this?

I don't know if it would be fitting for him to disapprove, but he definitely would disapprove.

What made him an anti-naturalist and why do you buck the trend?

In terms of what made him an anti-naturalist, there's, like, two ways of reading that. What's the sort of biographical, psychological explanation for what led him to be an anti-naturalist? And, what's the philosophical argument that he took to be justification for his anti-naturalism? In terms of the former, I don't know. I can't tell you the biographical reasons. I think that many people believe that a naturalistic picture would somehow undermine the importance of morality in particular, or value in general. I think that many people regard naturalism as in some way deflationary. That if we give a naturalistic account for example of moral wrongness, somehow this trivializes morality, makes it less important, makes it less worthy of our attention or some such thing. So, I think that could play a psychological role, although I can't say for sure.

The arguments that he gives for non-naturalism, I cannot say that I was ever particularly moved by them. My recollection was, he will quote various naturalists and then say, "this person simply misunderstands the fundamental nature of normativity," "fails to have truly normative concepts" or some such thing. I don't recall any particularly strong arguments against naturalism in his work, but I may not have studied them as closely as I should have.

You feel confident in straying from the true path then?

“ The arguments that he gives for non-naturalism, I cannot say that I was ever particularly moved by them. My recollection was, he will quote various naturalists and then say, “this person simply misunderstands the fundamental nature of normativity,” “fails to have truly normative concepts” or some such thing. I don't recall any particularly strong arguments against naturalism in his work ”

No, I don't feel confident straying from the true path. Although, I do feel confident in straying from *his* path. I might feel confident *de re* in straying from the true path, but not *de dicto*.

I see. Looking at a world post-Parfit. When I was in graduate school... Political Liberalism was just published and just there was this sort of samizdat publication of *Climbing the Mountain* circulating with Parfit, there was this thing called "T2" that was circulating amongst Rawls's students that we were studying. While I think Rawls is a very important writer, I'm not sure of his lasting influence. I think he's one of those writers who is huge in his time, but doesn't necessarily persist. What would you say of Parfit's prospects?

I think that I don't agree with the assessment of Rawls. I think that Rawls had an enormous contribution largely in sort of setting out the conceptual framework in which people thought about political philosophy. I think that that has endured so that lots of people up until and including now still frame a lot of the discussions within political philosophy in sort of Rawlsian terms. I think it's fair to say that I think that Rawls remains the most influential political philosopher, at least twentieth century political philosopher within the field. I think that Rawls now has a bigger influence within political philosophy than probably Parfit has ever had within ethics.

In terms of Parfit's lasting contribution, one difference between Rawls and Parfit, I see Rawls as someone who was a wonderful systematizer but a poor arguer. Parfit might be the converse. I think that Parfit is superior to Rawls both in the quantity or proportion of his work that consists in arguments and in the quality of those arguments. So, argumentatively speaking, I think Parfit is vastly superior, but he wasn't a systematizer the way that Rawls was. I think Rawls sort of pointed the way to generations of political philosophers - here's how you can think about political philosophy. I think Parfit had that kind of impact on some members of the profession, but not nearly as many members of the profession. But I think he will have a lasting impact. I think it may be that there's fewer people writing about Parfit now, or Parfitian themes now than there were for example in the immediate aftermath of the publication of *Reasons and Persons*, but my guess is interest in Parfit will perhaps never reach the level it had at its very height. I think it's reasonably high and suspect it will continue for quite some time to be reasonably high.

Do you think *On What Matters* fell a bit stillborn from the press? That it didn't have anything like the ground-shaking impact of *Reasons and Persons*?

Well, yes and no. I think comparatively speaking, "falling stillborn from the press" would be an exaggeration. There were some people like Peter Singer who regard it as an absolutely extraordinary masterpiece. Some people regarded it very highly. I don't think it had nearly the reception that *Reasons and Persons* had and I think that it's for good reason. I don't think it's as great a book as *Reasons and Persons* by a large margin, but I certainly don't think it's of no value.

“ I see Rawls as someone who was a wonderful systematizer but a poor arguer. Parfit might be the converse. I think that Parfit is superior to Rawls both in the quantity or proportion of his work that consists in arguments and in the quality of those arguments. ”