12,552 wc

The Very Present Self

Stephanie Beardman

ABSTRACT. I argue that the first-personal, deliberative standpoint of agency — and its attendant conception of the self — requires a distinctively Humean conception of reasons, according to which normative reasons are determined by one's present concerns and voluntarist commitments. In particular, rational decision-making about the sort of life one wants to live, or the sort of person one wants to be, depends solely on one's commitment-avowals at the time of decision. A striking consequence of this view, which I defend here, is that prudence (in the sense of acting with practical foresight) is not a requirement of rationality: there are no non-subjective reasons to promote one's future well-being. Both an objectivist account of reasons (and the notion of the self as equally real at all times) and Korsgaard's account of deliberative agency should be rejected because they cannot make sense of our reasons for becoming or avoiding becoming a certain kind of person, with different interests from those we have now.

Consider the prospect of getting Alzheimer's or a traumatic brain injury, being incarcerated, captured in war, fired from a job, cheated on, you or your lover get accidentally pregnant…. These are examples of unchosen future contingencies. As forward-looking creatures, beings who can worry and wonder about the future, we often do more than speculate about what might happen, and how we might behave, in such worst-case scenarios. Sometimes, we also make a decision now about what we will do — we form conditional intentions, conditional on various events transpiring.

We also often make decisions that govern future scenarios that are within our power to bring about. You decide to enroll in a music conservatory, get married, join a religious order, go to graduate school, sue your doctor, enlist in the Air Force, or adopt a child. Big Life Decisions such as these can significantly influence what sort of person you will become and what sort of life you will lead, and that's one reason such decisions (even the hypothetical ones) are not usually taken lightly. The choice of what career to pursue, whom to love, where to devote your energies and attention, all of these will strongly influence who you are, the kind of person you will become, the kind of life you

will lead, often in more or less predictable ways. Altogether, they substantially determine not only your style of living, but also whether you act with personal integrity. These are examples of the sort of decision-making that is widely taken to be based on your values and desires, and thus up to you to settle for yourself.

We can also explicitly commit to being a certain kind of person, in the sense of settling on traits and policies that might guide or provide a standard for other decisions. Do you value the familiar or the new, conformity or rebellion against authority; what is your attitude toward risk; and how thoroughly do you like to assess probabilities? Are you primarily a truth-seeker, willing to risk triviality and falsehood, or do you prefer to risk ignorance and skepticism to avoid falsehoods? Related to the issue of truths: is the preferred policy to kill the messenger, or to set aside the painfulness of learning certain information to discover more? Are you willing to sacrifice yourself for the greater good?

These are decisions that help determine the sort of person you are and want to become (or avoid becoming). You may not engage in this sort of deliberative process often, or even explicitly, but it's an intelligible phenomenon. Prima facie, as long as you're not making mistakes in reasoning or in apprehending relevant non-normative facts, such decisions are rational. It's natural, moreover, to think that the rational basis of such decisions involves a notion of commitment, understood not as a state such as a belief or a desire, but as an act of volition. (I use the term 'commitment' rather than 'will' because the features to which I want to draw attention are not present in all decisions.) Commitment makes it possible to determine who you are in the face of an unknown and unknowable future. This is true even for conditional commitments, as well as for decisions that bind your future self (such as those involving contracts and possibly advance directives).

In this paper, I want to focus on decisions with a certain structure, involving the possibility and prospect of value conflict within a self over time. As an example, consider a young socialist, who stands to inherit an enormous fortune when she turns fifty. Because of her political ideals, she intends to give her riches away to the poor. But she knows that over time, her ideals may fade. So, to guard against the possibility that she won't later do what she cares about most now, she puts the future inheritance into an irrevocable trust and signs it over to a trustworthy organization that will automatically

distribute the funds to the poor.[1] The central question is this: How should we take our future reasons into account in deciding what to do now, and to what extent is it rational now to make decisions that conflict with the preferences and interests of one's future self?

I believe that commitment provides the key to answering this question. I'll explore what's involved with this phenomenon, and in particular, what implications it has for our notions of self, the relevance of personal identity, and the rationality both of setting up foreseen inter-temporal value conflict and of standards for addressing such conflict. As I'll argue, this intuitive picture is at odds with a common understanding of reasons, and in particular, with the normative force of *prudential reasons* — future-based considerations that have present normative force independent of one's present desires and other attitudes.[2]

My account of rational motivation is broadly Humean. There are many forms of Humeanism, but the salient element for this account is that one's future has only as much normative status as is granted by the present self, through one's present attitudes. Irrationality, on this view, is confined to various kinds of internal inconsistency or getting the non-normative facts wrong. Those who disagree believe that some subjective states — whether desires, values, or volitions — are objectively and intrinsically irrational (because they are not responsive to real reasons). They hold that normative facts constitute reasons that hold independently of an agent's present subjective

[1] The structure of this example closely follows that of Parfit's Nineteenth Century Russian, except without the promise-making and its bindingness that figure in his case. (1984, 327) I mirror here, rather, the young nobleman's attempt to pre-commit his future self. He wants now, as a young socialist, to give away the estate he will inherit, even knowing (fearing) that later he will reject his earlier ideals and want to keep the riches for himself (or, presumably equally distressing, to devote them to some other ideal or cause that's opposed to his current ideals). I return to a fuller discussion of this case in section 5.

[2] Acting with prudence means appropriately attending to future considerations. Thus, 'prudence' here refers to acting with practical foresight and not necessarily with self-interest, as when one acts in the best interest of one's children's future without regard to one's own interest. On this view, it is irrational to disregard future interests or allow them to be outweighed by insufficient present considerations. For example, if you believe that next year you will want to go to Istanbul – though at the moment you may be feeling entirely sedentary and have no desire for future travel – then you have a prudential reason to save up for the plane fare. There are other definitions of prudence; perhaps most common is that it is concerned with considerations of subjective well-being, as opposed to morality. I am more interested in the normative relations between present and future reasons, and so want to focus on the future-directed understanding of prudence.

attitudes (and what follows from those attitudes).[3] I will argue that the deliberative standpoint, the agential perspective, provides support for a Humean account of rational motivation, and against an objectivist account. This is because, on my view, a rational agent is not free to do without a conception of herself as here now. I embrace the typically Kantian view that a proper conception of agency constrains what might count as an adequate theory of reasons and shows that we should reject prudence as a principle of rationality.[4]

Sections 1 and 2 will sketch the relevant notion of commitment and its relation to the deliberative perspective on agency and to the notion of the diachronic self, respectively. Section 3 argues for the irrelevance of the notion of personal identity. In section 4, I argue that commitment requires the possibility of inter-temporal conflict. It turns out that the capacity to respond to the prospect of conflict of this sort is central to our notion of autonomy and agency and therefore is rational if anything is. The overall argument has four parts. a) Commitment in the relevant sense is a form of avowal, and as such, requires the deliberative stance. b) To have rational authority over the future self, the deliberative perspective requires that you think the future self is you. c) If only the bare notion of the self is guaranteed from the deliberative perspective, along with one's evaluative and affective attitudes on one's attributes, then there are no resources *from this perspective* for recognizing the normative authority of the future as such beyond one's present concerns for that future. d) Commitment, in the sense identified here, requires the possibility of inter-temporal conflict. Section 5 turns to an objection from the requirements of morality; my response relies on the significance of the self/other asymmetry and the conception of the diachronic self. With several pieces in place, section 6 develops more fully the positive account of the normative authority of the very present self. An objection concerning the fact that action takes time is addressed in Section 7.

[3] A few examples of those who have an objectivist notion of normative reasons are Dworkin (1996), Scanlon (1998), Raz (1999), Nagel (1986), Shafer-Landau (2003), Enoch (2011), and Parfit (1984 and 2011).

[4] This paper is a descendant of chapter 3 of my dissertation *Affective Deliberation: Toward a Humean Account of Practical Reasons* (2000, Rutgers).

1. The deliberative perspective

A focus on deliberative agency reveals two ways that a subject can think of her future (essentially: third- or first-personally), and these shed light not only on how a subject conceives of herself over time, but on the nature of present-tense commitments about future prospects. There are two standpoints one can take toward one's conduct: practical (or deliberative) and theoretical (or descriptive). The first is the subjective perspective of the agent as such, the second, the impersonal observer perspective of empirical science. When we *commit* in the relevant sense, we can do so only from the deliberative stance, which means that commitments are avowals, rather than reports about one's mental states. To distinguish the relevant phenomenon from other related ones, we can call these *commitment-avowals*. Avowals express one's mental states — you can avow your beliefs, desires, or intentions. In contrast, reports merely describe one's mental states and have the same propositional content that could be affirmed from a third-person standpoint. Avowals are essentially first-personal, they are authoritative without relying on evidence or being mediated by judgments, and they are a way of making up your mind. As I apply the term to commitments, commitment-avowals are expressions of the will that often have as their object being a certain kind of person or living a certain sort of life, but their objects could also express deep attachments and values.

Though it's best brought out by examples, here is a rough characterization of the phenomenon.[5] Commitment in my sense is an act of the will, though not all intentional acts are commitments in the relevant sense. Commitment-avowals identify and produce meta-reasons, reasons we have to avoid, promote, or continue having certain sort of reasons.[6] They can rationally justify further actions. The values they express function to

[5] The term 'commitment' is invoked in myriad ways. Aspects of the general phenomenon we're focusing on has been variously spoken of (with subtle and sometimes significant differences between them) in terms of 'wholeheartedness,' 'second-order volitions', 'self-governing policies', 'categorical desires,' 'putting one's will behind an option,' and so forth (for a small sample from a large literature, see Frankfurt 1988, Bratman 2007, Williams 1973, Chang 2009). I say more about the nature of conditional avowals about future prospects in "Prudence as a Synchronic Principle," in progress.

[6] On the capacity of the will to create value, Joseph Raz writes,

frame the decision scenario, as well as to render certain considerations deliberatively relevant while excluding other potential considerations as irrelevant. They both exclude and include certain considerations that would otherwise lack or have normative status. Thus, they can render certain considerations that would otherwise constitute reasons as counting for nothing at all (rather than merely outweighing them). I will illustrate these points with examples in section 4.

The primary significance of the deliberative perspective, for our purposes, is the self-conception that it requires of agents when making a commitment that creates meta-reasons. I'll argue that this way of understanding commitment-avowals, from the deliberative perspective, forecloses objectivist accounts of reasons, and especially those that hold it's irrational to thwart future interests. That is: *If* you think the rough picture above is correct and that this is a plausible way to understand rational commitment, *it follows*, from that picture, that it can be rational to make plans that undermine one's future interests. This is because of the nature of agency. When we look at the deliberative, agential, conception of the self, we see not only no rationally necessary commitment to the future self's well-being, but also that this agential authority must allow for the rational possibility of acting imprudently. Otherwise, the very core commitments that are central to deciding who you want to be would be ruled out.

> One creates values, generates, through one's developing commitments and pursuits, reasons which transcend the reasons one had for undertaking one's commitments and pursuits. In that way a person's life is (in part) of his own making. It is a normative creation, a creation of new values and reasons.... But it is not like the change of reasons which is occasioned by loss of strength through age, or the absence of money due to past extravagances. Rather it is like the change occasioned by promising... its making transforms one's reasons, creating a new reason not previously there. Similarly, the fact that one embraced goals and pursuits and has come to care about certain relationships and projects is a change not in the physical or mental circumstance in which one finds oneself, but in one's normative situation. It is the creation of one's life through the creation of reasons. (1986, 387)

However, I disagree with Raz's view that the primary role of commitment is to resolve indeterminate or underdetermined choice situations. (cf. Raz 1986, 388-9; 1999, 10, 99ff.). For me, the will can always upset the balance of reasons; it's not merely relied upon as a pragmatically useful element to decide or resolve what to do, but is an essential and ineliminable part of agency. I do not think one can recognize this truth while limiting the role of the will to whatever reason itself doesn't already determine. (As I argue below.) This is also one of my central disagreements with Chang's account of the will, see footnote 11.

2. Commitment and the diachronic self

What is the source of one's felt authority to act on one's own behalf? The first point is obvious and identifies the barest, most minimal condition on all volitions in order for them to count as acts of the will. The real basis for the felt authority of a commitment is that I think the future it is determining is mine (the first-person pronoun here might help make the meaning clearer). It's *my* future. That I think so is a necessary condition of my having the authority to make arrangements that determine my future reasons. That I think it's *me* — that it's possible for me to regard my future from an engaged, subjective perspective — is a condition of properly exercising the concern I have for myself now and at other times past and present. To have the rational authority to make decisions for myself, I have to assume diachronic self-sameness from the deliberative perspective. This is the basis of the first-personal felt authority, and it's what gives one a right to make decisions for the future self.

It's important to understand that such identification is wholly in the present; it is your self-conception now. You are only one self over time, on this view, and there are not multiple selves strung together or differently related, from the first-person perspective that I'm pointing to here. Now, of course, you might not think of your future in this way — and, in my view, that is not only psychologically, but rationally, permissible. If we stay focused on the agential, present, first-personal self that is actually making decisions, from that point of view, I can only act with prudential, egoistic, concern for the one I take myself to be. This doesn't ensure, of course, that I will in fact be that self in the future. (This is why I don't endorse the view as a metaphysical account of personal identity.) But it is what governs my deliberations and decisions now, in the present moment — the only moment in which I, myself, can act.

One can also think of one's future from the theoretical, third-person perspective; this involves seeing oneself as others might see one. While this perspective isn't the source of one's felt authority to act on one's own behalf, it doesn't counteract that authority as long as one recognizes that that future *is* one's future. After all, one can see oneself third-personally and also know that she is oneself. But from the theoretical perspective, it is possible to consider one's future self as not oneself, or not sufficiently

oneself in normatively relevant ways. In that case — if one considers the future being in a third-personal way while not really believing or accepting that it is oneself — then on the face of it one would have no rational authority over that self, and so would have no business determining her life. Here two conflicting intuitions emerge. On the one hand, there might seem to be no special reason to consider the preferences of the future self, or to defer to her wishes, since she would have no more connection to the present person than would anyone else. On the other hand, one's close causal and physical embodiment relations to that future self might make it the case that one has special custodial duties to consider the effects on her well-being in one's present decision-making. (We will return to these third-person questions in Section 5.) In either case, if one does happen to think that there *are* in fact two completely separate selves at issue, on the face of it, we'd have a different set of questions that are, in principle, no more (and no less) interesting than considering how one should treat other people.

In any event, it's notable that people generally care what happens to the future self, however changed. But here's the thing. Identifying the future self as oneself does not rationally require that one thereby *cares* for that self. Affective detachment toward one's future self is rationally permissible (on my view), but it does not necessarily undermine the rational authority of the present self over the future. So a tension arises when one can foresee that one will become relevantly *dissimilar* to who one is now (or, to who one wants to be or has reason to be in certain circumstances one can foresee). What gives one normative authority to bind one's future self? From the deliberative perspective, the answer is obvious: *It's me*. All that's necessary to meet the most minimal necessary condition for their felt authority is that one think the future self is one's same self.

It will be helpful to distinguish two notions of the self. One is the thin notion of the self as nothing more than the bare indexical 'I'. I'm gong to claim (without argument, at this point) that the basic condition for the felt authority of commitment, from the agent's perspective, is just the indexical 'I' that's identified directly and not on the basis of any criteria. The second notion of the self is thick: it has descriptive content and consists in the attributes and other traits of a person, her personality, character, dispositions, physical, and relational properties: what she is like as a person, her hopes

and fears, whether she is brave or sad, introverted or kind, unflappable and patient, cynical or snide, stubborn or easily distracted.[7]

We can further distinguish different attitudes we can take toward this robust notion of the self: we can evaluate it in thick and thin evaluative terms, we can approve or disapprove of it, and we can take an affective interest in and care about this trait conception of the self. You can like or dislike, tolerate or spend hardly any time at all reflecting on yourself. You can be disgusted, horrified, compassionate, or exceedingly pleased with yourself — and relieved or angered by, proud of, averse to, and so on, concerning these first-order traits (as well as second-order ones such as self-esteem). All these attitudes, as well as conceiving of the self in this way, of course requires too that you think you are the one who bears these attributes. So these attitudes and judgments will assume the bare indexical. However, and importantly, it is not necessary to have an imaginative connection to your future self — to know, say, what things will be like, in order for it to count as *your* future; you can still retain a practical identification with that self without that. When it comes to the indexical I, neither the imagination nor evaluative attitudes have anything to get a grip on. It follows from this that you can think a self is you without endorsing her or knowing anything about her. Now, can you also, rationally speaking, think she's you and not care about her, not be concerned for her well-being? It's psychologically possible to care about the bare indexical, perspectival, self.[8] But *must* you, on pain of irrationality? If all that's required for the first-personal authority of diachronic rationality is the indexical self, it seems a further argument would be needed to show that you must, in order to be rational, also care about that self.

Note that not all decisions involve self-conceptions or representations of the self in either sense. Indeed, most don't. But commitment-avowals do, because they are about the sort of person one wants to be, so they involve trait self-conceptions essentially. Moreover, from the deliberative perspective, you cannot engage with the process of

[7] There is a nice discussion of this distinction in Nichols (2014). In that paper, Nichols emphasizes the ways in which the thin 'I' is implicated in episodic memory; I want to focus here instead on the observation that this indexical notion is necessarily present with all commitment-avowals. A focus on the agential perspective yields further truths about the indexical self, in addition to those that are discovered when exploring its role in episodic memory.

determining who you are (and therefore, of changing or retaining various of your attributes) without assuming the bare indexical self.

3. The irrelevance of personal identity from the agential perspective

Whether there *is* such a thing as the thin self is a matter of controversy. It may be an illusion. However, all that is needed at this stage in the argument is the acknowledgement that such a notion is in play — that such a self seems to exist — from the deliberative, first-person perspective. It's only from the deliberative standpoint that, phenomenologically, you're tied to a certain notion of yourself over time. All that is needed is that it seems to you that your self (in a simple and unified sense that secures your particularity and is conceptually independent of the traits you possess) persists through time. Whether the notion of a self in this sense is in the end coherent or intelligible, let alone real, does not affect the central point that one relies on such a prima facie conception when making diachronic commitments. We as agents have to assume a certain conception of ourselves when we act, but don't have to assume that that's the true story from the impartial, third-person perspective. Indeed, there may be good reason to think the self is not real. *Even so*, because we cannot do away with the deliberative perspective entirely, we cannot help thinking of the indexical self as real while making the sort of big life decisions under discussion. The fact that the self might be an illusion is one that we must forget, or at least, that we act as if it isn't true, whenever we commit truly.

The metaphysical nature of personal identity itself is irrelevant from the deliberative perspective (though one's beliefs about identity or survival might of course influence what's rational to do)[9]. This is because the sense of the indexical self can come

[8] Just as one can accept responsibility or feel responsible for her acts: from the deliberative perspective, praise and blame for one's acts attaches to the indexical self.

[9] This point is underscored by David Velleman when he rejects the idea that there's a single all-purpose notion of the self and observes: "Unfortunately, philosophers sometimes assume that the qualities essential to a person's sense of who he is are in fact constitutive of who he is and therefore essential to his remaining one and the same person, numerically identical with himself and numerically distinct from others." (2006, 5) One doesn't have to agree with the particulars of Velleman's account of the self to agree with the general conceptual distinction. I endorse his understanding of the self as a "psychological relation that holds between subjects who are on

apart from whatever (essentially non-perspectival) criterion of identity is proposed — such as any theory based on trait continuity and connection over time, the biological conception of identity, and so forth. In order to emphasize how tenacious is the sense of the indexical self, it's worth drawing attention to a striking dissimilarity among indexicals (such as here, now, and me). Consider episodic memory, the recollection of past experiences.[10] When we remember our experiences,

> … instead of *now*, we get a specific *when*, and instead of *here*, a specific *where*. By contrast I* [the mental representation of the indexical self] doesn't get replaced with descriptive content. This feature of the I* in episodic memory enables the sense of identity that accompanies typical episodic memory. When I remember my first kiss, this comes with a sense of identity with that distant and qualitatively different individual. [Nichols 2014, 144]

Episodic memory and episodic forecast rely on the thin, indexical, conception of self. Moreover, as Nichols (2014) argues, trait continuity is neither conceptually nor empirically necessary nor sufficient for bare identification of the self as such. The first-personal sense of the indexical self doesn't depend on those. Finally, from a deliberative perspective, the indexical self isn't itself robust enough to ground a metaphysics of the self.

However, from the theoretical perspective, a normative connection between selves at different times *would* have to rely, implicitly or explicitly, upon a theory of personal identity. That is, if you were looking at everything purely from the theoretical point of view, you would have to take into account metaphysical conditions, such as that of personal identity, concerning whether you're the same person over time (and to what degree, perhaps) — how else could you identify the relevant individuals? But you don't do that (and can't do that) from the deliberative perspective itself. When the bare

---

first-personal terms." (2006, 193) According to this terminology, my past and future selves refer to those whom I can think of reflexively, in the first person. This notion of the self is to be distinguished from the notion of personal identity, understood as a metaphysical relation that holds "among those who are one and the same person, and I'll describe them as the same person rather than as selves." (193) This has the added advantage that it doesn't assume a self is a person.

[10] In contrast to semantic memory, which allows one to retrieve knowledge of facts, episodic memory is the process of retrieving (or, in the case of episodic foresight, projecting) experiences. The same episodic processes are involved in thinking of past and future experiences as part of the subject's own past and future.

indexical is on the scene, there are no 'degrees' of oneself, in the way that there might be of identity.

Some may find this fact, that the deliberative view of the self does not depend upon a theory of personal identity, somewhat alarming, since it means that what is rational to do is not tied to what sort of metaphysical relation obtains between various temporal parts of the self. One might suspect I say the question of what constitutes identity over time is irrelevant because I do not think that there is a deep truth that secures personal identity. And it is true that I do not think there is a metaphysically deep fact of the matter about what constitutes personal identity. But I also think the deliberative perspective forces this view.

It's perhaps worth noting that I do not claim that personal identity is in general not relevant, of course. In particular, it's relevant to (i) the third-person stance one takes toward oneself, as well as (ii) how others should treat one. These domains both raise significant and difficult moral questions. Regarding (i), I allow that the theoretical perspective is always available. I claim only that that perspective will require dissociating from the agential perspective, or considering the future self outside the scope of one's avowals. Regarding (ii), nothing about how third-parties should treat one with respect to value conflicts over time follows from the deliberatively-based arguments offered in this paper. Whether pre-commitments, past directives, promises, and the like should be honored, enforced, released or whatever — once a person has changed enough in relevant respects — is a topic outside the scope of this paper.

4. The possibility of inter-temporal value conflict

The central counterintuitive consequence of this view is that there are no rational constraints on what one can be committed to (though there might be psychological constraints, given the sort of creatures we are, and there certainly are moral constraints). I'm happy to bite this bullet: this is a consequence of the theory. However, the metaphor here isn't entirely apt. Rather than being willing to pay the costs, catching a bullet between my clenched teeth, it would be more accurate to say that, properly understood, these purported counterexamples are to be embraced and hailed, for they

are indications that the normativity of our decisions resides within our subjectivity, which is always only here now.

The tradition of biting the bullet here is well-developed in the literature. Many Humean defenses are intent on showing that the purportedly counterintuitive consequences of the theory are actually rather intuitive, once all the details of the particular cases have been spelled out. I call this the 'argument from plausibility.'[11] In addition to this classic strategy, there is an under-appreciated second approach in the Humean's arsenal that comes from recognizing the normative authority of the deliberative self.[12] What follows in this section employs both of these arguments.

The first point to make, which cannot be emphasized enough, is that we need to include the present attitudes about the entire state of affairs, including one's present beliefs about the future. Normally, one will care a great deal about one's future interests, and not want to frustrate them. We fret and deliberate and try our best to figure out how to make things go well for ourselves over the long term. Very few of us, who are not

---

[11] I pursue such a strategy in my discussion of Warren Quinn's example of a person who wants to turn on radios for its own sake, where I argue that the Humean is not burdened with a purely functional account of desire, and that desires can be rationalized by recognizing the role of affective states ("Reasons and Other Turn-Ons," unpublished ms.). For more thorough arguments from plausibility, see Richard Hare's discussion of ideals and moral fanaticism (1963, 193-9 et passim, especially ch.9) and Sharon Street's discussion of what she calls ideally coherent eccentrics (2009). Williams has also emphasized the gist of the point often: "If an agent really is uninterested in pursuing what he needs; and this is not the product of false belief; and he could not reach any such motive from motives he has by [a sound deliberative route from the motivations he already has]; then I think we do have to say that in the internal sense he indeed has no reason to pursue these things. In saying this, however *we have to bear in mind how strong these assumptions are, and how seldom we are likely to think that we know them to be true*." (1981, 105, italics added.)

[12] Commitment-avowal is a distinct activity of the will. The inclusion of explicitly voluntarist elements in my Humeanism makes my account what Chang calls a hybrid account of the sources of normativity. This means that I reject the prevalent assumption that practical normativity has a univocal source, that is to day, that "*all* practical reasons have their normative source in irreducibly normative facts, or in the agent's desires, or in her will. (Chang 2009, 246) However, my hybrid account does not have the hierarchical structure central to Chang's hybrid account, according to which "voluntarist reasons — reasons whose normative source is an act of will — depend on there being *non-voluntarist* reasons — reasons whose normative source is not an act of will. More specifically, it holds that an agent cannot have a voluntarist reason unless her non-voluntarist reasons have 'run out'." (246) In contrast, on my view, one's will is always at the ready, standing by to make oneself over, to lay a claim to, and to decide, who one is — the sort of person one is and will be. While this capacity is not always engaged, it is never rendered impotent or silenced by non-voluntarist reasons, and it never requires, as it were, permission to create or act, thereby constituting new reasons. Thus, as I argue in the text, one's will can usurp considerations that, in the absence of the act of will, might otherwise have counted as reasons.

mistaken about any of the non-normative facts available to them, don't exhibit extreme care and concern for our future selves. Because of this, in fact, for beings like us, cases of rationally frustrating future interests will usually involve something of enormous significance — an important value, driven by one's particular passions and life projects — that is at stake. We'll see examples of this in what follows. These are the passions people have that can often make life itself worth living; it's unrealistic to think they will be like counting blades of grass or preferring pain on Tuesdays. And, as I'll argue below, these passions will not necessarily take into account one's future interests, even, one's foreseeable future passions. But first, let's stay with the present point about the psychological plausibility of certain purported counterexamples.

One can deny that future reasons have independent normative force in the present while maintaining a deep concern for one's future self and warmly desiring her well-being.[13] Another way of bringing out the point is that it is consistent with the Humean view to have a zero rate of time discount, so that future reasons behave normatively on a par with present reasons because one now cares about those future reasons. In fact, as I've been stressing, one does generally bear a normatively intimate relation to one's future self; one cares about one's future, one wants it to go well. Because of this concern, one would normally be quite interested to know what one's future desires and beliefs would be.

For example, say I've seen a movie about Antarctica and am so impressed that I form the desire to go there next year. Knowing the nature of such enthusiasms of mine, I know that my desire will not last. It might seem that, according to the Humean account, I have a reason to prepare for Antarctica. But we should keep in mind that a reason is a prima facie, defeasible, reason. So my desire not to look foolish with my arctic gear and no place to go (after my desire for a visit to Antarctica has vanished) can override a reason to prepare for Antarctica. The present desire to travel to Antarctica would be a reason to pack my arctic gear only if I believe that packing will lead to my going to Antarctica.[14] For this, I must believe that there is some chance that I will go. But in the

[13] Here again the language of "future self" becomes especially strained. For it is one's own self as a temporally extended self that is cared for. See Richard Foley (1978) for a desire-based account of prudence that assumes such a conception, along with a present desire for self-respect.
[14] Johann Frick calls this the 'Non-Wastage Requirement': "If I believe that doing M is costly, and that the sole potential value of doing M is its instrumental value as a means to doing E, then

given scenario, I have independent reason to believe that my present desire to travel will be extinguished before I have a chance to satisfy it. No matter what preparations are made, I know from the beginning that I will not go. And so I don't have even a prima facie reason to pack, on the Humean account.

However, this clarification won't address the central anti-Humean worry, brought out by those cases in which my preparing now to do something I will have no reason to do later *is* a means to (in the sense of increasing the chance of) realizing my present ends. Take, for example, the case of now wanting to have a Mercedes in five years, and so making the payment now for a Mercedes to be delivered to me in five years' time. Wouldn't it be irrational for me to make such a payment if I know that in five years I will not want the Mercedes, no matter how great my present desire to have it then? Or take the case of a 20 year old who does not want to live past the age of 60, regardless of what his 60-year-old self may want. Does he have a reason to now take a delayed-action poison that will kill him at 60?

In response, it's important to highlight the fact that in the first scenario, I also now have a belief that in the future I will be getting something I do not want, and in the second, the 20-year old would have the belief that he may be making it the case that he dies at a time when he may very much not want to die. So, it cannot be left out that the total state of affairs in the first case (similar thoughts apply to the second) is not just

---

it is rationally permissible for me to intend to do M only if I believe that it is sufficiently likely that, if I do M, I will do E." ("Self-Knowledge and Structural Rationality," ms. prepared for LSR seminar). I believe this is a condition on all instrumental reasons. Note the second condition of this principle rules out cases in which we set ourselves aims or goals we don't think we can attain, but for some reason other than actually attaining them. For example, a good way to jump as far as you can, even if you know you can only jump about six feet, is to try to jump ten feet. I do not want to dispute intuitions here, and there are the obvious problems surrounding the Toxin puzzle lurking. The point is just that we may have other reasons – reasons not directly related to the goal of being on an Antarctic expedition – to prepare to do now what we will have no reason to do later. So, it may be that one can rationally prepare to do some act A in order to accomplish some goal other than doing A. Suppose I know that I won't have a reason to go to Antarctica because I know that I will be dead before I have a chance to finish all the preparations necessary for me to take such a trip. Perhaps I would be happier spending the time I have left to live preparing for the trip than sitting around watching television shows about such trips. The aim of trying to keep my energies focused on something that interests me gives me a reason to prepare to do what I know I will later have no reason to do. (Of course, to say this is rational is not to say that it is psychologically possible. Even though one might have reason to continue preparing for the trip, it may not be possible to form the intention to do what one knows will be

that I satisfy my present desire to have a Mercedes five years later, but also that I get a Mercedes that I will not then want. And of course, I may now have an attitude about that future state of affairs. I may not want my future self to have an unsatisfied preference (to not own or have bought a Mercedes). Indeed, that is highly likely. In that case, the present desire for a Mercedes, or whatever, would be, for me, dependent on the continuation of that desire. But, too, it may not be so conditional, and in that case it will be rational to make the payment now for future delivery. It's possible then that I may now want such a total state of affairs to be true in five years, or I may be indifferent to it. So, I may have no reason to prepare for what I know I will have reason to do in the future. As these examples indicate, however, it's even worse than that: given the fact that one's present and future values can come apart, I may actually have reason now to try to prevent my future self from bringing about what she will have reason to bring about.

There are some cases in which anti-Humeans agree it is rational to plot against one's future self: namely, when one believes one's future self will be deluded or mistaken about one's reasons. Consider the case of Odysseus and the Sirens. Odysseus knew that no sailor had lived who had heard the Sirens sing. Their songs were so enchanting that men found them literally irresistible, and died in the shallow waters trying to join the singing Sirens. Knowing that he, too, would find the Sirens irresistible, and wanting both to hear them and to live, Odysseus had his sailors bind him to the mast of his ship while they stopped up their ears. Odysseus guarded against the possibility of acting upon future desires that were likely to thwart his larger plan. In this case, it seems rational for Odysseus to now prepare for something – sailing straight past the Sirens – that he later will think he has reason not to do. And similarly, it seems rational for Odysseus to now make it impossible to do something – joining the Sirens – that he will later appear to have reason to do. That is, it is rational for him both to take the means to some end that he will not have, and to prevent the means to some end that he will have later. Odysseus is protecting himself against an irrational future impulse, and so is now rational in binding himself to the mast.

But the cases with which we are concerned are not like this. In the cases in which the future self's reasons really are reasons at that time, the objectivist thinks it is

---

impossible to do. There are also related issues here concerning the rationality of thinking there

obviously irrational to plot against this future rational self. As Nagel, for example, puts it: "One does not have reason now to ensure the frustration of what it will be rational to do in the future." (1979, 40 fn.1) So let's confine the discussion to cases in which present and future reasons conflict, but of which it is not right to say that either is mistaken or irrational. Consider the following example. On that same journey in The Odyssey, Odysseus sails near the island that is now thought might be Corfu, where Alcinous, king of the Pheacians, rules. Alcinous has a daughter, Nausicaa. In Homer's story, the princess finds Odysseus on a beach, and they are lovestruck. For our example, however, we can imagine that Odysseus has heard about Nausicaa before his arrival at the beach, and that he is certain that Nausicaa would find him attractive. (She has heard of his legendary feats, has found none other to be worthy of her, and so forth.) And let us suppose that he also knows it is likely that he will find her as worthy as people have said she is, that they will be mutually attracted to each other, and that he will want to stay with her and perhaps marry her. At the same time that he can see all of this, he is in love with Penelope. Contra Homer's version of the story, though, in the one we are constructing here Odysseus has not married Penelope yet, nor has he asked her to wait for him, nor does she expect him to return for her (he's been gone so long, already). So, at present, Odysseus has no obligation to Penelope, though he loves and misses her. And he knows that, if he visits Corfu, he is likely to forget Penelope and to fall deeply in love with Nausicaa. Life there would be very good, living under the protection of Nausicaa's father the king, succeeding him to the crown – better than his life back in Ithaca. Odysseus' future self would not be deluded, nor would he be irrational in staying with the princess and forgetting about Penelope. In other words, that future self really does have reason to stay with Nausicaa.

But his present self has no such reason. Right now, the thought of losing his love for Penelope pains him, and he wishes to avoid the distraction of beautiful and beguiling princesses. He has a meta-desire that certain of his desires and emotional attachments not change. It even makes sense for him to change his plans, given what he knows about the likelihood of falling in love with Nausicaa, and to avoid laying foot on the beaches of Corfu (binding himself to the mast, if need be). This would be to prevent his future self

---

is some small chance one will survive, even when one believes that one won't.)

from developing a rational desire to stay with Nausicaa, but what of it? Is that possible future reason even a prima facie reason for the present Odysseus? Is it even a consideration to be weighed in favor of sailing toward Corfu? It seems to me that his possible future reason provides no present reason at all — not even a pro tanto reason — and it's rational for Odysseus to stay clear of Corfu. This is an example of meta-reasons being generated by present commitments.

One might worry about the counterfactual nature of the imagined case I've presented. One might think that the story shows only that when present choice can influence the future course of events, then present commitments determine one's reasons. This would already be a large concession to the Humean, since it would mean that, to be rational, one wouldn't have to weigh the reasons for all relevant prospects. But I believe that the same is true of future reasons that are not counterfactual.

In addition to the example of the young socialist presented at the beginning, another common and compelling sort of conflict between present and future reasons can be seen in the case of a young neuroscientist, ambitious and devoted to her profession and making important discoveries, who wants to undergo an irreversible surgical procedure that would make it impossible for her to conceive a child. She also fully believes evidence that, let's assume (and contrary to fact) indicates that the majority of women who make such decisions live to regret them all things considered. She knows that there is nothing special about her that makes it unlikely that this will be her fate, too: regretful that she made such a decision and desperate to conceive a child. But she wants now to thwart that possibility. In other words, she is plotting against (what she is now certain will be) the rational desires of her future self.[15]

[15] The case raises the question of whether and to what extent it can be rational to court future regret. Harman (2009) argues that reasonable preferences ("I'm glad I did it") formed after an act is performed don't necessarily tell us what is reasonable to prefer prior to that act, and thus don't necessarily bear on the question of whether we should, all things considered, perform that act. She leaves open the possibility that this might also be true for reasonable aversions formed after an act (if such exist). I agree with her rejection of the principle that "if a preference is reasonable in the future, then it would have been reasonable in the past (if nothing is learned in the meantime)." (193) As she argues, the fact that it's possible to form preferences that are reasonable *conditional* on the existence of certain attachments doesn't show those same preferences to be reasonable before the formation of those attachments. She frames her point as a rejection of a Reflection principle for desires, analogous to van Fraassen's principle of Reflection for belief (which she grants might be true). For an argument against the reading of epistemic Reflection she endorses (which is focused on the question of whether the future self is

My response to this case reflects the position I urged above concerning Odysseus. This woman does not seem to me to be irrational in deciding now to prevent her future self from becoming a parent, given her present self-conception and devotion to being a professional success without motherhood. She cares about her work so much that she wants to ensure she does not succumb to a strong future (rational) temptation that could distract her from her future capacity to continue her research. However, one might think that the case of the neuroscientist is different from that of Odysseus in a way that points to a potential problem with the latter case. The neurobiologist knows that she will later regret not bearing a child, no matter what. In contrast, Odysseus knows only that *if* he goes to Corfu, he will fall in love with Nausicaa, and that if he doesn't, he won't. One might say that in the former case, the researcher's future reason to conceive a child is now a *settled* fact, in the sense that it is true in all of her possible futures that she will have reason to conceive. The intuition behind the present objection is that there is a normative difference between blocking oneself from acting on an actual future reason one knows will obtain (in such a way that ensures future regret) and preventing the formation of a possible future reason. The logical distinction here is real and perfectly clear. But normatively speaking, the continuum between these cases seems to me to be quite fuzzy.[16] The same thing that rationally authorizes you to prevent you from having future reasons permits you to block acting on settled future reasons. Although the cases are distinct, the present commitment has force both in not developing reasons and in generating reasons to block future reasons, even foreseeable ones. The claim that future settled facts have normative force while possible future reasons do not strikes me as arbitrary. Furthermore, restricting future reasons to those that will obtain in all possible futures — that is, those that are already settled at the time of deliberation — would rather severely restrict the set of future reasons that one would be rationally required to take into account.

---

a reliable expert or authority), see my (2013). I defend a Reflection principle for reasons in "Prudence as a Synchronic Principle," in progress.

[16] It's important to keep in mind, for example, that one way to prevent one from having certain reasons is to exploit 'ought implies can,' and make it impossible to act on those reasons. But I do not want to make too much of this observation. The important point is the one that follows in the text.

Moreover, the fact that a present reason will be rationally extinguished doesn't give one reason now to abandon it. We're talking about things that matter a great deal – the kind of person one is, the sort of life one can inhabit from the inside, as it were, as one's authentic self. Here we arrive at the second point to make in response to those troubled by the counterfactual nature of the case of Odysseus.

There is, in fact, an independent virtue to considering a case that involves counterfactuals such as the fact that Odysseus would have reason to marry Nausicaa if he were to visit her island. The benefit of focusing on counterfactual reasons, of the sort that Odysseus has, is that it highlights a phenomenon that I think the anti-Humean cannot accommodate.[17] This is the phenomenon of deliberating about the kind of person one wants to be. This sort of deliberation necessarily involves the evaluation of counterfactual reasons. This is a process that involves not merely discovering, but also deciding, who one is. In the case of the neuroscientist, she is deciding now to be the sort of person whose greatest passion is her career, a passion which she wants to honor even in the face of predictable (and all-things-considered) regrets. She is making this commitment-avowal *not*, however, because she thinks her current reasons *outweigh* the reasons given by those regrets (and, per hypothesis, in the future they won't outweigh them; the regrets will loom larger), but because those future regrets count for nothing now; *they are not normatively motivating* (nor should they be, given her current reasons). Her commitment to ensure that she continues her career without certain distractions is not conditional on the persistence of her desire for such. Likewise, Odysseus has a present reason not to have a future reason to be with Nausicaa rather than Penelope. The objectivist cannot recognize the reasons that are generated by such present meta-desires (i.e., those in which the preference change is itself the object of a commitment-avowal). More generally, that view cannot make sense of our reasons for becoming or avoiding becoming the sort of person we want to be.[18]

---

[17] This might be too strong. My worry is that anti-Humeans who think that prudence is rationally required cannot recognize the full extent of the arena available to rational deliberation. One of the central things one gets to decide for oneself, in my view, is whether one will be a prudential sort of person.

[18] Here I echo a point of Christian Piller's, that an objectivist account of reasons as 'timeless,' such as that defended by Nagel (1970), cannot make sense of the project not to become a person whose reasons derive from interests that are opposed to one's current interests. If you believe that you will nevertheless become such a person, then according to such a theory, "you cannot

Finally, one might object that perhaps all that the examples show is that one needn't worry about commitments that are not subject to rational constraints, such as those stemming from emotions and attachments such as love. This objection seeks to restrict the scope of practical reasoning, this time by excluding reasons that are generated by emotion that may not be subject to rational constraints. My response here is that such a restriction on a theory of reasons is unacceptable – it would exclude *too many* of what we consider to be bona fide reasons. But the other point to make is that this objection misses the driving force of the example, which is that the love for Penelope is not just a feeling but a commitment.

5. An objection from morality: abusing the power of the present

So far, the discussion has been about whether prudence is a requirement of rationality. But to many, the examples considered trigger the worry that the present self can be a sort of thug or despot with regard to future selves. One wants to know, how can such thuggishness be rationally justified if it is morally unjust or wrong? Is the present self rationally justified just because it can set the agenda for what is to be done, being that she is in the present and the future self, well… is not? If so, then locating the source of rational normativity in the present seems to allow violations of morality. The worry here is that the very present self has an unfair advantage, as it were, and isn't morally justified in abusing it.

---

sensibly pursue your project of *never having reasons of a certain kind*, because if you believe that you will acquire some interests then you have to admit that you *already have reasons of the kind you do not want to have*" (1997, 198). This is exactly right, though I disagree with Piller's requirement that the agent think there is some small chance that one will succeed in resisting an expected preference change, and also with the view of persons implicitly assumed in his (and Nagel's) talk of 'future selves.' Piller's defense of Humeanism invokes desires that one not have the future desires one expects to have. His argument does not bring in the notion of counterfactual futures, which as I emphasize here are central to uncovering the lessons to be learned from the phenomenon of such meta-desires. It also relies on the thought that there is "some chance" that committing to a plot to undermine the future (and unwanted) shift in desires will succeed in preventing that shift (cf. pp. 196ff.). As I argue in the text, in the case in which the future is fixed (or, in which the future reasons are "settled"), I do not think one needs to believe there is some chance of preventing this in order to make rational one's present commitment to undermine the future. On my view, it is in the nature of certain commitments

As should be clear by now, this way of phrasing things as a conflict among 'selves' at different times altogether exemplifies the wrong way to conceptualize the issue. From the deliberative perspective, that's not the correct model. Furthermore, the objection relies on a view that what's rational is constrained by substantive reasons, which the Humean would deny. (We may be appalled at one who completely disregards her future reasons, but that's a substantive moral standpoint; nothing in the Humean position prevents one from regarding such a stance as rational.) But let me grant for a moment this way of conceiving of things. If one is tempted by this line of reasoning, it seems that one must be relying upon moral intuitions concerning people at different times. For example, in *Reasons and Persons* Parfit argues that the basis for prudence is moral: "we should claim that great imprudence is morally wrong. We ought not to do to our future selves what it would be wrong to do to other people." (1984, 320)

In his example of the 19th c. Russian, Parfit describes two alternative views about the diachronic self that exemplify this disjointed way of thinking of the self over time. In the story, a young socialist signs a legal document that can only be revoked by his wife, ensuring that the estate he will inherit will be given to the peasants. He then says to his wife: "'Promise me that, if I ever change my mind, and ask you to revoke this document, you will not consent.' He adds, 'I regard my ideals as essential to me. If I lose these ideals, I want you to think that I cease to exist. I want you to regard your husband then, not as me, the man who asks you for this promise, but only as his corrupted later self. Promise me that you will not do what he asks.'" (1984, 328) In due course, when the man's ideals have faded and he tries to release his wife from her earlier promise to his younger self, she does not think he has the normative standing to release her, because he is not the same person to whom she made the promise. Parfit explains: "We may regard some events within a person's life as, in certain ways, like birth or death. Not in all ways, for beyond these events the person has earlier or later selves. But it may be only one out of the series of selves which is the object of our emotions, and to which we apply some of our principles." (328) As a statement of people's attitudes, this is sometimes true. Whether the wife in this case is behaving reasonably or well — and whether her husband really is not the same person to whom she made the promise so that she can never be

---

that intransigence in the face of even purported full belief (that the future will bring unwanted

released from it — are topics for another time. For now, I want to draw attention to a more plausible reaction that Parfit leaves out. As he interprets the case, the later man does not have the moral authority to release his wife from the promise she made to his earlier self, because he is not sufficiently similar to the person he was then. But there is another conclusion that can be drawn. Noting (as Parfit does) that the reason the man now wants to release his wife from her earlier promise is precisely because his values have changed, and thus his two selves are not sufficiently similar, suggests that it is the younger man who did not originally have the moral authority to bind his future self (by getting his wife to promise to honor his past values no matter what) to begin with.

Moral objections can be raised within the context of each of these competing models. I'll first consider the case in which the future self is (normatively on a par with) another person distinct from you now. Next, I'll give some reasons to think that this is not in fact the model we usually invoke in deliberative agency. Finally, I'll consider the case in which the future self counts as yourself.

The first point is that, from the deliberative perspective (as discussed in section 2), *if* we really don't — perhaps because we cannot — think of the future self as *my* self, then we lose our agentially-based authority, and we have no rational authority to govern that self. Well, who else might have authority, if not you? Good question. If the future self will be incapacitated, there might be good reasons for it to be you, just as you have some say over how your property or your corpse is treated after you die. But if you believe it will be just a radically different self — as in the case where you have been destroyed in a teletransporter and entirely replaced with a different self, for instance — then, on the face of it, perhaps you should hold off on long-term planning, beyond doing things that would set that self up with the means to be able to act with autonomy. (Let's leave out how paternalistic one would be permitted to get; my answer to that is just this: however paternalistic one is permitted to be with other people, those same rules would apply here.) But notice how rare it is to think like this. In real life, we do not think of ordinary changes we undergo — including the value shifts involved with the Russian nobleman, the neuroscientist, and the examples involving Odysseus — we do not think of these futures as not us in the bare indexical sense. We think we are uniquely placed to

---

reasons) is rational; some might even call it behaving with integrity.

influence the life that these beings lead, and not merely because we can. Rather, from the first-person, agential, perspective, this is our life — from the present perspective, *one* life — and we are now determining the meaning, the shape, the value, that this life will have as a whole.

So Parfit's discussion of the identity of the nobleman concerns only the trait conception of self. The indexical self seems untouched; indeed, it's plausible to think that the older version remembers his younger self well, rueing the pact his wife made with him long before. Another reason for thinking that we generally think the same self persists is brought out by contemplating the following question. *Why* is the future change in values perceived as such a threat to oneself, when it is? Precisely because I think it will still be me, and *I don't want to be that person.* If I thought that the future self was normatively akin to someone else, why would I care? Otherwise, one should rationally, and from a self-interested perspective, care no more about it than one would about the state of having died. And even though one sometimes says about some future prospect, "If *x* happens, I might as well be dead," what one means is that *x* is actually worse than death, because it *damages* (and doesn't merely replace) oneself.

Now to turn more fully to the second reading, on which the future self is just oneself with more or less changed attributes. I don't want to deny that we might have duties to ourselves. But whatever duties we might have do not suddenly appear when we're looking at the future — if they exist, they must apply consistently to the self that is accessible in the present from the deliberative perspective. If it's permissible (all things considered) to harm oneself now (by, say, smoking) — and I think it is — then it's permissible to do the same to one's future self. Here, we can invoke a well-known asymmetry between self and other: it is morally permissible to impose burdens on oneself that would not be permissible to impose on others.

Most people take it for granted that we have a kind of discretion about how we treat ourselves that we don't have toward others. Indeed, this gets at the core of what's important about the distinction between persons. The boundary between persons is a fundamental fact underlying morality, and its significance is brought out in the fact that there is an asymmetry in what one owes others in contrast to what one owes oneself. Correspondingly, one has greater discretion in how one treats oneself. This seems to me true even if there are duties to the self. (And I think the discretion goes both ways: not

only may one treat oneself worse than one treats others, but one may also require more of oneself morally than one may require of others.) It's permissible for a person to impose sacrifices within her life that it's not permissible to impose on others for others' benefit. We have a greater *discretion* how we treat ourselves.

I'm relying here on the intuitive notion that there is such an asymmetry regarding what's permissible between self and other. To argue for it, or to say more about it, would require that one settle large questions about morality (which is the right normative theory? What is the relation between morality and rationality? Are there moral duties toward the self, and if so, what are they?) All I need here is a foothold in the claim that some sort of asymmetry is likely to hold. Certain views about morality will rule out such an asymmetry; I will for now assume that they are wrong and rely on the sense that there is a robust intuition, often taken for granted and operating in the background, to the contrary.

6. The normative authority of the deliberative self

My argument in these pages is grounded in a conception of the self as a decision-maker. It relies heavily on the fact that, from the deliberative standpoint, all decisions are made in the present. One's future has only as much normative status as is granted by the present self, through one's present attitudes and intentions.[19] This view is rooted in the very phenomenon of agency: a rational agent is not free to do without a conception of herself as here now.

It may be useful to contrast my voluntarist Humean account with that defended by Christine Korsgaard, since there are central elements of her view with which I agree. In her (1989), Korsgaard provides a Kantian defense of a view of the self as intimately tied to agency and practical reason. Her main point in that paper is to argue against Parfit's view of the self as a series of experiences, and on this point I agree with her: I reject a purely experience-based account of the self. I also agree with Korsgaard's claim that "a focus on agency makes more sense of the notion of personal identity than a focus on

experience. There is a necessary connection between agency and unity which requires no metaphysical support." (115) We both reject the common philosophical view that the assumption of a metaphysically robust self, a psychological subject that persists through a series of experiences, undergirds the intuition that one has a special relationship to one's future and past selves. As discussed in section 3, one need not take a stand on the metaphysical question of personal identity in order to settle questions concerning the rationality of action.

However, I part ways with Korsgaard in holding that it is because of *that* — the fact that the connection between agency and the unity of the self requires no metaphysical support — that the present looms large in the proper account of reasons and justifiably looms so large. One can only act in the present. One can now decide to act in the future, and one can do things now that will bring certain states of affairs about; but one cannot now act in the future. (The same is true for avowals; they can only be expressed in the present.)

Moreover, the notion that reasons distributed over time have the same normative force as reasons at a present time isn't really at home with a Kantian notion of respect for persons. Counting all moments equally in evaluating a choice at any particular moment is at odds with treating a person as an end in herself: It is only because the agent herself values her future that her future reasons should be taken into account by others. While I think Korsgaard might endorse the spirit behind this claim, I'll argue in the next section that her view of temporal unity produces a mistaken view of prudence as rationally required. But support for my claim that taking prudence to be rationally required is not compatible with respect for persons can be found in Korsgaard's remarks on paternalism. She opposes (most) paternalistic coercion to ensure prudential behavior for the following reason. "Even in the most straightforward case of compensation, where a burden is imposed on a person from which she herself will later benefit," she writes,

> compensation by itself does not do the justificatory work. That I will be compensated may give me a reason to accept a burden; it does not give you a reason to impose one on me. The only reason you have to impose one on me is that *I do accept it*.... In the case of an adult, it is the [present] acceptance, not the

[19] Note, though, that it is misleading even to put things in terms of 'present' and 'future' selves, as if they could conflict or war against each other. As Bernard Williams put it, "a man's future self is not another self, but the future of his self" (1981, 11).

> [future] compensation, that does the justificatory work. [130, I have added the essential modifiers of "present" and "future."]

The impositions of present burdens, even when offset by future benefits, can be justified only by the actual consent of the subject. To my mind, this claim would be true even if all parties know for certain that in the future, once the burden has been imposed, one would have been glad for the trade. As Korsgaard says, it is not the future (in this case, compensatory) reason that is doing *any* justificatory work at all — rather, the justification of making sacrifices (or incurring any type of burden) for one's future is that one now accept that sacrifice. These burdens will not be uncompensated, since you will be benefiting from their consequences. But you will only see it this way if you now want to promote your future interests. If you don't, and if you see the burden as uncompensated, then you may not think that you have a good reason to be prudent — much less, a good reason to be forced to be so.

Korsgaard offers two different reasons to support the conception of the self as unified (at a time, and over time). Since my view is in alignment with only one of those, it's worth laying out our disagreement. Her first reason is "the raw necessity of eliminating conflict among your various motives." (110) The second reason concerns the "standpoint from which you deliberate and choose," the way you think of yourself as *you* who are choosing. (111, 119) We disagree about how the story unfolds from that standpoint, but the thing to note here is that my own account relies only on this second sort of reason. Korsgaard categorizes both reasons as "practical," but doing so mixes up two senses of the term that should be distinguished. The first is "practical" in an instrumental sense. The idea is that in order to be able to act, one has to resolve inner conflict. She thinks that this sort of consideration presents reasons for thinking you are a single subject. The second reason is "practical" in the sense of agential, and that is the only ground for my own position. This reason is not instrumental but is rather constitutive of agency — we cannot but think we are a single subject, from the deliberative standpoint.

7. An objection: time and prudence

I'd like to consider a prominent objection to my view, concerning the fact that the present is, in effect, too short a time to get anything *done*. Any result of a present act is owned by a future self, even if it be only a few seconds away. One might think that this rules out any 'dated' conception of the self – to act at all, it seems, one must be moved by one's future reasons. On this view, the observation that all action occurs over time supports the notion that prudence must be a requirement on agency.

But it is a mistake to think that the objectivist view is correct just because of the fact that action takes time. Prudential reasons have as their object some future state of affairs. It would be missing what is special about such reasons to claim that *all* reasons for action (except constitutive ones) must be prudential, because the end that is aimed at always happens at some future time from the initiation of the action. The sort of dissociation from the future that some objectivists worry about is not the same as the very radical sort of dissociation that would be involved if one did not think that one's existence in a few moments were as real as one's existence now. Perhaps it's true that an agent must think of herself as persisting over time to some extent. But this is not a conception of the self that's sufficient to secure an anti-Humean account of prudence. Furthermore, the difference in self-conceptions here is not merely a difference in degree.[20] When we speak of our future well-being, we have in mind something somewhat indeterminate, but nevertheless definitely quite a bit beyond the next few moments, or (usually) even the next few days or weeks. In disagreeing over the proper account of prudential motivation, both sides rely on an intuitive and fairly robust notion of 'future-directed' action.

[20] There is empirical support for the intuition I am defending in this paragraph. For example, studies run at the Center for the Study of Brain, Mind and Behavior at Princeton University by Jonathan Cohen and his colleagues indicate that different neural systems in the brain influence whether one will choose a smaller, earlier reward over a larger, later reward – or vice versa (see for example McClure et al., 2004). For my purposes, I am less interested in the inconsistency the researchers try to account for (preferring a later reward but choosing an earlier smaller one), than in the fact that there seems to be a categorical shift within the subjects when thinking about the near-term as opposed to the far-term. The limbic system (which is involved in emotion) is more active when considering what to do in the immediate term (such as next week), while the prefrontal cortex (where more complex calculations take place) is more active when considering what to do in the future, say a year or a month from now.

In her argument for the diachronic unity of agency, Korsgaard draws too tight a connection between prudence and the fact that action takes time. She begins with some unproblematic observations, such as the fact that "[s]ome of the things we do are intelligible only in the context of projects that extend over long periods. This is especially true of the pursuit of our ultimate ends…. On a more mundane level, the habitual actions we perform for the sake of our health presuppose ongoing identity." (113) She points out that the present self is not necessarily concerned with the quality of present experiences, or with present satisfaction. In her terms:

> In order to make deliberative choices, your present self must identify with something from which you will derive your reasons, but not necessarily with something present. The sort of thing you identify yourself with may carry you automatically into the future; and I have been suggesting that this will very likely be the case. (113)

Quite so. The present self can most definitely have plans and projects which carry her into the future, and give her reasons to make sacrifices now. And yes, most lives have this. But is it rationally required that all agents have such conceptions of their lives?

Korsgaard says, "insofar as I constitute myself as an agent living a particular life, I will not […] oppose my present self to future ones. And so I do have a personal reason, whether or not I also have a moral one, to care for my future." (126-7) I agree that one does not need a further reason to care for more than the quality of present experiences, if one does happen to care for more than that. In fact, I would say that there is no reason at all to care even for the quality of one's present experiences, if one does not. So, as a general matter of fact, one will tend to have a reason to care about one's future. But nothing rationally requires that I do "constitute myself as an agent living a particular life" *in such a way* that I must care about my future. This is the step that is missing in Korsgaard's argument. She is possibly mistaking what is a very common – and important – feature of living life as an agent, with something that is rationally required of all agents. Let's look at one more passage.

> [T]he choice of any action, no matter how trivial, takes you some way into the future. And to the extent that you regulate your choices by identifying yourself as the one who is implementing something like a particular plan of life, you need to identify with your future in order to be what you are even now. When the person is viewed as an agent, no clear content can be given to the idea of a merely present self. (113-114)

Here Korsgaard makes no distinction between immediate and long-term future. I suggested above that that is a mistake. But even allowing this here, it still would not follow that prudence is rationally required. While it may be true that "the idea of a momentary agent is unintelligible" (117 fn.), this will only be in the uninteresting sense that any action – or, more precisely, any plan for action – takes time. Korsgaard is not unambiguously correct, even by her own criteria, in saying that "no clear content can be given to the idea of a merely present self." There is, after all, one sense in which a self with merely present concerns is perfectly intelligible. And Korsgaard herself has implicitly admitted that one's concerns need not extend into the future. If an agent's "life plan" extended no farther than the next few moments, then that person's self conception would be correspondingly shortened in time.

Korsgaard might be misread as arguing that she could secure the rationality of prudence from the fact that agents have, essentially, plans (even, life plans), and that plans take time. I have argued that one cannot get something this substantive from the fact that plans have a duration. Surely one of the things open to an agent to determine for herself is whether to be a prudential sort of person or not. One who works against her future rational self is herself recognizably – and, most importantly, from her deliberative standpoint – an agent. There is nothing in that scenario to threaten her personhood or agency – which is not to say that she may not come to regret her actions.

This brings us to the largest disagreement I have with Korsgaard, and to the central feature of my own understanding of normative reasons. For all her sensitive and steadfast commitment to the deliberative standpoint, it is odd that Korsgaard does not mention (what I take to be) its most salient and striking feature: that such a standpoint must always be in the present, here now. I do not of course mean that one's concerns must be circumscribed by the present or that they must be concerns about the present; I mean only that, from the deliberative perspective, the concern itself is fully in the present. This present-based foundation is in fact partly constitutive of the deliberative perspective, as it is of commitment-avowals.

Just as we must view ourselves as free and responsible when deciding what to do, so we must also view ourselves as here now. We need not view ourselves now as equally real with our future selves in any sense that would require prudence – only, rather, in

the sense to be determined *by* the very present[21] self who is to act and decide. Moreover, it is because of some of the very same reasons that lead Korsgaard to reject the conception of the self as just a series of experiences – a conception that can be captured and exhausted from the theoretical standpoint, that I reject the claim that prudence is a principle of practical reason. I want not to overlook the temporal aspect of the two standpoints we've been evoking: the theoretical, according to which we are natural objects and part of the causal order of things; and the agential, according to which we are autonomous deliberators and actors, with the voluntary will to choose among our desires. The one is 'timeless,' the other, not merely tensed and tied to a temporal perspective, but always present in the now. It is as much always in the now as it is in the here. Can we think of ourselves as elsewhere? Of course; we can even plan (and prepare) to *be* elsewhere at another time. But the agential perspective is always here. And it is always now.

The problem of personal identity, both over time and at a time, is solved by the practical exigencies of acting. My view has an affinity with Korsgaard's in this respect. We both rely heavily on a conception of persons as agents, and we resist the reduction of agency to just another form of experience. We agree that both the conception of myself as one, and the conception of myself as existing over time as one and the same self, is just as much a feature of the deliberative standpoint as is the conception of myself as free to choose among my desires. However, I would add to these a third feature which is required by the practical standpoint: *the conception of myself as now here*. Each of these three conceptions of the self is "a practical necessity imposed" (111) upon us by the nature of agency.

8. Conclusion

In sum, there is no objectivist principle of prudence. I am not rationally required to take into account the interests of my future self in deciding what to do now. In particular, sometimes it is rational to work at cross purposes to one's future rational self, either by preparing now to do something for which one will have no reason to do when

[21] 'Present': a lovely word connoting both time and place *and* active attention.

the time comes, or by failing to prepare for something that one will have reason to do later. This follows from the practical necessity of conceiving of oneself as essentially now here.

Objectivists deny the priority of a certain conception of the self that I have argued is central to the conception of oneself as an agent. In doing so, their theory permits a kind of alienation from the decision-making self, the present self, the one without whom no decision can be made. Nagel has claimed that "[f]ailure to be susceptible to prudence entails radical dissociation from one's future, one's past, and from oneself as a whole, conceived as a temporally extended individual."[22] I am suggesting that a worse sort of dissociation looms for the person who is thinking of all reasons (including present reasons) as objective — a dissociation from the present decision-making self, the present self, the one without which no decision can be made. We must own the perspective we occupy — the view from nowhere is not a perspective from which we can *act*.

More generally, I don't see how an objectivist can make sense of our reasons for becoming or avoiding becoming a certain kind of person, with different interests from those we have now. The reasons we have for remaining or becoming the sort of person we want to be are found in our present desires and meta-desires, and commitments. Conflicts between present and future reasons must be arbitrated by the present self. We saw that Odysseus has a present commitment not to have a future desire for Nausicaa rather than Penelope. Similarly, our commitment-avowals are generative: rather than discover or posit reasons, they create them. One's present commitments to who one is and wants to be, as an evaluator of reasons, is rooted in the normative force of the deliberative, agential, self; they are products of one's active engagement from this standpoint, and they rely on an a priori commitment to one's own authority, as thinker, perceiver, valuer, planning, and desiring agent. Taking a view from nowhere attitude would dissociate oneself from this presentness, and thus disable true forms of self-commitments. What we require — and what is in any case practically inescapable — is the view from now here, the perspective of the present self.[23]

[22] (1970), 58.

References

Beardman, Stephanie (2013). "A Non-factualist Defense of the Reflection Principle," *Synthese* 190, 15: 2981-2999.

Beardman, Stephanie (2000). *Affective Deliberation: Toward a Humean Account of Practical Reasons*, Ph.D. diss., Rutgers University.

Bratman, Michael (2007). *Structures of Agency*. Oxford: Oxford University Press.

Chang, Ruth (2009). "Voluntarist Reasons and the Sources of Normativity", in *Reasons for Action*, Sobel and Wall (eds.), New York: Cambridge University Press, 243-71.

Dworkin (1996). "Objectivity and Truth: You'd Better Believe It," *Philosophy and Public Affairs* 25: 87-139.

Dworkin, Ronald (1993). *Life's Dominion: An Argument About Abortion, Euthanasia, and Individual Freedom*. New York: Alfred A. Knopf.

Foley, Richard (1978). "Prudence and the Desire Theory of Reasons," *The Journal of Value Inquiry* 12: 68-73.

Frankfurt, (1988). *The Importance of What We Care About: Philosophical Essays*. New York: Cambridge University Press.

Hare, R.M. (1963). *Freedom and Reason*. Oxford: Clarendon Press.

Harman, Elizabeth (2009). "'I'll be Glad I Did It' Reasoning and the Significance of Future Desires," *Philosophical Perspectives* 23: 177-198.

Korsgaard, Christine M. (1989). "Personal Identity and the Unity of Agency: A Kantian Response to Parfit," *Philosophy and Public Affairs*, 18: 101-132.

McClure, S. M., Laibson, D. I., Loewenstein, G. & Cohen, J. D. (2004). "Separate neural systems value immediate and delayed monetary rewards," *Science* 306, 503-507.

Nagel, Thomas (1970). *The Possibility of Altruism*. Princeton: Princeton University Press.

Nagel, Thomas (1986). *The View From Nowhere*. Oxford: Oxford University Press.

Nichols, Shaun (2014). "The Episodic Sense of Self," in J. D'Arms and D. Jacobson (eds.) *Moral Psychology and Human Agency*, Oxford: Oxford University Press.

---

[23] Acknowledgments.

Parfit, Derek (2011). *On What Matters, Volume One*. Oxford: Oxford University Press.

Parfit, Derek (1984). *Reasons and Persons*. Oxford: Clarendon Press, Oxford.

Piller, Christian (1997). "Humeanism and Prudence," in eds. K. Lehrer, & J. Marek, *Austrian Philosophy Past and Present: Essays in Honor of Rudolf Haller*. Vol. 190 Boston Studies in the Philosophy of Science, 189-202. Dordrecht: Kluwer Academic Publishers.

Raz, Joseph (1999). *Engaging Reason: on the theory of value and action*. Oxford: Oxford University Press.

Scanlon, T.M. (1998). *What We Owe To Each Other*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Shafer-Landau, Russ (2003). *Moral Realism: A Defence*. Oxford: Clarendon Press.

Street, Sharon (2009). "In Defense of Future Tuesday Indifference: Ideally Coherent Eccentrics and the Contingency of What Matters," *Philosophical Issues* 19: 273-298.

Velleman, J. David (1996). "Self to Self," reprinted in his *Self to Self: Selected Essays* (2006). New York: Cambridge University Press.

Williams, Bernard (1981). "Internal and External Reasons" and "Persons, Character, and Morality," in his *Moral Luck*. Cambridge: Cambridge University Press.

Williams, Bernard (1973a). "The Makropoulos Case: Reflections on the Tedium of Immortality," in *Problems of the Self*, Cambridge: Cambridge University Press.

Williams, Bernard (1973b). "The Self and the Future," in *Problems of the Self*. Cambridge: Cambridge University Press.