

Why Should One Care About One's Own Future?

Julien Bugnon and Martine Nida-Rümelin

University of Fribourg

Abstract

Our natural attitude is to care about the fate of a future person in a special manner once we know that this person is none other than ourselves. In the present paper, we defend the rationality of that attitude against Parfit's famous contrary claim that 'identity does not matter'. We argue that it is intrinsically bad for a conscious subject to have negative experiences, and that one therefore has reasons of a special kind (*de se* subject-relative reasons) to avoid having such experiences in the future. Our argument makes crucial use of the distinction between values *simpliciter* and values *for someone*. Values *for someone* are often grounded in the way it is like *for someone* to instantiate a given experiential property. We submit that this claim sheds some light on the deep interconnections between issues about value and issues about the nature of phenomenal consciousness. The disagreement about whether 'identity matters', it turns out, stems from a more fundamental disagreement about the nature of consciousness.

Keywords: Prudential rationality, Value for someone, Personal identity, Derek Parfit, Experiential properties.

1. Introduction

Derek Parfit famously defended a radical attitude towards the title question. One might think that there is an obvious reply to why one should care about one's own future, namely, simply because it is one's own. If I know that I will suffer an unbearable pain tomorrow unless I act now to prevent it, then I have a reason to act because it is bad for me to suffer pain and because I want to avoid what is bad for me. The fact that I am the one who will suffer gives me a reason to prevent the occurrence of an event involving my own suffering—a reason quite different from the reason I have to prevent similar events involving the suffering of other people. It is rational to care about what happens to oneself in the future simply because one wishes to avoid what is bad for oneself.

Such is the 'natural attitude' we wish to defend here. As is well known, Parfit disagreed and developed sophisticated arguments to show that it does not (or rather *should* not) matter for me whether I am the one involved in some future

event. On his view, whenever I am contemplating my own fate, the fact that the future person I am considering is none other than myself can be important only in a derivative manner. It might be important for instance because that future person is in a better position to realize projects I take now to be valuable. However, in itself, the fact that I am the one taking part in some future event should not play any role for my rational evaluation of the future. Nor should it for the way I am motivated to do something about that event.¹

We disagree with Parfit's view and we argue that it is deeply mistaken about the origin of values and the corresponding reasons that rationally guide us in acting in such cases. In this paper, we will not criticize in any detail the reasoning behind Parfit's radical and highly influential view about prudential rationality. We will rather focus on developing an argument in favour of the natural attitude he rejects. The natural attitude appears obvious and it is not hard to see the counterintuitive air of its denial. What is less obvious, however, is how one can voice in a clear manner what makes that denial so counterintuitive. We will put forward a diagnosis. In making explicit a way in which one may justify—*pace* Parfit—the importance of identity, we hope to reveal what underlies the intuition speaking in its favour. We will thus propose an analysis of what underlies the natural attitude and thereby support the claim that it *does* matter for my rational attitude towards future events whether or not I will be myself involved in them.

In order to be able to spell out our proposal precisely, we first need to distinguish between different kinds of values on the one hand, and different kinds of reasons and motivations on the other hand. Such will be the purpose of sections 2 to 4. In sections 5 and 6, we detail and motivate the premises of our argument for the claim that identity matters for prudential rationality. Sections 7 to 13 constitute the substantial bulk of the paper. There, we consider different ways a 'Parfitian' philosopher could react to our challenge. We provide detailed replies to the various Parfitian moves and show how this sheds light on where the more basic disagreement between us and Parfit lies. Finally, in section 14 we reflect on the wider philosophical value of the type of dialectic we have set up in this paper.

2. Values *Simpliciter*, Values for Someone, and Values for Someone at a Time

We presuppose that things being a certain way—states of affairs—can have value. They can be good or bad. If, for instance, a conscious being suffers at a given time, then this is, in itself, a bad thing. Of course, this being's suffering might also be instrumentally good, for its suffering might have positive consequences (such as progress in its intellectual development or its capacity to avoid later dangers). Yet, irrespective of its consequences, a conscious being's suffering is bad—that is, it is intrinsically bad when a conscious being suffers. We will not need any precise account of the difference between intrinsic and instrumental value in what follows. Rather, a different distinction will play a crucial role in our reasoning: the distinction between being good/bad *simpliciter* and being good/bad *for someone*. If a conscious being suffers, then (other things being

¹ For Parfit's arguments in favour of this view, see for instance Parfit 1984 chs. 12 and 13, and Parfit 1995.

equal), this is bad simpliciter. Of course, it also is bad *for* the specific subject who suffers. But the world is a worse place overall (other things being equal) in virtue of that subject's suffering, or so we want to say. A well-functioning medical system in a society, the establishment of the rule of law or progress in philosophy are all examples of states affairs that have a positive value simpliciter. If something is good/bad simpliciter in the sense we intend here, then it is good/bad in a non-relativized manner.² By saying that 'there are' such values we do not mean to 'reify' values. We do not mean to suggest that values are entities that should figure in one's overall ontology. Rather, by saying that there are values simpliciter, we merely want to express the very natural and quite harmless thought that states of affairs can be bad or good to a certain extent and in a sense which does not require adding "for a specific person", "at a specific time" nor any other relativization.

Now, in addition to values simpliciter, we want to claim that there also are *values for a person or for a subject* and that these should be carefully distinguished from the former. In many cases, things are bad for you because they involve your being in an intrinsically bad conscious state, for instance a state of suffering.³ One sometimes wants to express a relativist attitude about truth when one disputes the claim that all propositions are true or false simpliciter and insists that (at least some) propositions can be true for you and false for me. We do not endorse such a relativist view, neither about truth nor about value. Here we do not intend to introduce relativized values in a sense parallel to the claim that a proposition can be true for your, false for me and neither true nor false simpliciter. To say that suffering is bad for the person who suffers in a way in which it is not bad for anybody else does not in any sense imply the idea of 'retreat' from an absolute value claim (without relativization). When a relativist claims that some proposition is true for someone, she implies that claims to its truth simpliciter are somehow unfounded. No such idea is involved in our terminological proposal. That a state of affairs or an event is good/bad for someone does not exclude that it is good/bad simpliciter. To the contrary, if something is good/bad for someone, then it is often also good/bad simpliciter precisely because it is good/bad for someone.⁴ The relativist view about truth does not al-

² For discussions of relativized vs. non-relativized value and the relation between the two, compare Orsi 2015, ch. 4, and McDaniel 2014. It seems to us very plausible that goodness and badness can be further conceptualized by quantitative notions ranging from the very bad to the very good. For the present discussion, we entirely set aside questions as to how—in a precise language for moral philosophy—such notions should be further developed. We also set aside here epistemic worries about the conditions under which attributions of value are rationally justified.

³ Note that we do not endorse the view that only intrinsically bad conscious states can have value for someone. For instance, one may argue that it is bad for you if your son falsely believes after your death that you are responsible for a horrible crime. If so, then this is bad for you even though you will never know about it nor suffer from being unjustly blamed by someone you love. Nagel 1970 famously discusses a series of potential examples where something is bad for someone although it is never experienced as bad by the person concerned. We will remain neutral about that point and we will not presuppose any specific view on the kinds of states of affairs that are suitable for having value for someone.

⁴ This need not always be so, however. For instance, it might be instrumentally good for a person to have a car, because it gives her more freedom, without it being good simpliciter that she has a car, because she is a very bad driver. The principle restricted to in-

low for an analogous relation between truth for someone and truth simpliciter. Here is another way to emphasize the disanalogy. If something is good for you, then this is so objectively. It is a fact, it is not just ‘a fact for you’ (whatever that might be). If something is bad for you, then this is true simpliciter and not just ‘true for you’. It is, from the evaluative point of view, objectively true that it is bad *for you*.

Besides values simpliciter and values for someone, we would also like to introduce values for someone at a time. It seems very plausible that things can be good or bad for a person at a time while not being good or bad for her at some other time. This is obviously so for instrumental values. It may be good for you at a given time t to have a broken leg if this prevents you from being involved in a war at t . Yet it may be bad for you at a different time t' , for instance if having a broken leg at t' prevents you from enjoying mountaineering in the way you normally would have at t' . In the case of instrumental values, one might suspect that something may be good for a person at a time without being good for her in a time-independent manner. It is perhaps less plausible that something may have a positive or negative *intrinsic* value for a given person at a time without having such a value for that person in a time-independent manner. Typical examples of intrinsic values are instantiations of intrinsically good/bad properties of our conscious life such as suffering an unbearable pain. It is plausible to think—and we will actually argue for this claim in section 5—that it is bad for a subject (in a time-independent manner) to instantiate such a property at any point in its life.

One might propose to define ‘good/bad for a subject S ’ (in a time-independent manner) by quantifying over all moments in time, or at least over all moments during S ’s life. According to this proposal, for a state of affairs X to be good/bad for a subject S (without relativization to time) just is for X to be good/bad for S at all moments (or at all moments during S ’s life). We would like to resist such a suggestion. We agree with the intuition defended in Nagel (1970) that things can be good/bad for a subject even though they are *not* good/bad for that subject at any particular time. Furthermore, we will argue that something can be bad for a subject S in virtue of being bad for S at a specific moment while not being bad for S at other times.

For the purposes of this paper, we would like to leave open the question about the exact relation between values simpliciter, values for someone and values for someone at a time. We do not believe that any of these values can be reduced to any of the others but will not discuss further this intriguing issue here.⁵ We will however discuss in detail the specific issue of the relation between intrinsic value for someone and intrinsic value for someone at a time. A substantial assumption about that relation will be crucial for our argument against Parfit’s view on future-directed rational evaluation and prudential rationality.

trinsic value—i.e. ‘if something is intrinsically good for someone, then it has a certain positive value simpliciter’ (even though other aspects may of course outweigh that positive value)—seems less vulnerable to counter-examples.

⁵ Another question we would like to leave open is whether anything can be bad simpliciter without being bad for anybody. Perhaps every value simpliciter is ultimately grounded in value for some experiencing subject or other. One might also wonder if every value for someone is grounded in value for someone at a time. We doubt that this is the case, but we do not need to take a stance with respect to that issue either for our purpose here.

3. Values and Experiential Properties

Experiential properties in a broad sense are properties such that necessarily there is something it is like to have them for the one who instantiates them. On our understanding of experiential properties, they can only be instantiated by experiencing subjects, that is, by someone for whom it is like something to have them.⁶ It seems plausible that for instance thinking about a philosophical problem in the sense of active reflection requires one to have some phenomenally manifest awareness of the problem she is thinking about. If this is correct, then thinking about a problem is an experiential property in a broad sense. Experiential properties in a narrow sense fulfil an additional requirement: there is a specific way it is like to have them which is the same at any time for any subject instantiating them. Presumably, this is not so for the case just considered. To be thinking about e.g. the Newcomb's problem may be radically different at different times for the same person and it may be different for different subjects at different times. It is not an easy task to say what makes it the case that a given episode of thinking concerns the Newcomb's problem. However, it is plausible to assume that the relevant commonality between all subjects who think about that specific problem at some point in their life is not phenomenally constituted. In other words, it does not consist in any common way it is like for all subjects to engage in such thoughts.

The situation is different for those experiential properties that we will call *strict experiential properties*. There is some specific way it is like, for instance, to undergo a visual experience of the colour blue. The same applies, or so we would say, to many experiential properties involving emotions (e.g. being sad about something or being angry about someone) or—to take the standard example—to the experiential property of suffering pain. In all these cases, having the property at issue is at least in part phenomenally constituted. That is, having such a property at least partially consists in the way it is like to have it. We will work with the following understanding of strict experiential properties: a property E is a strict experiential property if and only if (a) there is a specific way it is like to have E (for any subject having E at any time) and (b) having E is at least partially constituted by the way it is like to have E.

Values for some subject S at a moment m in time can be grounded⁷ in the fact that S instantiates a given strict experiential property at m. Some strict experiential properties are such that having them at a given moment is intrinsically good/bad for the subject concerned at that moment. The claim seems very hard to resist if we consider for instance an experience of sustained and agonising pain accompanied by unbearable anxiety. Pain can be instrumentally good, of course. It usually warns you that something is wrong with a part of your body. It may contribute, in the long run, to your personal flourishing. However, if we leave such positive effects aside, it is a bad thing—in itself—for the subject concerned to instantiate that experiential property. That experience is intrinsically

⁶ For more on the notion of experiential properties and their relevance for the philosophy of mind, see Nida-Rümelin 2017, section 6, 2020a, section 2, and 2020b, ch. 5.

⁷ In using the notion of grounding here, we do not want to commit to the idea that values for S are ontologically 'nothing over and above' the event of S instantiating a certain experiential property. We intend to remain neutral on this issue. For views of grounding on which it is not the case that a grounded fact is nothing over and above its grounds, compare Audi 2012: 708-11 and Rosen 2010: 133.

bad for the one undergoing it. One way to put this obvious insight is to say the following: for every subject *S* and moment *m*, the fact that *S* instantiates such an experiential property at *m* is intrinsically bad for *S* at *m*.⁸

One might perhaps be tempted at this point to think that all values for a subject at a time are grounded in the instantiation of experiential properties and, more specifically, in the specific way it is like to have them. We will not presuppose such a strong claim about time-relative values for a subject. Here is a potential counterexample: perhaps it is intrinsically good for a subject at a moment *m* to gain insight into some deep philosophical truth. If so, then it is surely not the specific phenomenal character of the relevant cognitive experience that makes having that insight valuable for that person. Rather, or so one might argue, it is the cognitive access to truth which makes such an insight a good thing for the person concerned.⁹

For the sake of clarity, we will limit our discussion here to values grounded in the instantiation of strict experiential properties. We will assume as one of our premises that there are intrinsically good/bad strict experiential properties in the following sense:¹⁰

Definition 1

The property *E* is intrinsically good/bad iff any instantiation of *E* by a subject *S* at a moment *m* is intrinsically good/bad for *S* at *m*.

4. Kinds of Values and Kinds of Reasons

In everyday life, people are sometimes motivated by what we would like to call subject-neutral reasons. Recall our discussion of values simpliciter in section 2. Now, values simpliciter are what provide subject-neutral reasons, or so we want to say. We may define subject-neutral reasons in the following manner:

Definition 2

A subject *S* has a subject-neutral reason to strive for/try to prevent *X* iff *S* has that reason in virtue of *X* being good/bad simpliciter.¹¹

We mentioned above examples of states of affairs having positive value simpliciter such as a well-functioning medical system in a society, the establish-

⁸ One might want to ask for further justification here if one believes it possible for instance for someone to rejoice in experiencing pain. Even if this is so, it still seems very hard to deny that there are intrinsically bad experiential properties.

⁹ One might also think here of Nozick's experience machine as speaking in favour of the point that what can ground value for us is not exhausted by the phenomenal character of our experiences. Nozick's thought-experiment also highlights the significance that our access to truth or reality can have for what is good/bad for us. See Nozick 1974 (1999: 42-45).

¹⁰ To avoid complications that are irrelevant for our present discussion, we will also limit our definition of intrinsically good/bad properties to the case of experiential properties.

¹¹ We intend this formulation to be neutral with respect to the issue as to whether reasons are facts or mental states. (See amongst other Dancy 2000 and Alvarez 2010 for the former view, and see Davidson 1980 for the latter.) Saying that the agent has a subject-neutral reason in virtue of an axiological fact is compatible with the view that the axiological fact is itself the motivating reason. It is also compatible with the view that the recognition of that fact is the motivating reason. (We do assume, however, that one can have a reason and yet not recognize that this is so. Also, we would like to allow for cases where someone recognizes a reason without being moved by it.) Parallel remarks apply to the definitions of other kinds of reasons below.

ment of the rule of law or progress in philosophy. If one has a reason to bring about such states of affairs in virtue of their value simpliciter, then one has a subject-neutral reason. One may obviously also introduce a notion of subject-neutrality for a specific kind of motivation: someone's motivation for acting is subject-neutral just in case she is motivated by subject-neutral reasons or by apparent subject-neutral reasons (we should allow that subject-neutral motivation can be based on errors about value simpliciter). To put it intuitively, you act for subject-neutral reasons if your motivation for acting is that you would like to make the world a better place. For instance, honest politicians should arguably act for subject-neutral reasons. If one struggles for a world where everyone has access to food, to medical help, to education and culture etc., then one acts for subject-neutral reasons. One struggles for such a world because one recognizes that the realisation of such conditions would be a good thing—that it would be, in other words, *good simpliciter*.¹²

Arguably, it is more common for us in everyday life to be motivated by another kind of reasons, which we will refer to here as subject-relative reasons:

Definition 3

A subject S has a subject-relative reason to strive for/try to prevent X iff there is a subject S' such that S has that reason in virtue of X being good/bad for that specific subject S'.

If you try to prevent your child from being hurt, then you do not act in order to make the world a better place. When you try to put your child out of harm's way, you do not act to make sure that less people suffer overall. Your reason for acting is, rather, that you want to prevent something that would be bad for that particular subject. If you love a person, then a possible state of affairs being good for *that* person provides a reason for you to act in favour of its realisation. It is surely part of what it is to love someone that one is moved by subject-relative reasons concerning that very person. Of course, a situation where there is a subject-relative reason to act is often also a situation where there is a subject-neutral reason to act. For instance, you might help your partner in order to bring about something good for her. But in many (though not in all) cases, something being good for your partner makes it a good thing simpliciter.¹³ In such cases, you will normally be aware that what you intend to bring about is also good simpliciter precisely because it is good for her (and in virtue of being good for her). However, if you are like most of us (or so we claim), what moves you when you act in favour of your partner is not that the state of affairs you

¹² Values simpliciter can provide a subject-neutral reason to act for one person and yet fail to provide such a reason for someone else. For instance, say you are in a position to do something about the reduction of a dangerous chemical substance in drinking water. The fact that drinking water is contaminated is bad simpliciter and provides a reason for you to act accordingly. But that negative value may not provide other people with any reason for acting if they have no means to do anything about the drinking water. Hence, values simpliciter provide subject-neutral reasons to people depending on their specific situation. Such reasons are therefore not subject-neutral in the sense of being reasons for everybody independently of their specific role or position. Such reasons are subject-neutral in the precise sense (given by definition 2) of being grounded in values simpliciter.

¹³ See note 4 for a discussion of a situation where this is not the case.

strive for is good simpliciter. Rather, you are moved by your recognition of its value for the person you love.¹⁴

Now, some special cases of subject-relative reasons are of particular interest for our purpose here: cases where S and S' in definition 3 are identical. These are the cases of prudential rationality. For instance, we often act in favour of things we believe to be good for ourselves. If you take special care to eat healthy food, then you do not do so to make sure that there is an additional healthy person on Earth. You do so because you think that it is good for you. You do not listen to music because you wish to increase the total amount of aesthetic pleasure on the planet, even if you know that your listening to music does increase that amount. You do so for your own musical pleasure—that is, because it is something good for you. It would certainly be an over-intellectualization to assume that whenever we accomplish what is pleasurable for ourselves, we always do so in order to achieve what we take to be good for ourselves. But we surely do so in many cases. *Pace* Parfit, whose arguments at least partially undermine that apparently harmless assumption, we will assume for the moment that self-concern in that sense—being motivated to act by the recognition that something is good/bad for oneself—is not only perfectly normal and widespread but also rationally legitimate. We will return in more detail to this point in what follows. For now though, let us call reasons of that kind *de se* subject-relative reasons and define them as follows:

Definition 4

A subject S has a *de se* subject-relative reason to strive for/try to prevent X iff S has that reason in virtue of X being good/bad for S.¹⁵

Subject-relative reasons come in two different kinds, where each plays a different role in our cognitive architecture, or so we would like to suggest. Some subject-relative reasons are *de se* subject-relative reasons and they are related in obvious ways to *de se* attitudes when we recognize ourselves as having them. Other subject-relative reasons are directed at other subjects. We are motivated

¹⁴ In many real-life cases, it might not be easy to tell whether a person is moved by subject-relative reasons or by the corresponding subject-neutral reasons. A medical doctor helping a stranger in a case of emergency can be moved by subject-neutral reasons and, we suppose, normally is moved by such reasons. Certainly, it is not uncommon for a person with such a profession to be moved by both—by the subject-neutral reason that it is good simpliciter if the person in danger is rescued and by the subject-relative reason that it is good for that very person to get help. One can even imagine a situation where a medical doctor's sudden sympathy for a stranger makes her switch from pure subject-neutral motivation to primarily subject-relative motivation. In such a case, she first acted (mainly) in order to prevent a situation in which someone gets in serious medical trouble because she recognizes the negative value (simpliciter) of such a situation. After the switch, the medical doctor (mainly) acts in order to realize what is good for that particular person. Perhaps one might even argue that it is in fact desirable to be moved in both ways in situations of that kind. The issue might be worth pursuing further in the context of virtue ethics, for instance.

¹⁵ The definition allows that one can have a *de se* subject-relative reason and yet not know that this is so. For instance, you might have a *de se* subject-relative reason to avoid danger for yourself because you are in danger and you are in a position to avoid it, and yet not know that this is so because you haven't recognised the danger—or because you have recognised the danger for a specific person without understanding that the person in danger is yourself.

by those reasons for instance when the subject for whom something is valuable is someone we care about. Both kinds of subject-relative reasons have in common that one has them in virtue of something being valuable *for someone*.

To sum things up, we deem it helpful and enlightening to distinguish both between values simpliciter and values for someone as well as between corresponding kinds of reasons (subject-neutral and subject-relative reasons) and motivation (subject-neutral and subject-relative motivations). In addition, it will prove useful for our argument to introduce a specific kind of reason with respect to time (subject-time-relative reason), as well as a specific kind of motivation (subject-time-relative motivation). Hence, we obtain the following parallel definition of a subject-time-relative reason:

Definition 5

A subject S has a subject-time-relative reason to strive for/try to prevent X iff S has that reason in virtue of X being good/bad for a specific subject S' at a specific moment m.

It is not hard to find examples where someone has subject-time-relative reasons to act in a specific way. For instance, the fact that it will be good for your son later in his life to be able to cook tasty food is a subject-time-relative reason for you to make sure that he acquires cooking skills now. Or, to take a different example, the fact that it will be good for you to remember your happy holidays in a concrete manner in a few years is a subject-time-relative reason for you to take pictures during your present holiday.

One could be tempted to think that the relation between subject-neutral reasons and subject-relative reasons is perfectly parallel to the one between subject-relative reasons and subject-time-relative reasons. After all, from a formal perspective, one gets from one type of reason to the other by the mere introduction of a further parameter. One should note, however, that there is a crucial disanalogy here. As we've seen before, one may have both a subject-relative reason and a subject-neutral reason speaking in favour of the same action. Yet, it does not follow that one is motivated to act by both reasons in such a case. For instance, as noted earlier, you may have a subject-relative reason and motivation to act in favour of someone you love and yet have no corresponding subject-neutral motivation. You act to make your mother happy because you are moved by the reason that this is good for her. Your mother's happiness does make the world a better place; that she is happy is good simpliciter. However, it can well be the case that that value simpliciter does not motivate you when you strive for her happiness. It is possible to be moved by a subject-relative reason without being motivated by the corresponding subject-neutral reason, as this case shows. And such cases are even widespread. However, the analogous claim does not apply to subject-relative versus subject-time-relative reasons. The analogous claim would be this: you can be moved by a subject-time-relative reason without being moved by the corresponding subject-relative reason. Consider again one of our examples above. Can you be moved by the subject-time-relative reason that it is good for your son later in his life to have (later in his life) the capacity to cook without being moved by the corresponding subject-relative reason that it is good for him (time-independently) to have that capacity later in his life? This does not appear to be possible. If you do something because you think it is good for your son in ten years, then you also do what you do because you

think that it is good for him (time-independently), or so one might plausibly argue.¹⁶

5. Preparing the Argument: Transition from Intrinsic Value for Someone at a Time to Intrinsic Value for Someone

We are now able to introduce the premises of our argument for the claim that, contrary to Parfit's famous dictum, identity does matter. Let us first give a more precise content to that claim. The simple fact that a possible future event concerns you and is good or bad for you makes a legitimate and rationally justified difference for the kind of motivation you have for doing something about it.¹⁷

Let us consider a situation where a subject *S* knows at a moment *m* that a subject *S'* will experience sustained and agonising pain accompanied by unbearable anxiety at a later moment *m'*. Suppose the situation is such that it is possible for *S* at *m* to prevent that future misfortune. We will argue for the following claim with respect to this example. If *S* learns that she is *S'*, then this makes a legitimate and rationally justified difference for *S*'s motivation to prevent *S'* suffering at *m'*. Here is a more precise way to put the thesis we will be arguing for: (a) if *S* is identical to *S'*, then *S* has a specific kind of reason, namely a *de se* subject-relative reason, to avoid that *S'* is harmed at *m'* in virtue of that identity; (b) it is legitimate and rationally justified for *S* to be moved, in such a situation, by the relevant *de se* subject-relative reason. It would even be irrational not to be moved by such a reason, as we will argue below.

To experience sustained and agonising pain and unbearable anxiety are paradigmatic examples of intrinsically bad experiential properties in the sense of definition 1. Let us refer to the conjunction of these properties with the letter *E* so as to formulate the first premise of our argument as follows:

Premise P1 [Intrinsic value in virtue of phenomenal character]

The instantiation of *E* by *S'* at *m'* is intrinsically bad for *S'* at *m'*.

Premise P1 is hard to deny. If there are any intrinsically bad experiential properties in the sense of definition 1, then those chosen for our example are surely among them. Being in pain and in a state of anxiety are strict experiential properties: having them partly consists in the way it is like for any subject having them to have them. Furthermore, the specific way it is like to have them makes it a bad thing to have them. It is a bad thing to have them in virtue of their phe-

¹⁶ This disanalogy does not concern the relation between kinds of values or kinds of reasons but only that between kinds of motivations.

¹⁷ Parfit's claims concerning prudential rationality are motivated by his metaphysical view about identity. He argues that identity between an earlier and a later existing person—as we normally understand it—is an illusion. If this is so, then Parfit's claims about prudential rationality would be hard to deny. Given the metaphysical foundations of his claims about practical rationality, one might think that in order to seriously undermine the latter one must address the former. We agree that a full discussion of Parfit's views about rationality must include a response to his metaphysical motivation. In this paper, however, we focus on undermining his views about rationality directly. Since accepting the metaphysical view put forward by Parfit involves quite unavoidably endorsing his view about rationality, undermining the latter indirectly undermines the former as well. The ontological issue about transtemporal personal identity is directly addressed in a number of other publications by one of the authors: see Nida-Rümelin 2009, 2013, and 2020c, lecture 3.

nomenal character—in virtue of the characteristic way it is like to have them. Premise 1 hardly needs any justification. Any reader who accepts the notion we introduced of an intrinsically bad experiential property will have to agree that the premise is justified with respect to the example at hand.

The next premise of the argument is surely more controversial. For Parfit and his followers it might be natural to suspect that it already contains the deep error which lies, in their view, behind the common and natural assumption that identity matters for prudential rationality. We will come back below to possible objections. Let us first focus on explaining the content of the second premise and its intuitive motivation. According to that premise, it is a bad thing for a subject to have intrinsically bad experiential properties at any time in his or her life. It is bad for your son to suffer later in his life. You do not want him to suffer later in his life because you think that this is bad for him (time-independently) and you do not want anything to happen which is bad for this particular person you love. You wish to avoid future harm for those you love because you think it is bad for them. You recognize that it is bad for them (time-independently) to suffer later in their life just like it is bad for them to have suffered earlier in their life.¹⁸ The idea is this: if having a given experiential property at moment *m* is intrinsically bad/good for any subject *S* at *m* in virtue of the way it is like to have that property, then it is a bad thing for any subject *S* to have that property at any moment in his or her life. To have such an intrinsically bad experiential property at any moment is bad for the subject concerned. One may formulate this premise as follows:

Premise P2 [Transition from intrinsic value for a subject at a time to intrinsic value for a subject]

If the instantiation of an experiential property *P* by *S* at *m* is intrinsically bad for *S* at *m*, then the instantiation of *P* by *S* at *m* is intrinsically bad for *S* (time-independently).

Perhaps one may defend the even stronger thesis that something's being intrinsically bad for a subject *S* at a time is—in general—intrinsically bad for *S*. Such a general transition from being intrinsically bad for someone at a time to being intrinsically bad for that subject may seem plausible, but we will not use that claim here as a premise for two reasons. First, the general thesis is more difficult to judge by testing one's intuitions. Second, it is always more cautious to use the weakest version of a claim which serves the purposes of one's argument. For our argument to be valid we only need the relevant claim about the special case of properties that are bad to instantiate in virtue of the way it is like for the subject concerned to have them.

The negative value of the instantiation of *E* by a subject *S* at a given moment *m* resides entirely in the way it is like for *S* to be involved in that particular instantiation. The intuition we would like to defend is this. If the instantiation of a property is bad in virtue of the way it is like to instantiate that property for the person *S* concerned, then to be involved in that instantiation is bad for *S*, no matter when during *S*'s life the instantiation occurs. According to this claim, it is intrinsically bad for a person to have intrinsically bad experiences at some mo-

¹⁸ Intuitions tend to be different for the future and the past case. We suggest that in both cases the transition from values for a person at a time to time-independent values for a person is equally valid. The difference underlying the impression that the two cases are not analogous is perhaps partly due to the fact that only in the future case we can still do something to prevent the realization of the state of affairs at issue.

ment m . On this view it is true to say of Anton at some arbitrarily chosen moment m' that his suffering at an earlier moment m is intrinsically bad for him, even though at m' he perhaps does not remember that suffering anymore. One must carefully distinguish two distinct claims in this context. It is one thing to say that such a statement about Anton is true when uttered at m' and quite a different thing to say that the past suffering is bad for Anton at m' . We claim with respect to Anton's past or future suffering only the former but not the latter. It is bad for Anton to suffer at some specific moment m in his life and to say this is to express a truth no matter when one says it. However, it might be false to say that it is bad for Anton at m' that he suffers at m if m' is a moment different from m . If so, then to say the contrary is false whenever one says it.

6. Completing the Argument: Transition from De Se Subject-relative Value to Adequate Subject-relative Motivation

We are now in a position to summarize and to complete our argument:

Premise P1 [Intrinsic value in virtue of phenomenal character]

The instantiation of E by S' at m' is intrinsically bad for S' at m' .

Premise P2 [Transition from intrinsic value for a subject at a time to intrinsic value for a subject]

If the instantiation of an experiential property P by S at m is intrinsically bad for S at m , then the instantiation of P by S at m is intrinsically bad for S (time-independently).

Premise P3 [Transition from intrinsic value for a subject to de se subject-relative reason]

If X is intrinsically bad for S , then S has a *de se* subject-relative reason to avoid X .

Premise P4 [Transition from de se subject-relative reason to de se subject-relative motivation]

If S has a *de se* subject-relative reason to avoid X , then it is rationally permissible for S to be motivated by her *de se* subject-relative reason to avoid X .

Assumption included in the description of the kind of situation at issue:

(A1) S is identical to S' .¹⁹

Consequences:

(C1) The instantiation of E by S' at m' is intrinsically bad for S (time-independently). *[From P1, P2, A1]*

(C2) S has a *de se* subject-relative reason to avoid the instantiation of E by S' at m' . *[From C1, P3]*

(C3) It is rationally permissible for S to be motivated by her *de se* subject-relative reason to avoid the instantiation of E by S' at m' . *[From C2, P4]*

Consequence C3, on its own, is not yet in contradiction with Parfit's dictum that identity does not matter. However, if the justification given by the above ar-

¹⁹ Since the Parfitian declares identity between people to be illusory, one might think that she would simply deny A1. However, Parfit puts forward a notion of transtemporal personal identity which is freed from our usual mistakes about it and which does obtain in normal cases (see for instance Parfit 1984, ch. 11). Therefore, on her interpretation, a Parfitian should be ready to accept these assumptions about the example.

gument is correct, then the assumption A1 that the subject concerned is identical to the one undergoing the intrinsically bad experience is essential. Assumption A1 is required in order to reach consequence C3. C3, therefore, applies to the kind of situation we consider *because* the agent is identical to the subject undergoing the later (possible) experience. We will see in the next section, however, that our Parfitian opponent has a strategy to undermine that reasoning.

Moreover, we want to suggest that a stronger conclusion is justified as well. Namely: the conclusion that it is rationally required of an agent S (in the kind of situation considered) to be motivated by her *de se* subject-relative reason to avoid that S' instantiates E at m', provided that S has the relevant information. The relevant pieces of information in question can be formulated as follows: S knows that the instantiation of the experiential property E at m' is intrinsically bad for S' at m', and S knows that she is none other than S'. With a further plausible premise that we will introduce in a moment, it follows that S knows that the instantiation of E by S' at m' is intrinsically bad for her. Now, if someone knows that something is bad for her and yet is not moved by the corresponding *de se* subject-relative reason, then she thereby exhibits a form of practical irrationality. Therefore, it is not only permissible for a subject to be moved by her *de se* subject-relative reason in such a situation; it is, rather, even irrational for a subject in such a situation not to be moved by that reason.

A precise formulation of the argument we just sketched can be articulated (see below) by making use of the three premises P1, P2 and P3 of the first version of the argument above. The second version also makes use of P3 as a presupposition of premise P4', which replaces P4, and it uses P2 as a presupposition of premise P5, which needs to be added. The argument uses again assumption A1 (identity between S and S'), as well as two additional assumptions about S's epistemic situation, which characterize the kind of situation at issue: S knows that S' instantiates an intrinsically bad experiential property at m' and S knows that she is S'.

Premise P1 [Intrinsic value in virtue of phenomenal character]

The instantiation of E by S' at m' is intrinsically bad for S' at m'.

Premise P2 [Transition from intrinsic value for a subject at a time to intrinsic value for a subject]

If the instantiation of an experiential property P by S at m is intrinsically bad for S at m, then the instantiation of P by S at m is intrinsically bad for S (time-independently).

Premise P3 [Transition from intrinsic value for a subject to de se subject-relative reason]

If X is intrinsically bad for S, then S has a *de se* subject-relative reason to avoid X.

Premise P4' [Transition from de se subject-relative reason to irrationality; presupposes P3]

If S knows that X is bad for her and is not motivated by her corresponding *de se* subject-relative reason to avoid X, then S thereby exhibits a form of practical irrationality.

Premise P5 [Knowledge of the content of P2; presupposes P2]

S knows that for any subject it is bad (time-independently) to instantiate an intrinsically bad experiential property at some moment.

Assumptions included in the description of the kind of situation at issue:

- (A1) S is identical to S'.
 (A2) S knows that she is identical to S'.²⁰ [Presupposes A1]
 (A3) S knows that the instantiation of E by S' at m' is intrinsically bad for S' at m'. [Presupposes P1]

We can now deduce the desired result in the following manner:

- (C1) The instantiation of E by S' at m' is intrinsically bad for S (time-independently). [From P1, P2, A1]
 (C2) S has a *de se* subject-relative reason to avoid the instantiation of E by S' at m'. [From C1, P3]
 (C3) S knows that the instantiation of E by S' at m' is intrinsically bad for S' (time-independently). [From A3, P5]
 (C4) S knows that the instantiation of E by S' at m' is intrinsically bad for her. [From C3', A2]
 (C5) S knows that the instantiation of E by S' at m' is intrinsically bad for her and she has the corresponding *de se* subject-relative reason to avoid the instantiation of E by S' at m'. [From C4, C2]
 (C6) If S is not motivated to avoid the instantiation of E by S' at m' by her *de se* subject-relative reason to do so, then S thereby exhibits a form of practical irrationality. [From C5, P4']

If one accepts P3, then it is hard to see why one would resist the premise P4'. It seems to us highly plausible that it is part of what it is to be a rational agent that one is motivated by those *de se* subject-relative reasons that one recognizes oneself as having. To recognize that one has *de se* subject-relative reasons for acting in order to avoid X is to see that one has such reasons *in virtue* of X being intrinsically bad for oneself. To see that one has such a reason and yet remain unmoved by it is clearly a form of irrationality, or so we would like to insist. We would like to say something similar with respect to P5. If one accepts P2, then P5 is quite unproblematic. Arguably, it is part of our normal implicit conception that it is bad for a subject to have a bad experiential property at some time in her life. Moreover, a Parfitian philosopher is likely to agree that this is so. If P2 is true, then having that conception amounts to knowledge.

7. A Parfitian Strategy to Undermine our Challenge

According to the argument developed in the previous section, identity matters: it is rationally required of a person S to be motivated by a particular kind of reasons (*de se* subject-relative reasons) to avoid the occurrence of an intrinsically bad experiential property in *her own* future life. Identity plays a role in our argument in two places: it is bad for a person S if an intrinsically bad experiential property is instantiated by a subject S' in case $S = S'$ and in that case the instantiation of that property is bad for S *because* $S = S'$.²¹ Furthermore, if S is not mo-

²⁰ A2 must be read as attributing *de se* knowledge to S. The same applies to the relevant consequences that we draw.

²¹ One may express in a somewhat technical manner how premise P2 leads to the result that identity matters. The transition from value for someone at a time to value for someone (time-independently) creates extensionality in the way required with respect to "S" in the assertion of value which now reads "X is bad for S". We can replace "S" by an expression referring to the same subject without changing the truth value of that assertion. The situation is different for the original sentence "X is bad for S at moment m". For a

tivated to act to avoid that future event (although S knows that the experiential property is intrinsically bad and that *she* is the one who is at risk of instantiating it), then S is irrational. S is irrational in that case because she has no such motivation despite her knowledge that she is identical with the subject concerned. Therefore, knowing that one is identical with the person a future event will involve does make a difference, according to our argument, for the question as to what one should rationally be motivated to do about that event.

We would now like to explore how a philosopher who is convinced of the contrary thesis by Parfit's arguments could or should react to our argument. In general, one will expect a counter-argument to a given argument to take the form of an attack on the premises used or the validity of the deduction leading to the relevant conclusion. In the present case, the dialectical situation is, however, of a special kind. A Parfitian is not opposed to the idea that, in normal life, we have reasons (or, perhaps even, are rationally required) to care about our own future. Rather, she opposes the idea that the fact that we have such reasons depends on that future being ours. Therefore, she need not reject the conclusion that S should be motivated to avoid the instantiation of an intrinsically bad experiential property by a subject S' when S learns that she is S'. However, according to Parfit, S has that reason only in virtue of a necessary condition for identity between S and S', a condition consisting in the fact that the mental states of S at m and the mental states of S at m' are causally related in the appropriate way. Following Parfit, we will call that causal relation '*relation R*', and we will refer to events standing in relation R to one another as being *R-related*.²² According to Parfit, as is well-known, identity between S (existing at m) and S' (existing at m') consists in the satisfaction of two conditions. First, the mental states of S and S' at the respective moments are R-related. Second, S' is the only subject existing at m' whose mental states at m' are R-related to those of S at m (*uniqueness condition*). The basic Parfitian idea here is this: whether or not the uniqueness condition holds should not make a difference for whether it is rational for S to be motivated to do something about the latter instantiation of an intrinsically bad property. Therefore, only the R-relation matters. S has a reason (at m) to prevent the instantiation of an intrinsically bad experiential property by S' at m' if S = S', but this is so only because the identity of S and S' involves that the mental states of S at m and the mental states of S' at m' are R-related.

Given the dialectical situation just described, the Parfitian need not base her response on a denial of our premises. She could accept them and go on to argue that our argument expresses an ill-founded motivation for our conclusion. On *our* interpretation of our premises, this is surely what the Parfitian must do in response to our challenge. The Parfitian cannot accept our premises on the reading we intended and have explicated when we introduced them. However, she can accept all our premises on a different reading, namely on their re-interpretation in Parfit's theoretical framework. She can then go on to argue that the premises of our argument, which involve talking implicitly or explicitly of

Parfitian, it is natural to assume that "S at moment m" is the relevant referential expression in that sentence. If so, then replacement *salva veritate* is only possible if the moment considered is the same as well. Even without that Parfitian assumption the sentence "X is bad for S at moment m" does not allow for the conclusion that X is bad for S at the earlier moment at which S deliberates about what to do.

²² See Parfit 1984 (1986: 215).

identity between subjects, are true only because identity between subjects requires the right kind of causal connection between mental states. The Parfitian may therefore adopt the following general strategy in order to undermine our argument for the claim that identity matters: for each of our premises, she proposes an analogue within her framework which does not involve identity but only causal connectedness by the R-relation. These premises then lead to a more general conclusion which implies our conclusion. According to Parfit's account of identity (although not according to our understanding of identity concerning subjects), his proposed new premises will imply our premises. If the Parfitian can successfully do so and if each ersatz premise is acceptable, then she will have shown the failure of our argument. We mistakenly thought that our crucial assumptions depended on identity between S and S' while in reality only the R-relation matters. The Parfitian will then have done two things at once. She will first have shown that we were wrong in concluding that S has a reason to avoid the later misfortune for S' in a case of identity between S and S' because of that identity. At the same time, she will also have proposed an error theory: we were under the illusion that identity plays a crucial role for reaching our conclusion because we overlooked that—on the correct account of transtemporal identity between subjects—our premises follow from premises which do not involve identity. We will now present and discuss a Parfitian rejoinder to our argument that follows the strategy just outlined.

8. A Parfitian Version of Premise P1

In formulating our argument, we make crucial use of the notion of value for someone and value for someone at a time. One may doubt, however, that these notions are acceptable for a philosopher sharing Parfit's metaphysical views on which persons need not be mentioned in a complete description of reality. We may plausibly assume that Parfit would have defended the same metaphysical view with respect to the more general case of experiencing subjects. According to that view, to fully describe mental reality one only needs to talk of mental events and their causal interrelations, without saying anything about the existence of subjects and their identity across time.²³ It is indeed difficult to see how one might incorporate the idea that events have values for a subject into a view of reality devoid of subjects. Perhaps Parfit would have simply rejected the notion of value for a subject (simpliciter or at a time), as well as the notions of subject-relative reason and subject-relative motivation, since the latter notions are defined in terms of the former. As a consequence, he would thus have had reasons to reject all premises of our argument since they all use one or more of these notions. Resisting the argument in this manner is probably the most radical of all possible Parfitian replies. The Parfitian motivation for rejecting subject-relative value, reason, and motivation would be metaphysical in this case. In order to defend our terminology against that reaction we would have to argue directly against Parfit's metaphysical views of experiencing subjects. One may in-

²³ This is the view explicitly defended in Parfit 1984 and it is the view we call 'Parfitian' here. We are not going to discuss if and how later variants of Parfit's view might motivate a different reaction to our argument. For a later variant of Parfit's view, see for instance Parfit 1999. Note that the view defended in this later paper would actually make it harder for the Parfitian to reject our premises on the ground that they use notions such as subject-relative reasons.

deed argue against Parfit's metaphysical views in a number of different ways.²⁴ There is no room, however, to review these arguments in any detail here. That being said, we will later return to what is perhaps the most basic reason for us to reject his metaphysical views, a reason closely related to our understanding of experiential properties. For the moment though, let us put aside the radical Parfitian reaction that flatly rejects our terminology and let us move forward to the Parfitian strategy described above.

Our first premise P1 (restated below) presupposes that there are intrinsically bad experiential properties in the sense of our definition: properties that are intrinsically bad in virtue of the way it is like for the one who has them to have them. As a result, P1 only tells us something about the kind of examples one should have in mind here, i.e. cases where the person S' instantiates at moment m' one of those intrinsically bad experiential properties.

Premise P1 [Intrinsic value in virtue of phenomenal character]

The instantiation of E by S' at m' is intrinsically bad for S' at m'.

A Parfitian will not deny that there are intrinsically bad 'experiential events'. She can agree, furthermore, that such experiential events are intrinsically bad because they are instantiations of intrinsically bad experiential properties. But when it comes to the question as to what experiential properties are and why they are intrinsically bad (in those cases where they are), then the Parfitian will have trouble following our suggestions. According to our proposal, for a property to be an experiential property means that—in any instantiation of that property—there is some subject for whom there is a way it is like to instantiate it. It seems obvious that a Parfitian will reject that proposal. To accept our suggestion would amount to admitting that experiences cannot occur without a bearer, without someone undergoing it. This would mean that—contrary to Parfit's metaphysical view—reality cannot be fully described without mentioning experiencing subjects.

It does not yet follow that a philosopher following Parfit must reject the claim that to have an intrinsically bad experiential property at a moment m is bad for the subject concerned at that moment. However, the Parfitian philosopher who accepts premise P1 must reinterpret the notion of 'something being bad for a subject S at a given moment m' within her own metaphysical framework. The natural way for a Parfitian to interpret that locution is along the following lines: there is a mental event occurring at m which is part of the chain of mental events constituting subject S and that event is the instantiation of an intrinsically bad property.

It will be useful for the purpose of our discussion to use the notion of 'chains' of mental events that are causally related to one another in the relevant manner. According to Parfit, specific chains of mental events constitute subjects. The mental events they contain must be causally related in the right manner (by the relation R) and the chain must neither fuse nor divide. In what follows, we will use the term *R-chains* for chains of mental events containing only events that stand in relation R to one another.²⁵

²⁴ Compare Nida-Rümelin 2009 and 2020c, part IV.

²⁵ This is not a locution that—to our knowledge—Parfit himself uses. By talking in that way, we do not mean to imply that Parfit must be interpreted within a four-dimensionalist framework.

It is interesting to note that, according to the Parfitian interpretation, the order of definitional reduction is inverted with respect to our own suggestion. While we defined what it is for a property to be intrinsically bad in terms of what it is for something to be intrinsically bad for a subject *S* at a time *m*, the Parfitian proposal under consideration explains intrinsic subject-time-relative value in terms of intrinsic value of properties. Given this, something gets lost—or so we would say: the claim that the experiential event at issue is bad for a subject at a time is not explained (contrary to our proposal) by the way it is like for that subject to have that property at that time. In our view, there is an essential link between the nature of experiential properties and value for a subject at a time. The locution ‘for a subject’, we submit, occurs in the same sense within the attribution of subject-relative value as it occurs in the locution ‘there being something it’s like for someone’. There is no such essential link between value for a subject and the nature of experiential properties within the Parfitian framework. A Parfitian will regard this as a virtue; we hold it to be a vice. We will return to that point shortly.

In light of the above considerations, it seems obvious how a Parfitian can and should formulate her replacement for our first premise, P1. Namely: the experiential event at issue (the instantiation of *E* by *S* at *m*) is intrinsically bad. To say further that it is *bad for S' at m'* does not add anything if we presuppose—in the description of our case—that it is a mental event of *S'* occurring at *m'*; in other words that it occurs at *m'* in the chain of causally interconnected events which constitute *S'*. On a Parfitian view, premise P1 is only a strange way to express that the experiential token event at issue is intrinsically bad. We thus arrive at the following Parfitian replacement of our first premise:

*Premise P1** [Intrinsic value in virtue of phenomenal character]

The instantiation of *E* by *S* at *m* is intrinsically bad.

According to our description of the general Parfitian strategy, P1* should imply P1 on the Parfitian view. This is in fact the case. P1 is a trivial consequence of P1* if we assume the above reading of what it is for an experiential event to be intrinsically bad for a subject *S'* at a time *m'*: namely, to be intrinsically bad and to occur at *m'* within the R-chain constituting *S'*. The first of these conditions is explicitly expressed in P1*. The latter conditions are expressed by calling the event an instantiation of a property by *S* at *m* which presupposes—by Parfitian lights—that the event occurs at *m'* in the R-chain constituting *S'*.²⁶

9. Criticizing the Parfitian Replacement of Premise P1

If one thinks—as does the Parfitian according to the above diagnosis—that the intrinsic value of an experiential property is conceptually and ontologically more fundamental than subject-relative intrinsic value at a time, then the idea must be that the qualitative character ‘alone’ so to speak (of instantiations of that property) is constitutive of that intrinsic value. In our view, however, the idea that qualitative character could occur ‘alone’—without there being someone for whom it is like something to instantiate that property—does not make sense. The only sense that one can give to the locution that ‘a certain phenomenal quality occurs at a given moment’ is, in our view, that a specific experiential

²⁶ For simplicity, we are working here with a notion of consequence which includes among the consequences of a claim its presuppositions.

property is instantiated at that moment. Furthermore—as restated a moment ago—according to our understanding of experiential properties, any instantiation of an experiential property E necessarily involves a subject S such that it is like something for S to instantiate E on that occasion. Accordingly, there is no way to separate qualitative character from there being something it is like for a subject at a given moment. Any instantiation of qualitative character is an instantiation of qualitative character for someone.

As a result, we would suggest, there is no way to understand what makes an experiential property intrinsically bad without having in mind the way it is like for subjects having that property to have it. *It is the way it is like for subjects instantiating an intrinsically good or bad property which grounds its intrinsic value.* In other words, an experiential property can only be intrinsically bad in virtue of the way it is like for subjects to have that property at a given time. Hence, there is no way for intrinsic value to be grounded in phenomenal qualities in isolation. Such an idea turns out, or so we claim, to be conceptually confused. A given instantiation of an experiential property can be intrinsically bad in virtue of its qualitative character only because it is intrinsically bad for the subject whose instantiation of that property is the instantiation at issue.

Furthermore, we cannot agree with P1* on the Parfitian reading. As explained above, the Parfitian reading excludes that one explains why the experience is intrinsically bad by reference to the way it is like for the subject concerned to undergo it. In our view, this reveals not only a misunderstanding of intrinsic values of experiences but also a more fundamental misunderstanding of what experiential properties are, by their very nature. If our analysis is right, then a misconception of experience leads the Parfitian philosopher to a misconception of value as well.

We just sketched our reasons for rejecting P1* on its Parfitian reading—reasons that are based on our understanding of what it is for a property to be experiential. Our view on the nature of experiential properties motivates our rejection of the Parfitian metaphysical view of experiencing subjects as well. Only mental events and their causal relations would have to be included in a complete Parfitian description of reality. Given our view on the nature of experiential properties, we insist that mentioning mental events, at least when they consist in the instantiation of experiential properties, involves, unavoidably, mentioning experiencing subjects as well. If we are right that an instantiation of an experiential property requires there to be someone for whom there is a way it is like to have the property at issue, then no experience can occur without someone undergoing it. We can now say how we react to the more radical Parfitian reply that rejects the notions of value for someone (time-independently and at a time). This reply is ultimately based on a metaphysical view which we take to be mistaken because it misconceives the nature of conscious experience.

10. A Parfitian Replacement for Premise P2

Our second premise relates value for someone at a time with time-independent value for someone:

Premise P2 [Transition from intrinsic value for a subject at a time to intrinsic value for a subject]

If the instantiation of an experiential property P by S at m is intrinsically bad for S at m, then the instantiation of P by S at m is intrinsically bad for S (time-independently).

Our own motivation for P2 is simple and apparently obvious. It is bad for the subject S to instantiate an intrinsically bad experiential property at some moment m in his or her life, because it is subject S for whom there is a way it is like to have that property at m. It is bad for you to suffer at a later moment because it will be you for whom there is a way it is like to suffer at that moment. In general, in virtue of the way it is like for S to have an intrinsically bad experiential property at a given moment, it is bad for S (time-independently) that she instantiates that experiential property at m.

A philosopher developing an argument against our challenge who follows the strategy described above must find a premise P2* which satisfies the following conditions simultaneously:

- (a) Within the Parfitian framework, P2* must imply P2.
- (b) According to P2*, if the instantiation of an experiential property by a subject S at m is bad for S (time-independently), then this is *not* so *because* the subject undergoing the experience at m is identical with the one for whom this is bad (time-independently). Rather, it must be so in virtue of a relation which is compatible but does not require identity between the subject who undergoes the experience and the subject for whom we assess its intrinsic value.
- (c) The left-hand side of the conditional of P2* must coincide with P1* so that we can apply modus ponens.

Bearing in mind the constraints (a) to (c) for the Parfitian replacement of P2, the following proposal seems very natural:

Premise P2 [Transition from intrinsic value of experiences to intrinsic value for R-chains]*

If the instantiation of an experiential property P is intrinsically bad, then the instantiation of P is intrinsically bad for any R-chain in which it occurs.

Premise P2* obviously satisfies (c). So, let us have a look at the other two constraints. For (a) to be satisfied it must be the case that—within the Parfitian framework—the left-hand side of P2 implies the left-hand side of P2* and that the right-hand side of P2* implies the right-hand side of P2. This is in fact the case. Within a Parfitian framework, the only sense one can give to the left-hand side of P2 is to say that the experiential event at issue is intrinsically bad. If so, then the left-hand side of P2 trivially implies (within the Parfitian framework) the left-hand side of P2*. (The latter implication also holds in our understanding of these assertions.)

To see that in the Parfitian framework the right-hand side of P2* implies the right-hand side of P2, we must recall Parfit's theory with respect to transtemporal personal identity (which, we assume, applies also to experiencing subjects). According to that theory, a subject S who exists at an earlier moment m is identical with a subject S' who exists at a later moment m' if and only if (i) the mental states of S at m are causally connected in the right way with the mental states of S at m' (that is, these mental states are R-related) and (ii) no other subject existing at m' has mental states at m' which are R-related with those of S at m. Within that framework, experiencing subjects are special R-chains; they are R-chains where no fusion nor fission occurs. It follows that the right-hand side of P2* im-

plies the right-hand side of P2. This completes our argument that P2* does indeed satisfy constraint (a), as is required for the Parfitian argument to succeed.

According to the argument we are considering, P2 is true in virtue of P2*. In other words, it is intrinsically bad for a subject S if it instantiates an intrinsically bad experiential property at some point in his or her life only because the instantiation of that property belongs to the R-chain which is that subject. P2* does not in any way capture the intuition underlying P2. The idea that the instantiation of an intrinsically bad experiential property at some point in a person's life is intrinsically bad for that person (time-independently) because it is that person for whom there is a way it will be like to have that property is lost in the proposed replacement for P2. However, that in itself is no argument against the proposal from the Parfitian point of view. To the contrary, this is, according to the Parfitian view, exactly the way it should be, since on that view the intuition underlying P2 is misguided. The underlying intuition, or so the Parfitian will tend to argue, is incompatible with the metaphysical view of mental reality that we are rationally required to adopt.

11. Criticizing the Parfitian Replacement of Premise P2

P1* has a reading which allows us to accept that premise as true. There is no such reading in the case of P2*. On the contrary, there seem to be a number of decisive reasons for rejecting P2*. P2* generalizes the notion of something being bad for an individual X in such a way that the notion becomes applicable in cases where X is not an experiencing subject. We find it hard or impossible to make sense of such a generalized notion. In our view, understanding the relevant sense in which things can be bad for a person—in particular, understanding the sense in which token experiences can be intrinsically bad for an individual X (time-independently)—presupposes that X is a subject. But chains that fuse or divide are not subjects.

An important reason for rejecting P2* has to do with what we take to be the best motivation for P2. According to P2* whether an experience occurring at a later moment is your experience or someone else's does not play any role for whether or not the occurrence of that experience is intrinsically bad for you. This means, given the nature of experiential properties, that, according to P2*, for the intrinsic value a token experience has for you, it is irrelevant if there will be a way it is like for you to undergo it. If one understands and accepts the notion of experiential properties as introduced before, then one will have to say that this result is plainly absurd. Of course, a convinced Parfitian will not be impressed. She will rather try to show that our notion of an experiential property is somehow based on a cognitive illusion. She might argue that this notion makes no sense—or perhaps that even if it makes sense, it has no extension. Even though it will be difficult, at this point, to continue exchanging arguments, one will have made progress in reaching a better understanding of where the deeper disagreement lies. The disagreement has its origin, once again, in a disagreement about what experiential properties are—in other words, in a disagreement about what conscious experience consists in.

A further reason for rejecting P2* stems from its metaphysical motivation, which we take to be mistaken. A typical Parfitian argument in favour of P2* relies on the idea that there is no genuine factual difference in the relation between a subject S existing at moment *m* and a given future token experience (a) in a

case where the experience is an experience of S, and (b) in a case where the experience is only R-related to the mental states of S at m but does not belong to S. The Parfitian account implies that there is no such genuine factual difference since the only difference between the two cases consists, according to that account, in the satisfaction versus violation of the uniqueness condition. From our point of view this consequence of the Parfitian theory is a *reductio ad absurdum*. There is a factual difference between the two cases, we would insist, a difference clear enough to be grasped: in the case of identity, there is a way it is like for S to undergo that experience, whereas in the case of R-relatedness without identity, there is none. A Parfitian must insist that we only appear to be able to grasp a genuine difference when we understand the previous claim. Parfit often admitted that his claims about identity are hard to believe. We would rather claim that the present consequence is not only hard but impossible to seriously endorse. As already noted, there are a number of further reasons for rejecting Parfit's metaphysical claims about mental reality. Every such reason is, a fortiori, a reason to reject P2*, for P2* is based on those (in our view untenable) metaphysical assumptions.

There is a related worry here, which is not itself directly a reason to reject P2*. It is rather a way to object to the role P2* is supposed to play in the Parfitian overall counter-argument. For the strategy to work, P2* should imply P2—and it does, according to the Parfitian thesis about the transtemporal identity of experiencing subjects. However, we reject that thesis. We reject it precisely because it overlooks the essential link between (i) S being identical to the subject undergoing a future experience and (ii) there being a way it is like for S to undergo that experience. Parfit and his followers can accept that the identity at issue holds if and only if (ii) is the case. However, their account of identity implies the following claim: that there is a way it is like for S to undergo the future experience E consists in the fact that the future experience is R-related to S's mental states and the fact that there is no other subject existing at m' whose mental states are related to S's states at m in this manner. But Parfit himself provides the best reason for saying that this explanation cannot be adequate. According to that explanation, whether or not you are the one for whom there will be a way it is like to instantiate a given experiential property at a future moment depends on whether there is more than one subject R-related to your mental states now. If one acknowledges that there is a factual difference between there being a way it is like for S to undergo a future experience and there being no way it is like for S to undergo it, then it seems clear that this difference cannot depend on the fulfilment of the uniqueness requirement.²⁷ We claim that one must concede that there is such a factual difference. If so, then the Parfitian account of transtemporal identity for experiencing subjects must be rejected. The Parfitian will surely disagree. She is likely to insist that the intuitions on which we rely (i.e. that there is a factual difference between the two cases) are cognitive illusions and should not be taken seriously. Perhaps then the substantial progress of this discussion is that it shows that one cannot accept the apparently harmless notion of experiential properties we propose if one follows Parfit in his views about identity and prudential rationality. Following Parfit in that respect commits one to deep suspicion about the way we naturally think of conscious expe-

²⁷ For a more elaborated version of this reasoning, see Nida-Rümelin 2006, sections 3.10, 4.3 and 4.6.

rience itself—a way captured, or so we claim, in an intuitively accessible way in the account of experiential properties we suggested.

12. Completing the Parfitian Parallel Argument

It seems clear from Parfit's writings that he accepts that one has reasons to care about one's 'successors'.²⁸ He thus implicitly accepts the existence of what one may call *de se* (R-)chain-relative reasons. A subject-relative reason (in our terminology) is a reason one has because something is good/bad for a particular subject. Analogously, a chain-relative reason is a reason one has because something is good/bad for a particular R-chain. A *de se* subject-relative reason is a reason an agent has if she is the one for whom something is good/bad. Analogously, *de se* chain-relative reasons are reasons one has because something is good/bad for an R-chain to which one's present mental states belong. We assume in what follows that a Parfitian can accept the proposed terminology and would be ready to use it in order to reject our argument. With these preparations in mind, it will be obvious how one must complete the Parfitian argument. Let us first recall premises P3 and P4 used in the argument above:

Premise P3 [Transition from intrinsic value for a subject to de se subject-relative reason]

If X is intrinsically bad for S, then S has a *de se* subject-relative reason to avoid X.

Premise P4 [Transition from de se subject-relative reason to adequate motivation]

If S has a *de se* subject-relative reason to avoid X, then it is rationally permissible for S to be motivated by her *de se* subject-relative reason to avoid X.

The Parfitian can replace these premises by the following assertions:

Premise P3 [Transition from intrinsic value for an R-chain to de se chain-relative reason]*

If X is intrinsically bad for an R-chain and if S's mental states at moment m belong to that R-chain, then S at m has a *de se* chain-relative reason to avoid X.

Premise P4 [Transition from de se chain-relative reason to adequate motivation]*

If S has a *de se* chain-relative reason at m to avoid X, then it is rationally permissible for S to be motivated at m by her *de se* chain-relative reason to avoid X.

On a Parfitian understanding, *de se* subject-relative reasons are reasons only in virtue of the fact that what is bad/good for a subject is bad/good for an R-chain. Only chain-relative reasons really count: subject-relative reasons ultimately reduce to chain-relative reasons. If one accepts this view of the relation between subject-relative reasons and chain-relative reasons, and of the relation between subject-relative values and chain-relative values mentioned earlier, then one will conclude that P3* implies P3 and that P4* implies P4. The important idea here is that P3 is only true in virtue of P3* being true and P4 is only true in virtue of P4* being true.

In order to be able to state the Parfitian parallel argument with all its premises, we have to recall the premises P1* and P2* discussed in detail above and we must add the following assumption A1* (formulated within the Parfitian framework), which replaces the assumption of identity between the subject suffering at moment m' and the subject deliberating about what to do at moment m:

²⁸ See among others Parfit 1984, ch. 13.

*Premise P1** [Intrinsic value in virtue of phenomenal character]

The instantiation of E by S' at m' is intrinsically bad.

*Premise P2** [Transition from intrinsic value of experiences to intrinsic value for R-chains]

If the instantiation of an experiential property P is intrinsically bad, then the instantiation of P is intrinsically bad for any R-chain in which it occurs.

(A1*) There is an R-chain C such that the mental states of S at m and the mental states of S' at m' belong to C.

On the Parfitian view, A1* is a consequence of the assumption that—in the situations relevant for our argument—S is identical to S'. We can now deduce the desired result in the following way:

Consequences:

(C1*) The instantiation of E by S' at m' is intrinsically bad for any R-chain in which it occurs. [From P1*, P2*]

(C2*) S at m has a *de se* chain-relative reason to avoid the instantiation of E by S' at m'. [From C1*, A1*, P3*]

(C3*) It is rationally permissible for S at m to be motivated by her *de se* chain-relative reason to avoid the instantiation of E by S' at m'. [From C2*, P4*]

If one concedes all the premises of the Parfitian parallel argument, then one must admit that it successfully shows that identity does not matter. On the Parfitian view, C3 must be reinterpreted as a trivial consequence of C3*. The Parfitian will then have shown that C3 is true only in virtue of the agent and the later experiencer being causally connected in the appropriate way and not in virtue of their assumed identity. This is so, she may claim, since C3 can be adequately motivated by and deduced from C3*, which in turn follows from the proposed premises and features of the example without any use of the assumption of identity between the agent and the later experiencer. Furthermore, the Parfitian has an elegant error theory at her disposal. Our premises (which crucially involve the assumption of identity) follow from Parfitian premises which do not involve identity. From a Parfitian perspective, all these premises are acceptable when understood as motivated by the corresponding Parfitian premises. This is where our main mistake lies, from a Parfitian point of view: we overlook that the real justification of each of our premises can be expressed in a framework which does not speak of transtemporal identity of subjects. According to the Parfitian, no agent ever has a reason for avoiding later suffering by a later existing experiencer in virtue of being himself or herself that later experiencer. Furthermore, if it is, in a given case, rationally permissible to be motivated to act in order to avoid that later experience, then this is never so in virtue of the agent's identity with the one who will undergo the experience.

The Parfitian philosopher can even accept the stronger conclusion we argued for in section 6. She can thus agree with us that it is a form of practical irrationality not to be moved by one's *de se* subject-relative reason when one has the relevant information. However, she can apply again the same kind of strategy and argue for a result that implies ours but that can be deduced from premises that do not involve identity nor knowledge about identity with the subject suffering at the later moment. Let us first recall the premises P4' and P5 used in the version of the argument for the stronger result concerning irrationality:

Premise P4' [Transition from *de se* subject-relative reason to irrationality; presupposes P3]

If S knows that X is bad for her and is not motivated by her corresponding *de se* subject-relative reason to avoid X, then S thereby exhibits a form of practical irrationality.

Premise P5 [Knowledge of the content of P2; presupposes P2]

S knows that for any subject it is bad (time-independently) to instantiate an intrinsically bad experiential property at some moment.

The Parfitian who follows the strategy just mentioned can then replace these premises by the following assertions:

Premise P4' [Transition from de se chain-relative reason to irrationality; presupposes P3*]*

If S knows that X is bad for an R-chain she belongs to and is not motivated by her corresponding *de se* chain-relative reason to avoid X, then S thereby exhibits a form of practical irrationality.

Premise P5 [Knowledge of the content of P2*; presupposes P2*]*

S knows that for any R-chain it is bad to instantiate an intrinsically bad experiential property at some moment.

The epistemic assumptions (characterizing the kind of situation at issue) concerning identity and intrinsic value can be replaced as follows:

(A2*) S knows at m that there is an R-chain C such that her mental states at m and S's mental states at m' belong to C.²⁹ [Presupposes A1*]

(A3*) S knows that the instantiation of E by S' at m' is intrinsically bad. [Presupposes P1*]

The desired result can then be deduced as follows:

(C3*) S knows that the instantiation of E by S' at m' is intrinsically bad for any R-chain in which it occurs. [From A3*, P5*]

(C4*) S knows that the instantiation of E by S' at m' is intrinsically bad for an R-chain she belongs to. [From C3*, A2*]

(C5*) S knows that the instantiation of E by S' at m' is intrinsically bad for a R-chain she belongs to and she has the corresponding *de se* chain-relative reason to avoid the instantiation of E by S' at m'. [From C4*, C2*]

(C6*) If S is not motivated to avoid the instantiation of E by S' at m' by her *de se* chain-relative reason to do so, then S thereby exhibits a form of practical irrationality. [From C5*, P4'*]

In the Parfitian framework, C6* implies our result about practical irrationality. Recall that the Parfitian can readily agree with our claim that it is irrational not to be moved by one's *de se* subject-relative reason if one has the relevant information. Yet she can insist that this is so only because *de se* subject-relative reasons are a special case of *de se* chain-relative reasons. Moreover, the Parfitian will suggest that any *de se* subject-relative reason is a reason to act in a particular way only because it is a *de se* chain-relative reason. In perfect analogy with her reaction to our argument for the weaker claim, the Parfitian can accept our result and yet insist that our claim concerning practical irrationality does not essentially depend on any assumption about identity or knowledge about identity.

²⁹ There might be an epistemic issue about A2* raised by a referee of this paper. How can the subject concerned know of a future subject that she stands in psychological continuity to her? We propose to simply include this assumption by stipulation as an element characterizing the kind of situation we consider here.

13. Completing our Evaluation of the Parfitian Reply

It will already be quite clear why we do not accept the Parfitian reply to the argument concerning rational permissibility. We have explained at some length our objections to P1* and P2*. We also said why we disagree with the claim (which underlies the Parfitian error theory) that P2 is a consequence of P2*. Of course it would be enough in order to defend our argument to undermine one of the Parfitian premises. But it will contribute, or so we hope, to a clearer understanding of the nature of our disagreement with Parfit if we also critically discuss the remaining premises of that part of the Parfitian reply presented above, namely P3* and P4*.

De se chain-relative reasons a person has at a given moment to avoid a possible future experience are reasons the person has in virtue of the causal relation existing between her present mental states and the future experience. However, we cannot see how such causal relations can be relevant at all for having such a reason. Why should I care about a future suffering because that suffering is R-related to my present mental states? I might care about the fate of a future person who is in the right way causally related to me because she is disposed and able to carry on my present projects.³⁰ I might even want to avoid suffering for her as far as it might prevent her from carrying on my present projects. But we are interested here in the reasons one has for avoiding suffering in virtue of its intrinsically bad phenomenal character. However, it is hard to see how such a non-instrumentally motivated avoidance of intrinsically bad experiences can be justified by the fact that they are R-related to my present mental states. Putting all instrumental reasons aside, it is not clear at all why one should care more about a future intrinsically bad experience when it is causally related to one's own present mental states than when it is not so related. We conclude, therefore, that P3* lacks any intuitive motivation.

We deny, of course, that P3* implies P3 on the reading we have in mind and we reject the Parfitian reading of our premise P3. On our understanding of *de se* subject-relative reasons, they are not a special case of *de se* chain-relative reasons. They would indeed be such a special case if such reasons for a given agent existed in virtue of the fact that something is bad for an R-chain to which the present mental states of the agent belong. But this is exactly what we deny, for we cannot see how subject-relative reasons could exist in virtue of belongingness to a given R-chain.

As far as P4* is concerned, we do not want to put it into question. If it is true that one has a *de se* chain-relative reasons to act in a certain way, then it is certainly rationally permissible that one is moved by that reason. However, since we reject the view that *de se* subject-relative reasons are a subclass of *de se* chain-relative reasons, we disagree with the Parfitian assumption that P4 is true in virtue of the truth of P4*. For this to be the case, P4* should at least imply P4 but, on our understanding of *de se* subject-relative reasons, this is not the case.

Let us finally look at the Parfitian reply considered above for the second version of our argument, the version which aims to show that it is irrational not to be moved by *de se* subject-relative reasons in the kind of situation considered. In addition to our criticism of P1* and P2*, we have two main objections against the Parfitian reasoning here. First, P4'* presupposes P3* and, as argued

³⁰ See for instance Parfit 1984 (1986: 310).

before, we do not see any good reason for accepting P3*. Second—as explained earlier as well—on our understanding of *de se* subject-relative reasons, they are not a special case of *de se* chain-relative reasons. Therefore, the result of the Parfitian version of the argument concerning irrationality does not imply our result. Hence, the Parfitian strategy fails against the second version of our argument as well.

14. Is Our Argument Question-begging?

There is little doubt that some readers will have thought all along that our argument against the Parfitian claim that identity does not matter might be question-begging. The terminology used and at least some of the premises (in particular premise P2 and premise P3) will immediately appear ‘anti-Parfitian’ to those familiar with Parfit’s views. Our response may surprise some of our readers. We would like to admit that—on some legitimate understanding of what it is for an argument to be question-begging—our argument is indeed question-begging against the claim that identity does not matter or against the Parfitian view this claim is part of. However, we do not share the maybe widespread background assumption that arguments which are question-begging (in some defensible sense of that term) are generally of no philosophical use or interest. We would like to briefly motivate this attitude.

One should remember in this context that it is no easy task to say what exactly question-begging arguments are if one wishes to avoid the absurd conclusion that any sound argument falls into that category.³¹ According to a common preliminary understanding, a question-begging argument presupposes or uses its conclusion already in its premises. Obviously, the concept should not be restricted to trivial cases where the conclusion explicitly occurs in the premises. Any definition of what it is for an argument to be question-begging must cover cases where the conclusion is presupposed in the premises in an implicit and hidden manner. Now, an easy way to say what it is for a set of given assumptions A_1, A_2, \dots, A_n to implicitly presuppose a certain claim C is to say that $A_1 \dots A_n$ taken together are incompatible with the negation of C . But this is just to say that $A_1 \dots A_n$ imply C . Obviously, on such a naive understanding of ‘implicit presupposition’, any sound deductive argument would be question-begging. In order to serve its purpose of characterizing an argumentative mistake, one needs to develop a notion of ‘question-begging’ based on a different understanding of ‘implicit presupposition’ or one must replace it in one’s account by something else.

There is no room here to get into that interesting methodological debate—and we do not need to do so for what we would like to say. We would like to motivate our conviction that one need not worry too much whether a given philosophical argument is question-begging. For that purpose, no specific account of what it is for an argument to be question-begging is required. Our point is the following: it is a vice for an argument to be question-begging if its primary purpose is to convince the person or theorist the argument addresses. It is not clear at all, however, that arguments serving different purposes must not be question-begging. We would like to suggest that philosophical arguments often or even normally do not have the primary function or purpose to convince one’s philosophical opponent.

³¹ See Lippert-Rasmussen 2001.

When an argument is used with the primary purpose of convincing, then the premises should—even upon careful reflection—be such that the addressee can hardly deny them. Ideally, they should already be part of what the addressee believes anyway. The argument's purpose then is to show that the addressee must admit the conclusion as well given her own convictions. If an argument's primary purpose is to convince, then the conclusion should not be implicitly assumed in its premises. Otherwise, or so one may argue, the addressee—in the interesting case where she is not already convinced of the conclusion—will not be disposed to accept the premises. These remarks would certainly require further clarification but one can surely defend the following claim: arguments whose primary purpose is to convince the addressee of the truth of their conclusion are unlikely to do so (at least for a rational addressee) if they are question-begging. This is why it is commonly assumed that question-begging arguments are to be avoided. An account of what it is to be question-begging should respect the pragmatic constraint of capturing those cases where the conclusion is in some way 'contained' in the premises of the argument. For that makes the argument of no use to convince other thinkers that they are rationally forced to accept its conclusion in the light of their own views.

That being said, we are convinced that the value of exchanging arguments in philosophy does not exhaustively depend on their potential to convince philosophical opponents that they must abandon their view and accept a conclusion that is incompatible with their theoretical outlook. Even though philosophers do and should strive for the truth, the main point of the exchange of arguments in a philosopher's daily work is not to strive for general convergence of opinion among philosophers working in the same field. Philosophers are not naive enough to always expect that the arguments they develop in their writings will contribute to a substantial convergence of philosophical opinion. The real purpose and the intellectual value of exchanging detailed arguments in philosophy is not primarily their potential to convince the opponent (although this may happen occasionally) but rather their contribution to other aspects of philosophical progress. Elaborated arguments in a given field may contribute to our understanding of what really is at stake when theorists deeply disagree about a specific issue. They may contribute to clarifying the landscape of coherent views with respect to a given problem. They may uncover interesting interrelations between issues in distinct fields of philosophy. They may help us to understand the intuitions underlying our natural convictions and the way they are interconnected. They may help us to see why we tend to find a given view intuitively attractive or rather counter-intuitive. All these purposes may well be served by an argument which has little or no potential to convince the opponent because its premises are neither part of what the opponent already accepts nor part of what she tends to accept upon reflection nor part of what she is rationally required to accept. There is no reason to assume that a question-begging argument cannot serve all these and further valuable purposes and thereby contribute to philosophical progress.

If these reflections are on the right track, then one should not worry too much whether the argument we develop in this paper is question-begging. Even if it is, we hope that reflection on that argument and the various ways a Parfitian may respond can contribute to some of the purposes mentioned. For instance, we hope that it might contribute to a better understanding of what is really at stake in the dispute as to whether identity matters for prudential rationality. We

also hope that it sheds some light on the deep interconnections between issues about value and issues about the nature of phenomenal consciousness.³²

References

- Alvarez, M. 2010, *Kinds of Reasons: An Essay in the Philosophy of Action*, Oxford: Oxford University Press.
- Audi, P. 2012, "Grounding: Toward a Theory of the In-Virtue-Of Relation", *The Journal of Philosophy*, 109, 12, 685-711.
- Dancy, J. 2000, *Practical Reality*, Oxford: Oxford University Press.
- Davidson, D. 1963. "Actions, Reasons and Causes", *Journal of Philosophy*, 60, 23, 685-700.
- Garcia-Carpintero, M. and Guillot, M. (eds.), *The Sense of Mineness*, Oxford: Oxford University Press, *forthcoming*.
- Hale, B. and Hoffmann, A. 2010 (eds.), *Modality: Metaphysics, Logic, and Epistemology*, Oxford: Oxford University Press.
- Harris, H. 1995 (ed.), *Identity*, Oxford: Oxford University Press.
- Koons, R.C. and Bealer, G. 2009 (eds.), *The Waning of Materialism*, Oxford: Oxford University Press.
- Lippert-Rasmussen, K. 2001, "Are Question-Begging Arguments Necessarily Unreasonable?", *Philosophical Studies*, 104, 123-41.
- McDaniel, K. 2014, "A Moorean View of the Value of Lives", *Pacific Philosophical Quarterly*, 95, 4), 23-46.
- Nida-Rümelin, M. 2006, *Der Blick von innen*, Frankfurt a.M.: Suhrkamp Taschenbuch Wissenschaft. (Translation in preparation with OUP.)
- Nida-Rümelin, M. 2009, "An Argument From Transtemporal Identity for Subject Body Dualism", in Koons and Bealer 2009, 191-211.
- Nida-Rümelin, M. 2013, "The Argument for Subject Body Dualism from Transtemporal Identity Defended", *Philosophy and Phenomenological Research*, 86, 3, 702-14.
- Nida-Rümelin, M. 2017, "Self-awareness", *Review of Philosophy and Psychology*, 8, 1, 55-82.
- Nida-Rümelin, M. 2020a, "Experiencing Subjects and So-Called Mineness", in Garcia-Carpintero and Guillot 2020, *forthcoming*.
- Nida-Rümelin, M. 2020b, *Conscious Individuals and their Place in Nature: Sketch of a Theory*, Oxford: Oxford University Press, *in preparation*.
- Nida-Rümelin, M. 2020c, *Philosophical Fundamentals for Scientific Studies of Consciousness: Jean Nicod Lectures 2019*, Cambridge, MA: MIT Press, *in preparation*.
- Nagel, T. 1970, "Death", *Noûs*, 4, 1, 73-80.

³² We would like to thank the participants at the *Sharif University - Fribourg Conference*, in July 2019 in Tehran, and at the *EXRE Colloquium*, in April 2018 in Fribourg, where we presented earlier stages of this work. Many thanks to Donnchadh O'Conaill and to an anonymous reviewer for their comments on a previous version of this paper. Work on this paper was carried out as part of our research on the project *Phenomenal Consciousness and Self-Awareness*, funded by the Swiss National Science Foundation (project number 100011_166320).

- Nozick, R. 1974, *Anarchy, State, and Utopia*, New York: Basic Books (1999, Oxford: Blackwell Publishers Ltd).
- Orsi, F. 2015, *Value Theory*, London: Bloomsbury Academic.
- Parfit, D. 1984, *Reasons and Persons*, Oxford: Clarendon Press (1986, New York: Oxford University Press).
- Parfit, D. 1995, "The Unimportance of Identity", in Harris 1995, 13-45.
- Parfit, D. 1999, "Experiences, Subjects, and Conceptual Schemes", *Philosophical Topics*, 26, 1-2, 217-70.
- Rosen, G. 2010, "Metaphysical Dependence: Grounding and Reduction", in Hale and Hoffmann 2010, 109-136.