Routledge
Taylor & Francis Group

First personal modes of presentation and the structure of empathy

L. A. Paul[1]

University of North Carolina, Chapel Hill, NC, USA

ABSTRACT
I argue that we can understand the *de se* by employing the subjective mode of presentation or, if one's ontology permits it, by defending an abundant ontology of perspectival personal properties or facts. I do this in the context of a discussion of Cappelen and Dever's recent criticisms of the *de se*. Then, I discuss the distinctive role of the first personal perspective in discussions about empathy, rational deference, and self-understanding, and develop a way to frame the problem of lacking prospective access to your future self as a problem with your capacity to imaginatively empathize with your (possible) future selves.

There are a host of fascinating philosophical issues that concern our understanding of the self, its relation to the first personal perspective, and its connection to the structure and content of conscious experience.

These issues connect to work in the philosophy of language involving the nature of *de se* content and the role of perspective. They concern the role of indexicals in broader philosophical theories and the nature of the semantic content that indexicals contribute to our linguistic and conceptual representations. Further metaphysical and epistemological questions arise when we ask whether we need first personal content or other essentially perspectival philosophical tools to adequately present certain features of the world to ourselves.

Ideally, we use the semantics to capture relevant metaphysical and epistemic structure. A popular philosophical thesis holds that we employ the

---

CONTACT L. A. Paul ✉ lapaul@unc.edu
[1] Arché Research Centre, University of St. Andrews

semantics of 'I', often described as the 'essential indexical,' to capture distinctive metaphysical and epistemic features of the first personal perspective needed for our expressions to have *de se* content. Capturing this content allows us to use indexical expressions to represent perspectival features of reality, in particular, features concerning ourselves and our relationship to other parts of the world.

Recent work by Herman Cappelen and Josh Dever attacks this idea.[1] Cappelen and Dever focus their attack on the philosophical significance of the first personal perspective, arguing that there is no deep philosophical contribution made to our understanding of the world by the use of perspectival or indexical notions. They argue that phrases like 'essential indexicality' and 'irreducibly *de se* attitudes' are not well defined, and deny that theories of indexicality and perspective are doing important explanatory work in philosophy.

One of the important lessons I draw from Cappelen and Dever's work is that we need to look harder at the implicit metaphysical and epistemological commitments made in various discussions of the essential indexical.[2] Their work shows that deep and difficult philosophical issues about the self and the first personal perspective are not captured *merely* by developing the semantics for indexical expressions such as 'I' and 'now'.

However, the limitations of our semantic tools do not imply that we can simply dispense with perspectival philosophical assessments of ourselves and of the world. There are important philosophical issues about how we are to understand the nature of reality and the way we understand ourselves that involve the exploration of the first personal perspective.[3] Even if we work out the semantic issues about perspective and indexicality to our satisfaction, theories of the first person and the self are philosophically rich and productive – *because the philosophical action is elsewhere*. That is: the action isn't in the semantics. It's in the metaphysics, mind, and epistemology.

Below, I'll discuss Cappelen and Dever's arguments against the essential indexical, and suggest that we can better understand the *de se* by employing the subjective mode of presentation (or, if one's ontology permits it, by defending an abundant ontology of perspectival personal properties or facts). Then I'll present some reasons for thinking there is an important role

---

[1]Herman Cappelen and Josh Dever, *The Inessential Indexical*. The book is essential reading for anyone interested in the philosophical questions concerning the nature and representation of the self.

[2]Although I'm focusing on Cappelen and Dever in this paper, there are many other important contributions made to the philosophy of language literature on this topic, such as those by Egan 2006, Ismael 2007, Ninan 2013, Stanley 2011, Titelbaum 2013, etc.

[3]I develop my views about the importance of the subjective perspective in Paul 2014, 2015a, 2015b, 2015c.

for philosophical discussions of the first personal perspective when we are thinking about empathy, rational deference, and self-understanding.

The inessential indexical

The literature on the *de se* has grown up largely around two central examples, given by John Perry and David Lewis, respectively:

> [**Shopper**] 'I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle on one side of a tall counter and back down the aisle on the other, seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally it dawned on me. I was the shopper I was trying to catch… I believed at the outset that the shopper with a torn sack was making a mess. And I was right. But I didn't believe that I was making a mess. That seems to be something I came to believe.' (Perry 1979, 3)

> [**Two Gods**] 'Consider the case of the two gods. They inhabit a certain possible world, and they know exactly which world it is. Therefore they know every proposition that is true at their world. Insofar as knowledge is a propositional attitude, they are omniscient. Still I can imagine them to suffer ignorance: neither one knows which of the two he is. They are not exactly alike. One lives on top of the tallest mountain and throws down manna; the other lives on top of the coldest mountain and throws down thunderbolts. Neither one knows whether he lives on the tallest mountain or on the coldest mountain; nor whether he throws manna or thunderbolts. Surely their predicament is possible. (The trouble might perhaps be that they have an equally perfect view of every part of their world, and hence cannot identify the perspectives from which they view it.)' (Lewis 1979, 520)

We can agree that 'I' is playing an indexical role in these examples, and that some sort of discovery is made when Perry discovers he is the shopper making the mess or one of the gods (say, Zeus) discovers he is the god on the tallest mountain. The question is whether this means there is a proper subset of 'I'–involving indexicals that are somehow special or more essential than other semantic tools, and whether this sort of indexicality is philosophically distinctive. Is 'I' *the* essential indexical?

With characteristic style, Cappelen and Dever argue that it is not. They argue for this by showing, over and over again, how the examples involving 'I'-indexicality are instances of the familiar phenomenon of generic opacity. That is, they show how examples like *Shopper*, which have been interpreted as involving discoveries of special, distinctive 'I'-facts, should instead be interpreted as involving discoveries stemming from the way that co-referential referring expressions cannot be substituted *salva veritate* in action-explanation contexts.

The central move is brought out clearly by Cappelen and Dever's Superman/Clark Kent example, a modification of *Shopper*:

> [**Superman/Clark Kent**] Pushing my cart down the aisle I was looking for CK to tell him he was making a mess. I kept passing by Superman, but couldn't find CK. Finally, I realized, Superman was CK. I believed at the outset that CK was making a mess. And I was right. But I didn't believe that Superman was making a mess. That seems to be something that I came to believe. And when I came to believe that, I stopped looking around and I told Superman to clean up after himself. My change in beliefs seems to explain my change in behavior. (Cappelen and Dever 2013, 33)

It seems that we have just the same structure here as in Perry's original example. If so, there is no so-called 'essential' contribution made by the indexicality in *Shopper*. There is a change in belief. But there's no special first-personal, essential indexicality involved. Rather, Perry's case is just an 'I'-involving example of a familiar phenomenon: the failure of substitution in certain contexts. Cappelen and Dever's discussion of Lewis's *Two Gods*, in addition to drawing out some very interesting and subtle assumptions that are particular to Lewis's picture, suggests that the same sort of failure of substitution is occurring there.

The critical discussion of *Shopper* and *Two Gods* forces us to re-think the nature of our implicit philosophical commitments in discussions of the *de se*. As I've noted, I'm very sympathetic to Cappelen and Dever's view that these commitments deserve more scrutiny than they've usually received: They successfully show that semantics alone isn't enough to support philosophical theories of the first personal perspective.

This does not mean, however, that the first personal perspective is philosophically uninteresting. Theses involved in, for example, philosophical theories of the nature of the self, first personal models of causal reasoning and decision-making, explanatory models of moral psychology, and utility assessments for social choice theory, require substantive underpinnings in metaphysics and epistemology. Looking to philosophy of language and in particular, to *indexicality* to explain all that is philosophically distinctive about the first person is looking in the wrong place.

Metaphysical facts about the self

If you are interested in the metaphysical and epistemological structure of first personal experience and the self, you should look to related debates in metaphysics. In particular, it seems natural to look to discussions of time

and temporal experience that focus on the indexicality of 'now' and the metaphysical nature of the present.

A.N. Prior famously illustrated the irreducibility of tensed expressions to tenseless ones:

> One says, e.g. 'Thank goodness that's over!', and not only is this, when said, quite clear without any date appended, but it says something which it is impossible that any use of a tenseless copula with a date should convey. It certainly doesn't mean the same as, e.g. 'Thank goodness the date of the conclusion of that thing is Friday, June 15, 1954,' even if it be said then. (Nor, for that matter, does it mean 'Thank goodness the conclusion of that thing is contemporaneous with this utterance.' Why should anyone thank goodness for that?) (Prior 1959, 17).

Now, in contemporary metaphysics of time, such irreducibility is uncontroversial. The focus in current work is not on the translation or reduction of tensed expressions to tenseless ones. Rather, the debate is over whether temporal reality (or sometimes, fundamental temporal reality) includes 'tensed' properties of nowness, pastness, and futurity in addition to 'tenseless' relations of earlier than, later than, and simultaneous with (Paul 2010). The question is whether these tensed properties exist, and whether differences in our use of language mark a metaphysical difference between tensed and tenseless properties.

On a permissive metaphysics of temporal properties, I can have nowness properties (I'm presenting a paper now) as well as tenseless properties (at 3:30 pm on 6 January 2016, I present my paper). To have a full understanding of my temporal properties, I need to grasp my tensed properties as well as my tenseless ones. Connecting back to the debate about 'I'-involving indexicals, we could do a version of this with personal properties. To have a full understanding of my self-involving properties, I need to grasp my personal properties as well as my impersonal ones. This is a natural way to treat the structure of reality as metaphysically rich enough to ground perspectival claims.

One could use this to propose that the source of the perspectivality in *Shopper* involves the metaphysical difference between impersonal facts (or impersonal properties, or whatever your favorite metaphysical, nonlinguistic kind of entity is) and personal facts (properties).[4] So the interesting issue here involves the existence of perspectival, personal facts, not something special and essential about mere indexicality.

So, what Perry discovers when it dawns on him that he is the shopper making the mess is a perspectival fact; he knew that there existed an

---

[4]Cappelen and Dever mention in a footnote that Andy Egan suggested a similar point in conversation, but they do not develop the idea.

impersonal fact concerning a shopper who had the property of making a mess. What he discovers is a new personal (perspectival) fact: *he* is the one exemplifying the property of making a mess.

On this view, 'Someone is making a mess' picks out an impersonal fact, while 'I am making a mess' picks out a personal fact, a fact that is no less real than the other objectively existing facts that ground the mental states of a conscious individual. The impersonal facts and the personal facts are different, equally real, kinds of facts. This move involves a somewhat permissive ontology on the face of it, but there are various further reductive techniques one might employ in order to minimize the ontological commitments.

Alternatively, if you want to reject this sort of permissive ontology (I'm inclined to reject it), you might argue that there are no such personal *facts*, but there are personal relations of grasping or 'knowing in a new way' various impersonal facts. This is just a preference about which kinds of entities you are allowing into your ontology. Perhaps you think epistemic relations, not nonlinguistic facts, ground the perspectival/nonperspectival distinction.

The epistemic view suggests that the way we should think about the essential indexical cases involves a standard distinction in the philosophy of mind between the descriptive mode of presentation and the subjective mode of presentation. (This fits well with Cappelen and Dever's argument that the indexicality involved is just a species of opacity.) The appeal to the subjective mode of presentation is familiar from extant discussion of the knowledge problem: before she leaves her black and white room, Jackson's Mary might know the propositions about what it is like for her to see red under one mode, the mode that presents these propositions to her in terms of the scientific facts about her mental states and other dispositions. But when she leaves her room and sees red for the first time, she is presented with the relevant facts in a new way, and learns what it is like to see red under the subjective mode of presentation (Jackson 1982; Lewis 1988; Loar 1990).

On this view, we should think of Perry, as he shops, as knowing that someone is making a mess under only one mode of presentation. But when he discovers that it is *he* who is making the mess, he recognizes the same fact or proposition but under a new mode of presentation, a subjective mode.

So, just like there's a distinctive way to understand what it's like to see red involving the subjective mode of presentation of propositions about red experience, there's a distinctive way to understand what it's like to be me involving the subjective mode of presentation of propositions about an individual. Call this the *de se* mode of presentation. We can use this to understand Perry's *Shopper* case. Just like a person can discover a new way to grasp what it's like to see red when he has a retinal implant and sees red

for the first time, he can discover a new way to grasp the proposition that someone is making a mess when he discovers that he is the source of the mess. On this approach, the distinctions between modes of presentation, especially the important and distinctive roles of (various sorts of) subjective modes, are what ground the perspectival semantics.

We can do something similar with Lewis's *Two Gods*. Lewis's view is that to believe a property-content is to self-ascribe the content. So when Zeus needs to know whether *he* is the god living on the tallest mountain, what Zeus needs to do is to knowledgeably self-ascribe a particular content.

Cappelen and Dever give a lovely assessment of Lewis's view here, showing how Lewis's approach is motivated by his larger commitments to propositions as sets of worlds and individuals as sets of properties across worlds, as well as Lewis's preference for describing worlds qualitatively.[5] They argue that:

> Lewis presents another kind of argument in favor of Essential Indexicality. He gives an informational puzzle, asking what sort of information could be gained by one of the two gods when that god learns his location. He then proposes a specific kind of perspectival *de se* content to solve this puzzle – mental state contents are properties/ centered worlds, rather than propositions. This content is perspectival because it delivers truth conditions not absolutely, but only relative to a choice of agent/center. Lewis's … solution to his informational puzzle proceeds by combining his novel theory of properties-as-contents with a specific view of the cognitive role of those contents – that to believe a property-content is to self-ascribe the content. The role of self-ascription is little emphasized either by Lewis or in the subsequent literature. But the work of solving the informational

---

[5] I admit to puzzlement about why we should embrace Lewis's view that propositions are sets of worlds and individuals are sets of properties across worlds. For Lewis himself, it makes sense. This is because Lewis approaches this entire question through the lens of his reductive approach to modality, where we have a particular sort of ontology of possibilia, we reduce necessity and possibility to classes of possible worlds, we maximize reductive simplicity and strength, and we aim to reduce as much as possible. This is Lewis's central motivation behind taking propositions to be sets of worlds, for taking properties to be sets of individuals across worlds, and then for moving to mental state contents as centered worlds. So, because of his commitment to these reductive theses, Lewis starts big and winnows down. But if you don't share all of Lewis's reductive goals, why endorse all this machinery?

Note that the support for the reductive thesis does not stem merely from a commitment to modeling using possibilia (you can like possible world semantics just fine and reject the reductive commitments). And I don't see Lewis's machinery as giving us a natural approach to a theory of truth and meaning once it's been divested of its reductive motivations. It's unnatural as a theory of properties and as a theory of individuals: we can *model* propositions as worlds and model properties as sets of individuals across worlds for some purposes, but why take the Lewisian view that properties *are* sets of individuals across worlds? It's unnatural as a theory of propositions: why think that propositions *are* sets of worlds? These extra claims can be motivated by Lewisian commitments to a certain way of building one's ontology, again, they were well-motivated for Lewis himself, because they stemmed from his broad commitment to reducing modality to possibilia and his modal realism. But these are fairly idiosyncratic methodological commitments, and they get more idiosyncratic if we reject modal realism and shift to a view where we reduce necessity and possibility to ersatz entities. Since I don't embrace this part of Lewis's approach to ontology, I've never been motivated to accept these other features of his preferred approach. It's not clear to me why others do.

puzzle rests entirely on the self-ascription leg of the proposal, as can be seen by the fact that equivalent solutions can easily be given by combining other theories of content (such as possible-worlds content) with analogous attitudes such as self-believing.

So the important work is being done by self-ascription. My response is to suggest that we understand self-ascription along the same lines as my response to Perry. If we agree that Lewis needs an explicit account of self-ascription (and that he can't get it indirectly, or on the cheap, via functionalism), this fits very naturally with a need for distinctive modes of subjective presentation.[6] In particular, self-ascription might be what is involved when one stands in the subjective mode of presentation to the set of properties in W that is the individual in W.

Very roughly: as Cappelen and Dever point out, for Lewis, to hold an attitude involving a property-content C is to self-ascribe that attitude. So we need a substantive account of self-ascription to understand what's really going on here. One natural way I see to go (without worrying too much if Lewis himself would go this way), is to say that to self-ascribe, and in particular to self-ascribe an attitude involving a property-content C, involves standing in the subjective mode of presentation to property-content C.

If we assume that individuals are sets of properties, when Zeus discovers that *he* is the god on the tallest mountain in W, he discovers that he is a particular set of properties in W. Likewise, when Zeus discovers that *he* is the one hurling down manna from on top of the mountain, what he discovers is that he should self-ascribe a content, that is, he discovers the unique set that he is. This is a substantive discovery, much like discovering what it is like to see red when seeing red for the first time. What Zeus discovers is which distinctive set defines him. And he discovers which unique set he is, by standing in a distinctive subjective mode of presentation, the *de se* mode of presentation, to the property-content defined by that set, the Zeus-set. In sum: Zeus is subjectively presented with the property-content defined by the unique set that he is.

(On a more permissive ontology, not Lewis's and not mine, we could say that self-ascription follows when Zeus discovers a special metaphysical fact, the perspectival fact that he is the one hurling manna. Here, the important move is the self-ascription of a property involving this perspectival fact.)

Just a few more remarks on the idea I'm sketching here. First, if you make a move like this, you can say that the *de se* involves knowing-how. On a broadly Lewisian picture, this is the idea that Zeus gains a new ability rather

---

[6]Note that Lewis 1988 effectively endorses modes of subjective presentation in his reply to the knowledge argument.

than learning new information (i.e. rather than learning new facts) when he self-ascribes.

But second, there's going to be an element of the *de re* here, because purely qualitative modes of presentation won't be good enough, not for the Lewis case nor for the Perry case. I already built that in with the comments about 'the unique set' or 'the particular set' above. Having the 'that's-me' or a 'sense of myself' experience presented to you isn't enough for *de se* content, you've got to have the right individual picked out by 'me' or 'myself'. (So there's an essential role here for acquaintance or indexicality after all, even if it's a shallow role in Cappelen and Dever's sense. That is, we can agree that indexicality isn't playing an essential role in information-gathering, even if it is needed for other reasons; Ninan forthcoming).

So in a lot of ways I agree with Cappelen and Dever. I agree that the Lewisian approach needs a more robust metaphysical account of self-ascription (as opposed to some sort of deflationary or functionalist account). I agree that we need to spell out the *de se* in more metaphysical detail, and some will think this means we need special perspectival facts and a more permissive ontology. For anyone who wants to accommodate the structure of subjective experience in their metaphysics (and certainly as those who want to endorse richer temporal ontologies want to do), these are perfectly legitimate philosophical commitments to explore.[7]

But I'm not in agreement with the idea that there is nothing of philosophical significance about the first personal perspective. There is still something deeply interesting about the nature of the first person, even if the parts played by indexicality and (a thin sense of) perspective are not doing much work. Arguments against the semantic theses of Lewis and Perry don't show that there isn't lots more to say about the issues in metaphysics and mind concerning the treatment of the self and how we might regard the *de se*. (There are also further things to say about the metaphysics, especially about metaphysical theories of personal identity and the self.)

So, I take Cappelen and Dever's arguments to show that the action really is elsewhere. Indexicality isn't distinctive. Nor is the thin sense of perspectivality. But the thick sense of perspectivality is important, and may well involve an irreducible mode of subjective presentation (or, if you prefer, irreducibly perspectival facts). This makes perspective and the first person, as something in the world deeply related to the self and conscious experience, philosophically distinctive. Perhaps it isn't especially distinctive with respect

---

[7]I don't take any of this to be a direct objection to Cappelen and Dever. While I'm betting they'll be skeptical of the kinds of moves I'm suggesting, they don't argue directly against them in their book.

to the semantics, but with regard to the metaphysics and epistemology and the philosophy of mind, it is a philosophically deep and important topic.

This feeds into questions about moral psychology and practical decision-making.

Imaginative empathy and self understanding

In this section, I'll discuss two issues in moral psychology and epistemology where the epistemological importance of a grasp on one's self and on the selves of others is important. The idea is to give some reasons for thinking there are philosophically interesting issues about the self that underwrite why we should care about the first personal perspective.

Empathy

One reason why it is important to develop the metaphysical and epistemological structure of the first personal perspective is because it can connect to interesting philosophical questions in moral psychology and social choice theory.

A place where some of these interesting questions come up is the debate about empathy. One question about empathy concerns just what it is. We need to work out the metaphysical and epistemic structure of the state. There's a fairly shallow interpretation of empathy as merely 'feeling what another feels' without any substantive cognitive component. But we can identify a richer kind of empathy, what I'll describe as 'cognitive empathy', where you perform a cognitive act that allows you to first personally represent some element of another person's experience. The ability to represent in this way and draw on what is learned from such a representation has a normative role to play in moral assessment and decision-making, and is also routinely explored in the psychological literature.

Cognitive empathy is often used to establish a basis for our tolerance or understanding of another's perspective.

> …reflect on how you know that someone else's physical pain is a bad thing. It may be that you know perfectly well what pain is, and know that some other people have it, but don't particularly see why that should concern you, even in cases where the other person is right there in front of you. You might find a guru or spiritual instructor who tells you, and occasionally reminds you, that you should be concerned about other people who're feeling physical pain. But this is very different to the usual case, in which an imaginative understanding of someone else's physical pain immediately drives a compassionate response to them. A compassionate response

in the ordinary case has an authenticity that the externally driven response does not. (Campbell 2015, 788)

One of the strongest ways to motivate tolerance for alien points of view is to argue that you should 'stand in the other person's shoes' or before you judge them you should 'walk a mile in their shoes' etc. These kinds of cognitive acts involve the exercise of epistemic capacities where we understand or grasp certain possibilities or certain facts as part of imaginatively representing the different subjective perspectives of the individuals we are judging. (I suspect that empathy is just one member of a class of cognitive acts where we open up ourselves to a new perspective to grasp something previously epistemically alien to us.)

To have the capacity for empathy seems to require a certain amount of shared experience, and to adopt an empathetic perspective, one needs to try and view the world from the experienced perspective of the other. You don't have to have had all the same experiences, but you have to somehow be able to 'try on' the beliefs and attitudes of the person you are trying to empathize with. This act of 'trying on' involves an experience where one attempts to experience or take on the perspective of the other, perhaps by attempting to partially cognitively model their perspective, or by attempting, in some properly attenuated sense, to grasp the belief or emotional structure of another through the lens of one's own perspective.

On the cognitive understanding of empathy, the empathetic task involves grasping some relevant feature of another person's first personal perspective. This grasp is understood first personally: that is, you have another person's first personal perspective, or some salient part of it, subjectively presented to you. You understand some dimension of what it is like to be that person, or how that person understands a given situation from her first personal perspective. This gives you a distinctive sort of understanding and the ability to make certain ethical and moral judgments. Note that there is no *de se* confusion here – it isn't that you confuse another person's first personal perspective or sense of self with your own. So, whatever the act of grasping another's first personal perspective involves, we should distinguish it from simply grasping *de se* content.

Drawing on the discussion of modes of presentation I gave earlier, we can use the structure I developed for the first personal perspective to enrich our account of what empathy is. Using the ontologically minimal account where we employ subjective modes of presentation, we think of the first personal perspective as a distinctive subjective mode of presenting content, the *de se* mode of presentation. Now the (cognitively rich) way to understand

empathy described above can be understood somewhat more clearly, because we can better understand the structure of the empathetic task.

When Zeus discovers that *he* is the god throwing down manna from the tallest mountain, he is presented with the content *a God is throwing down manna from the tallest mountain* in a distinctive, subjective way. Unusually for a God (and for Zeus in particular), Zeus is feeling empathetic, and so he also wants to understand what it is like for Hera to know that *she* is the God throwing down thunderbolts from the coldest mountain.[8] If Zeus wants to know this, now he has a further, rather interesting task.

On the model I sketched above, when Hera is *de se* subjectively presented with the content *a God is throwing down thunderbolts from the coldest mountain* (or she self-ascribes the relevant property) she knows that *she* is the God throwing down thunderbolts from the coldest mountain. What Zeus needs, then, is a mode of presentation that will embed Hera's *de se* mode of presentation into his own, such that he first-personally grasps Hera's first personal perspective, but he grasps it *as* Hera's first personal perspective.

So he needs to grasp a content in a complex way, namely, he needs to grasp *a God is throwing down thunderbolts from the coldest mountain* such that he (imaginatively) grasps what it is like for Hera to be *de se* subjectively presented with *a God is throwing down thunderbolts from the coldest mountain*, that is, he imaginatively grasps what it is like for Hera to know that she is the God throwing down thunderbolts from the coldest mountain. This sort of structure is interesting and complex. I doubt there is some sort of special semantic distinctiveness here. That's not the point. The point is that we need to be clear about the cognitive structure of empathy and the embedded structure of subjective modes of presentation if we want to understand what empathy is.

Empathy involves first personal understanding, and it is an important tool for moral understanding, moral motivation and moral assessment. I'm not arguing that empathy is somehow metaphysically necessary for morality. I'm arguing that it is often psychologically necessary for moral action, and also philosophically interesting: the structure of cognitive empathy involves an interesting way of embedding a representation of the first personal perspective of one individual into the first personal perspective of another.

We can see the interest and importance of empathy and moral understanding in a range of philosophical discussions. For example, it's the key to determining utility and preferences in social choice contexts, at least

---

[8]I'm ignoring the fact that I've introduced a gender asymmetry in the example. Gods have complicated gender issues anyway.

once we move past a simple behaviorist picture of revealed preferences. It connects to decision-making for others, to the extent that we rely on it to determine their preferences when we make decisions for them. It's pretty clearly the central task for Harsanyi, as we can see from his description of his influential social choice approach.

> [T]he basic intellectual operation … is imaginative empathy. We imagine ourselves to be in the shoes of another person, and ask ourselves the question, 'If I were now really in *his* position, and had *his* taste, *his* education, *his* social background, *his* cultural values, and *his* psychological makeup, then what now would now be *my* preferences between various alternatives. (Harsanyi 1977, 638)

There are many other ways an understanding of the first personal perspective connects to moral psychology. For example, in recent years, there has been more interest in giving analyses of the nature of friendship, love, and sympathy. A philosophically rich treatment of these relationships must include a developed account of how first personal experience and perspective is built into the metaphysical structure of such relations. (While such relations are not merely phenomenological, they aren't merely nonphenomenological either. Zombie love isn't love as we'd ordinarily understand it. Understanding the role and importance of grasping the first personal perspective and experience of your loved one, and developing the structure of this complex psychological attitude, requires a model of how this works.)

Self understanding

Another place where we need a rich model of the first personal perspective is when we try to understand first personal deliberation and decision-making. Prospective reasoning, or reasoning about what we should do, based on what we desire and how we could act, is an extremely important cognitive task. Swathes of the psychological literature are devoted to it.

Ordinarily, when making first-personal assessments of our preferences, we strike a balance between our first personal and third personal assessments of who we are, using memory and anticipation to represent ourselves over time and in different possible situations. Work in cognitive science suggests that performing this sort of cognitive task plays a central role in prospective reasoning.

> [These] combined observations suggest that the core network that supports remembering, prospection, theory of mind and related tasks is not shared by all tasks that require complex problem solving or imagination. Rather, the network seems to be specialized for, and actively engaged by, mental acts that require the projection of oneself into another time, place or perspective. Prospection and related forms of self-projection might enable mental simulations that involve

> the interactions of people, who have intentions and autonomous mental states, by projecting our own mental states into different vantage points, in an analogous manner to how one projects oneself into the past and future. (Buckner and Carroll 2006)

Further, prospective reasoning about our future selves under conditions of radical change suffers from a number of defects stemming from the role that experience plays in our ability to grasp what our future selves will be like, and the possibility of variation in the self that we are over time (Paul 2014, 2015b).

This brings out a need to develop the underlying metaphysics and philosophy of mind involved in the way we understand what a self is, and to better understand the way that first personal deliberation contributes to how we think of rational action and decision-making.

We can relate this back to our earlier discussions of empathy and modes of presentation of a self. One way to frame the problem of lacking prospective access to your future self is to think of it as a problem with your capacity to imaginatively empathize with your (possible) future selves. This is a kind of understanding of our future selves that we value, and if we have to make big life decisions without it, we lose something of epistemic value. In some of these contexts, we value the first person perspective, even if we don't strictly speaking *need* it, in order to make a rational decision about our future.

It will be helpful to make this clearer using a somewhat roundabout approach discussing moral deference. The debate over moral deference involves the view that there is something problematic about an individual adopting a moral belief solely on the basis of what a moral expert tells him to do. The thought is that there is something wrong with an individual's adopting a moral belief without judging and assessing the moral question himself, that is, there is something less than ideal with adopting a moral belief solely on the basis of the testimony of a moral expert. The sense of wrongness remains even if we assume the moral expert is in fact correct about the moral status of the belief.

The background assumption is that the individual lacks the capacity to assess the moral value of an act or an event on his own. For some reason, the agent cannot introspectively arrive at the right moral beliefs. And the thought is that, ordinarily, even if a certain amount of moral instruction is appropriate, we want a moral agent to employ his own moral judgment to arrive at his moral beliefs. That is, an appropriate part of the process involved in arriving at a moral belief involves a certain sort of first-personal assessment that delivers moral understanding, and gives the believer insight into why his moral belief is the right one to hold, or why the relevant action is

right or wrong. The idea is that, even if one should adopt the correct moral beliefs, it is better, all things considered, for a believer to adopt this moral belief on the basis of his moral understanding rather than to adopt it merely on the basis of the testimony of the moral expert.

Related issues arise with respect to intellectual deference and the virtue of understanding:

> Even if the student can know that the professor's judgment is true – on the strength of his professorial authority – still if this is his evidence the student has not gained insight into why it is that that conclusion is true, or that the relevant action is wrong (or right). (Enoch 2014, 9)

Finally, we can see a parallel here to the agent who is choosing to act based on advice from the scientific expert. Imagine you are faced with a big life choice, for example, the choice of whether to have a child. You have no idea what the best choice for you is: should you do it, or shouldn't you?

In this case, we assume you lack the capacity to discover or introspectively assess the value of the act on your own, and so you rely solely on the scientific expert for your beliefs about what you should do. So for this choice, you defer to the scientific expert, the sociologist or the psychologist. You make your decision of whether to have a child based solely on what he – the scientific expert – says you should do. This is what is ordinarily described as 'blind deference': you blindly defer to the scientific expert when you decide whether to become a parent. Let's also assume that the scientific expert is in fact correct about what you should do, such that his advice corresponds to what would maximize your expected utility. Thus, we will treat him as a rational expert, and deference to him is *rational deference*.

There seems to be something less-than-ideal about making the choice this way.[9] But what is it? Even if we assume that the scientist really does know what's best, there is something lost with rational deference. I suggest that, just as we think it is better for an individual to base her moral beliefs based on her moral understanding (perhaps with guidance from the moral expert), we expect an individual to base her life choices on her self understanding (perhaps with guidance from the rational expert).

The point is not that deference to the scientist creates a problem for self knowledge. Just as moral deference can yield moral knowledge, rational

---

[9]Viz., Paul (2015a). Also see Jones 1999.

deference can yield self knowledge.[10] The point is that something important is lost when one gains knowledge without understanding, which we can see by recognizing the loss when one merely has moral knowledge without moral understanding.[11]

Some think that deference, in itself, is problematic. But I'm not assuming that deference is problematic, since there is value in getting ourselves to the truth. The idea, instead, is that the unsatisfactory nature of deference highlights the value of understanding. David Enoch makes a strong case in favor of moral deference, but even he does not claim that there is no cost associated with it.[12] Rather, he argues that a need for moral deference gives us evidence that something else is lacking: there is a failure to achieve something of moral value.

> This, then, is what is fishy about moral deference. It is not that it is unjustified as grounds for action – it is so justified, indeed, some-times even required. It is not that it is epistemically unsound – here as elsewhere, expertise can ground knowledge, or help achieve any other epis-temic status. It is that forming a moral judgment by deference and then acting on it is much less of a moral achievement than forming the true judgment with-out deference, because it does not constitute the appropriate response to the morally relevant features of the case. Moral deference indicates failure to achieve something we are expected to achieve. (Enoch 2014, 27)

It seems clear that what we are expected to achieve, but fail to achieve, in cases of moral deference, is moral understanding.[13] If I understand why an act is morally right, I can use this understanding as a guide for future judg-ments, but I also exhibit a kind of moral virtue, because I have the capacity to understand why the act is morally right, that is, I grasp and employ the relevant reasons and can base my belief that it is right on these reasons (Hills 2009). More generally, understanding is distinct from knowledge, and can be instrumentally as well as intrinsically valuable in its own right.[14]

---

[10]There is an interesting connection here to a form of antirealism about rationality: I am assuming, with standard decision theory, the realist view that rationality concerns utility maximization. If we require rationality to require self understanding or self-construction, we can deny that the agent is acting ration-ally when she defers. The point is salient for the debate about rational deference as well as for the debate about moral deference. 'In general, the idea that one can understand a claim only if one independently grasps its justification conditions is a characteristic thesis of anti-realism; indeed, disagreement on this point is sometimes seen as definitive of the dispute between realist and anti-realist.' (McGrath 2011, 126.) Relatedly, if one thinks there are no facts about rationality, then the source of the problem with rational deference is that there are no facts about which one should be deferring.

[11]See Enoch 2014 and McGrath 2009.

[12]Sliwa 2012 defends the use of moral testimony to gain moral knowledge.

[13]As some of Richard Yetter Chappell's remarks suggest, (http://www.philosophyetc.net/2013/10/testimo-ny-and-moral-understanding.html) we could endorse moral testimony as an apt guide to moral knowledge even if we would not *settle* for moral knowledge without moral understanding. Also see Sliwa 2012.

[14]Hills, Kvanvig, and Pritchard all contrast understanding with knowledge and hold that understanding is distinctively valuable. Kvanvig 2003 holds that understanding is distinctive and valuable, as does Pritchard 2009 and Hills 2009.

We see exactly the same structure in cases of rational deference without self-understanding. Here, the virtues are epistemic as opposed to moral: there is epistemic virtue in acting rationally, and hence in deferring to the rational expert about what to do. But the need for rational deference arises because of a failure of a distinctive kind of achievement: a failure to achieve self-understanding. There is epistemic virtue in self-understanding.[15]

The distinctive problem, again, concerns the first personal perspective. That is, very briefly, the problem is that the agent cannot construct and represent her own first-personal causal model of how she'll evolve in response to becoming a parent. She can't imaginatively evolve her first-personal perspective forward under the proposed change, and as a result, she lacks a certain sort of epistemic capacity.

If I lack the ability to prospectively represent and imagine the nature of the experience I'll undergo, I cannot represent and imagine my responses and the first personal changes that I'll undergo as I respond to the experience, or imaginatively model my future point of view as it changes in response to the nature of the experience. Thus, I cannot project myself first personally into my future point of view. The future me, the person who I'll become, is first-personally inaccessible to me now.

Should I nevertheless defer to the rational expert when I choose? If rationality is the only epistemic goal here (or if being rational trumps my other epistemic goals) then I should defer. But in such a situation, there is a failure to achieve. When I defer to the rational expert to make my life choice, I am failing to exemplify the virtues of self-knowledge and informed self-construction that I am describing as the epistemic virtues of self-understanding. These are the virtues that we'd demonstrate if we chose with epistemic transparency.

Conclusion

Semantics alone isn't going to do the work we might have wanted it to do in exploring the nature of the first personal perspective and the self. To understand the deeper questions about nature of the self involved in philosophical treatments of the *de se*, we need to talk more about the implicit metaphysical and epistemological commitments made in various discussions of the essential indexical. In this paper, I've started a conversation about how to draw on well-established work in the philosophy of mind on

---

[15]This connects to the exchanges about using testimony without introspection to determine one's preferences in Pettigrew (2015), Dougherty, Horowitz, and Sliwa (2015), and Paul (2015b, 2015c).

modes of presentation, and on metaphysical theses about abundant facts and properties, in order to undergird theories of the first person and the self. Such theories are important for the philosophical work we want to do in many places, including work in moral psychology, metaphysics, epistemology, and philosophy of psychology.

Disclosure statement

No potential conflict of interest was reported by the author.

References

Buckner, Randy, and Daniel Carroll. 2006. "Self-projection and the Brain." *Trends in Cognitive Sciences* 11 (2): 49–57.

Campbell, John. 2015. "L. a. Paul's *Transformative Experience.*" *Philosophy and Phenomenological Research* 91 (3): 787–793.

Cappelen, Herman, and Josh Dever. 2013. *The Inessential Indexical: On the Philosophical Insignificance of Perspective and the First Person*. Oxford: Oxford University Press.

Dougherty, Tom, Sophie Horowitz, and Paulina Sliwa. 2015. "Expecting the Unexpected." *Res Philosophica* 92 (2): 301–321.

Egan, Andy. 2006. "Secondary Qualities and Self-location." *Philosophy and Phenomenological Research* 72: 97–119.

Enoch, David. 2014. "A Defense of Moral Deference." *The Journal of Philosophy* 111 (5): 1–30.

Harsanyi, John C. 1977. "Morality and the Theory of Rational Behavior." *Social Research* 44 (4): 623–656.

Hills, Alison. 2009. "Moral Testimony and Moral Epistemology." *Ethics* 120: 94–127.

Ismael, Jenann. 2007. *The Situated Self*. Oxford: Oxford University Press.

Jones, Karen. 1999. "Second-hand Moral Knowledge." *The Journal of Philosophy* 96 (2): 55–78.

Jackson, Frank. 1982. "Epiphenomenal Qualia." *Philosophical Quarterly* 32 (1): 27–36.

Kvanvig, Jon. 2003. *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.

Lewis, David. 1979. "Attitudes De Dicto and *De Se.*" *Philosophical Review* 88: 513–543.

Lewis, David. 1988. "What Experience Teaches." *Proceedings of the Russellian Society* 13: 29–57. University of Sydney.

Loar, Brian. 1990. "Phenomenal States." *Philosophical Perspectives* 4: 81–108.

McGrath, Sarah. 2009. "The Puzzle of Pure Moral Deference." *Philosophical Perspectives* 23 (1): 321–344.

McGrath, Sarah. 2011. "Skepticism about Moral Expertise as a Puzzle for Moral Realism." *Philosophy* 107: 111–137.

Ninan, Dilip. 2013. "Self-location and Other-location." *Philosophy and Phenomenological Research*. 87 (2): 301–331.

Ninan, Dilip. forthcoming. "What is the Problem of *De Se* Attitudes?" In *About Oneself: De Se Attitudes and Communication*, edited by M. Garcia-Carpintero and S. Torre, 86–120. Oxford: Oxford University Press.

Paul, L. A. 2010. "Temporal Experience." *Journal of Philosophy* cvii (7): 333–359.

Paul, L. A. 2014. *Transformative Experience*. Oxford: Oxford University Press.

Paul, L. A. 2015a. "What You Can't Expect When You're Expecting." *Res Philosophica* 92 (2): 149–170.

Paul, L. A. 2015b. "Transformative Choice: Discussion and Replies." *Res Philosophica* 92 (2): 473–545.

Paul, L. A. 2015c. "Transformative Experience: Replies to Pettigrew, Barnes, and Campbell." *Philosophy and Phenomenological Research* 91 (3): 794–813.

Perry, John. 1979. "The Problem of the Essential Indexical." *Noûs* 13: 3–21.

Pettigrew, Richard. 2015. "Transformative Experience and Decision Theory." *Philosophy and Phenomenological Research* 91 (3): 766–774.

Prior, A. N. 1959. "Thank Goodness That's over." *Philosophy* 34 (128): 12–17.

Pritchard, Duncan. 2009. "Knowledge, Understanding and Epistemic Value." *Royal Institute of Philosophy Supplement* 64: 19–43.

Sliwa, Paulina. 2012. "In Defense of Moral Testimony." *Philosophical Studies* 158: 175–195.

Stanley, Jason. 2011. *Know How*. Oxford: Oxford University Press.

Titelbaum, Michael. 2013. "*De Se* Epistemology." In *Attitudes 'De Se': Linguistics, Epistemology, Metaphysics*, edited by A. Capone and N. Feit, 307–325. Stanford, CA: CSLI.