Accidentally about me

Abstract: Why are *de se* mental states essential? What exactly is their *de se*-ness needed to do? I argue that it is needed to fend off accidentalness. If certain beliefs – e.g. nociceptive or proprioceptive or introspective beliefs – were not *de se*, then any truth they achieved would be too accidental for the subject to count as knowing. If certain intentions – intentions that are in play whenever we intentionally do anything – were not *de se*, then any satisfaction they achieved would be too accidental for the subject to count as intentionally acting. How states hook onto their referent is relevant in a systematic but underexplored way to whether they non-accidentally achieve their aim – truth in the case of beliefs, satisfaction in the case of intentions. In the relevant cases, the way of hooking on to a referent needed to avoid being accidental is the way a *de se* state hooks on to its referent.

(0) Introduction

De se states seem to play an essential role in certain cases of knowledge and intentional action (hereafter, action). It's easy to feel the pull of the cases that seem to show this. It's surprisingly difficult to say what explains this. This paper offers an explanation.

I begin with two cases:

Cognitive: Ruth lacks any non-trivial identity knowledge concerning herself (i.e. she doesn't know anything like 'I am Ruth' even if she does know trivialities like 'I am me' or 'Ruth is Ruth'). She is in a particular mental state: a state of pain. She attends to this pain via nociception – i.e. the normal way of observing pain that doesn't involve observing pain-behavior. As a result, something cognitive happens. She forms a belief whose truth-condition is: <Ruth, is in pain>.

Conative: Ruth lacks any non-trivial identity knowledge concerning herself. She has just moved a step backward. This resulted from something conative happening. She formed an intention whose satisfaction-condition is <Ruth, takes a step backward>.

There are many different beliefs whose truth-condition is <Ruth, is in pain>. Anyone who has been exposed to relevant works by Castañeda, Perry and Lewis will suspect that, in relation to *Cognitive*, a belief it would be most natural for Ruth (a monoglot English-speaker) to express by uttering the sentence 'I am in pain' – a *de se* belief – is somehow especially relevant. [1]

It takes a bit of care to say what the special relevance consists in. The belief nociception causes Ruth to form might *not* be of the 'I am in pain' sort. Perhaps attending to her pain causes Ruth to look up and catch sight of a doctor's clipboard on which is written 'Ruth is in pain'. Ruth forms a 'Ruth is in pain' belief, and her attending to her pain was a cause of her doing so. But one thing that seems harder to imagine is the following: Ruth forms a belief that is knowledge, on the basis of nociception, but the belief in question is not of the 'I am in pain' sort. *De se* belief seems essential for nociceptive knowledge.

---

[1] See Castañeda (1967), Perry (1979), Lewis (1979).

This raises an explanatory question. The name 'Kentucky Fried Chicken' builds in an obvious reference to chicken. It's not surprising that one would expect to be able to buy chicken in a Kentucky Fried Chicken shop. In setting *Cognitive* up, I glossed *nociception* as the 'normal way of observing pain that doesn't involve observing pain-behavior'. This gloss does not build in any reference to 'I'. If a belief that it would be natural to express using 'I' really is essential for nociceptive knowledge we should expect there to be something that explains why that is. If we can't give an explanation of why that is, we might even reopen the question of whether it is. But even supposing that isn't a serious possibility – e.g. because doing so involves violating clear-cut intuitions we have about a case like *Cognitive* – the purely explanatory question remains pressing.

Conative also raises an explanatory question. There are different intentions with the satisfaction-condition <Ruth, takes a step backward>. But, of these, the one Ruth would most naturally express by uttering the sentence 'I will take a step backward' – a *de se* intention – seems especially relevant. There are ways of fleshing out the details of the case, so that the intention Ruth formed turns out not to have been of this sort. Perhaps what happened is that Ruth formed an intention whose content she would report by uttering the sentence 'Ruth will take a step backward', one she expects to satisfy by calling Ruth up and giving her an order. The temerity she feels herself to be showing in forming this intention unnerves her, and her nerves show themselves in her unintentionally taking a step backwards. But one thing that is much harder to imagine is the following: Ruth *intentionally took a step backward* but the only intention on which she acted is of the 'Ruth will take a step backward' not the 'I will take a step backward' sort. This cries out for an explanation. The notion of *intentionally taking a step backward* does not build in any obvious reference to 'I'. So why should Ruth's intentionally taking a step backward depend on having her having an intention that she would most naturally express with 'I'?

In this paper, I answer both questions – which strike me as parallel – in a parallel way: the first by reference to a non-accidental truth condition on knowledge, the second by reference to a non-accidental satisfaction condition on action.[2] Both non-accidentality conditions are sensitive to how a state hooks onto its referent – to what I will call 'hyperintensional profile'. That makes the following strategy possible. We stipulate that we are dealing, in a given case, with knowledge or with intentional action. We infer that the associated non-accidentality condition is satisfied. We use that to reverse engineer what the hyperintensional profile of the knowledge-constituting belief, or action-producing intention, must be. In particular, we use it to explain why the required hyperintensional profile, if Ruth is to know in a case like *Cognitive* or to act in a case like *Conative*, is the very hyperintensional profile that *de se* states have.

What makes a case of knowledge 'like *Cognitive*' or a case of action 'like *Conative*'? I'll argue that the point about *Cognitive* generalizes to all beliefs that are based on *single-object relations* – knowledge-yielding relations that one can only stand in to oneself and one's own properties.[3] Other plausible examples of single-object relations are proprioception,

---

[2] See Williamson (forthcoming) for the more general idea that knowledge and action are duals.

[3] Readers familiar with the notion of immunity to error through misidentification (IEM) may wonder whether the question about scope could also be answered using that notion. Can we also say that a case is 'like *Cognitive*' if it involves knowledge constituted by a belief that is IEM relative to 'I'? I think that would be fine, since I think that beliefs are IEM relative to 'I' when and because they are based on single-object

kinesthesia, and introspection. I'll argue that the point about *Conative* generalizes not to a subset of intentional actions but to *all* of them: one needs a *de se* intention *whenever* one intentionally does anything. So there is a contrast in scope between the two generalizations from *Cognitive* and *Conative*. I'll argue this contrast in scope is explained by the fact that *all* our intentions impact the world by causing movements of our bodies (bodily action), or events in our streams of consciousness (mental action). By contrast, not all our beliefs involve single-object relations.

Before setting out the sections the paper, I provide a bit more detail on two key terms – 'hyperintensional profile' and '*de se*'.

A 'hyperintensional profile', as I use the term, is an answer to the question, asked about a mental or linguistic representation, 'Why does it have the object it has? (Equivalently: 'Why is it about what it is about?' or 'Why does it refer to what it refers to?' 'Why do its truth-/satisfaction- conditions involve the thing they do?'). Someone might have three different beliefs about Ruth, one about her 'as Ruth', another about her 'as the author of *Biosemantics*', and another about her 'as that woman'. There are three different 'hyperintensional profiles' in that there are three different answers to the question, asked of each of the three beliefs, 'Why does it have Ruth as its object?', each mentioning a different property Ruth has. If you prefer a different label for the same notion – e.g. 'reference-fixing story' – feel free to substitute it instead.

A '*de se*' state, as I use the term, is a representational mental state that has some special connection with the first-person pronoun, in English the word 'I' (I've glossed the connection by saying that they are *most naturally expressed* using that word). This word-oriented definition of a type of mental state leaves almost everything about the nature of *de se* states open. Ruth Millikan, who defends the view that 'the whole genre of indexicals is simply missing from thought' (1990:725), can accept that some of our states are *de se* in this sense. The phrase 'special connection with "I"' is in fact borrowed from her (see 1990: 732).

If one thinks *de se* states have other properties in addition to their connection to the first-person pronoun – e.g. if one thinks they have a certain hyperintensional profile – one will want to give an argument. There is in fact a standard and highly natural view about the hyperintensional profile of *de se* states: they are about their objects because those objects are their subjects.[4] The property of having this hyperintensional profile is the property of *de se* states that is crucial to my explanation of why they are essential. Part of the work of the paper is to show that they do have that property.

In (I), I argue, using examples that have nothing to do with the *de se*, that there is a non-accidental truth condition on knowledge that has a hyperintensional aspect, captured by a principle I call *Accord*. In (II), I use *Accord* in explaining why Ruth's nociceptive belief in *Cognitive* must be *de se* if it is to be knowledge. The point generalizes to all beliefs based on single-object relations. In (III), I extend the same kind of consideration from knowledge to action, in explaining why Ruth's intention in *Conative* must be *de se* if she is to intentionally take a step backwards. The point generalizes to *all* intentional actions. Section (IV) is comparative. It argues that the non-accidentality considerations developed

---

relations. However, it is only the notion of a single-object relation that does explanatory work in this paper. For discussion of IEM, see the essays in Prosser and Recanati (2012).

[4] E.g. Campbell (1994), Kaplan (1989), O' Brien (2007), Perry (1979) and Peacocke (2008) all hold that *de se* thoughts are governed by a rule according which any token of a *de se* thought refers to its subject.

in the earlier sections of the paper rule out positions about the *de se* defended by Ruth Millikan, Herman Cappelen and Josh Dever, and Ofra Magidor.[5]

(I) The non-accidental truth requirement on knowledge is sensitive to hyperintensional profile: Accord

The following is an intuitive and widely accepted idea about knowledge:

>*Non-accidental truth:* If one knows, then it is not accidental that one believes truly.

E.g. if I form a true belief ('It's 3pm!') about what time it is on the basis of the time the hands of a stopped clock are showing, my belief is *not* knowledge, because it is accidental that I believe truly.[6] It's an accident, for example, that the clock's hands didn't stop a minute later, or earlier, and that my relying on it for the time didn't cause me to form a false belief.

There are different proposals about how to capture more precisely the rough starting idea of a non-accidental truth requirement. One might say that knowledge requires belief whose truth is not *lucky*.[7] Or one might say that knowledge requires belief that is *reliably* true.[8] Or one might say that knowledge requires *safe* belief, where safety involves avoiding falsehood in all nearby worlds.[9] For my purposes here, any of these would have been fine. I stick with *non-accidental* truth.[10]

Suppose that I violate the requirement – I form a true belief but I don't gain knowledge because it is, in the relevant sense, accidental that I believe truly. An important thing to recall is that the scenarios that make this so may be ones in which I form a false belief that has a *different intension* from the belief I in fact formed. E.g., suppose I'm asked whether 71*35 is 2485 (correct) or 2585 (incorrect). I form the belief that it is 2485, on the basis of a coin-toss. Intuitively, I don't thereby come to know that 71*35 is 2485, because it is accidental that I believe truly. But the scenarios that make this so can't be ones in which 71*35 isn't 2485. There aren't any scenarios like that since it is a necessary truth that 71*35 is 2485. The relevant scenarios are rather ones in which I rely on the same coin-toss basis as I in fact relied on and this leads me to form a belief with a different intension – e.g. the intension that 71*35 is 2585 – that is false.[11]

Forming a belief about the answer to a multiplication question on the basis of a coin-toss is a bit bizarre. It might be suspected that the point the case illustrates is only relevant to some marginal class of beliefs. But that would be a mistake, as a case of Richard Kimberly Heck's brings out. In it the subject, Tony, forms a belief in response to the testimony of someone else, Alex, about Eric Blair (a.k.a. George Orwell):

>Suppose Eric Blair were to become amnesiac and check himself into a hospital. The doctor, Tony, deciding that she needs to have *some* name by which to call him, dubs him "George Orwell". And suppose further that Alex says—*not* intending to refer to Tony's patient—"George Orwell wrote *1984*" and that Tony

---

[5] See Millikan (1990, 2001), Cappelen and Dever (2013) and Magidor (2015).
[6] See Russell (1948:170).
[7] See Pritchard (2005).
[8] See Goldman (1976).
[9] See Williamson (2000) and Sosa (1999)
[10] See Unger (1968).
[11] See Pritchard (2012:182) for further discussion.

forms, in reaction to Alex's assertion, the belief she would express to other members of her staff as "George Orwell wrote *1984"*. This belief is true: Tony's new patient happens to be Eric Blair, that is, "the other" George Orwell. But surely it would not count as knowledge, even if Alex knows that George Orwell wrote *1984* (1995:95).[12]

Intuitively, the case is not one of knowledge because the truth of Tony's 'Orwell$_{\text{my patient}}$ wrote 1984' belief is accidental. Its being accidental cannot turn on its being accidental that the intension the belief in fact has is true. For there is a different belief with the very same intension that Tony could form and that would be knowledge. This is the belief Tony would express by saying 'Orwell wrote 1984', intending to refer to *whoever Alex is talking about*. Heck's passage doesn't explicitly say whether Tony has in fact also formed this belief. But, assuming she has, it seems like knowledge. It isn't merely accidentally true.

I think that what makes one of the two beliefs accidentally true, and the other not, is a difference in what I have been calling hyperintensional profile: i.e. a difference in why each of the beliefs has the object that is has. One belief has an appropriate hyperintensional profile, and the other does not. *Accord* is the general claim I want to defend about what makes for appropriateness of hyperintensional profile:

> *Accord:* If *S* gains knowledge of some thing by standing in a relation, *R*, to it, then the hyperintensional profile of the belief that constitutes that knowledge doesn't make it accidental that the belief is about something that *S* stands in *R* to. [13]

Accord doesn't exactly trip off the tongue. Before arguing for it, let me clarify how it applies to Heck's case.

S is Tony. *R* is whatever the relevant knowledge-yielding relation is when someone gains knowledge through testimony. Perhaps it is something like 'gaining testimony about x from one whose testimony expresses knowledge'. For brevity, I'll just say 'gaining testimony about'.

Consider Tony's 'Orwell$_{\text{mypatient}}$ wrote 1984' belief. The hyperintensional profile the belief has is as follows: it picks out a certain person because they are the patient Tony labeled 'Orwell'. Its having this hyperintensional profile *does* make it accidental that it picks out someone that Tony has gained testimony about. The relevant accident is that Tony's patient is the very person who Alex is talking about.

Consider Tony's 'Orwell$_{\text{whoeverAlexistalkingabout}}$ wrote 1984' belief. The hyperintensional profile the belief has is as follows: it picks out a certain person because they are the person Alex is talking about. Its having this hyperintensional profile *does not* make it accidental that it will pick out someone that Tony has gained testimony about. On the contrary, it ensures that the belief is about the very person Alex is offering testimony about.

The argument for *Accord* is inference to the best explanation. Alex says 'Orwell wrote 1984'. There is a belief that Tony could form that would be knowledge – the belief that picks out Orwell because he is *the person Alex is talking about*. There is a belief that Tony

---

[12] See Heck (1995).

[13] See Campbell (1994: 137) for a somewhat similar principle.

could form that wouldn't be knowledge – the belief that picks out Orwell because he is, as Tony would put it, '*my patient'*. I claim what distinguishes them, what best explains why one counted as only accidentally true, and not a case of knowledge, and the other doesn't, is that one belief violates *Accord* and the other does not.

This argument for *Accord* involves a case of testimony. But it need not have done. Much of our knowledge of the world comes from standing in potentially knowledge-yielding relations to things – e.g. hearing about them (testimony), seeing them, smelling them, hearing them, detecting them with a particle accelerator. These are genuine relations. If one sees *a*, and *a* is *b*, then one sees *b*. Yet, even if *a* is *b*, one can know *a is F* without knowing *b is F*. When one stands in a potentially knowledge-yielding relation to something, actually getting the knowledge the relation makes available requires forming a belief with an appropriate hyperintensional profile. *Accord* says what makes for appropriateness of hyperintensional profile.

I'll end this section's defense of *Accord* by considering a worry, a reply, a follow up worry, and a follow up reply.

The worry is that there is a whiff of paradox about appealing to the idea of an accident in relation to Tony's failure to know. One way of describing what is supposed to be accidental is as follows: if Tony were to form the belief 'Orwell$_{\text{mypatient}}$ is Orwell$_{\text{whoeverAlexistalkingabout}}$)' its truth would be accidental. But that belief has a necessary intension. At least in that sense, its truth couldn't be further from being accidental.

The reply is that we should refuse to identify the kind of non-accidental truth that consists in a belief having a necessary intension with the kind of non-accidental truth that is required by knowledge. This already was necessary in relation to the '71*35 is 2585' example (and, perhaps less obviously, in relation to the stopped clock example, if we think that the intension of my 'It's 3pm' belief is <3pm, 3pm, is identical to>).

The follow-up worry is that this reply makes the non-accidental truth requirement seem insubstantial. Perhaps the requirement cannot be understood independently of the notion of knowledge. In particular, perhaps *Accord* – which is intended to capture the hyperintensional aspect of the non-accidental truth requirement – cannot be understood independently of the notion of knowledge. In that case, can we really ever *explain* verdicts about knowledge by reference to the non-accidental truth requirement, or by reference to *Accord*?

Consider a fake barn case in which I believe 'That is a barn', of the only real barn in a county full of fakes. Is my belief knowledge? Is it accidentally true? Arguably our uncertainty about the second question tracks uncertainty about the first. If we end up saying the belief is not a case of knowledge, it isn't very informative to say that it isn't knowledge *because* it only accidentally true. If we'd gone for the view that the belief was knowledge, we would have said that it *is* non-accidentally true in the relevant sense.[14]

But this isn't the only kind of case. There are also the clear cases of merely accidentally true belief of the sort that provide the original motivation for the requirement. I look at a stopped clock at 3pm and form a true belief ('It's 3pm'). The belief isn't knowledge

---

[14] See Hawthorne and Gendler (2005) for a discussion of the instability of our intuitions around fake barn cases. See also Williamson (2000:100).

because it is an accident that I looked at the clock at the very time the hands are displaying. That this is accidental is something that can be appreciated independently of any verdicts about knowledge. In the stopped clock case, it really is informative to say I didn't know because the truth of my belief was accidental.

I submit that the same is true in Heck's case. Tony's 'Orwell$_{\text{my patient}}$ wrote 1984' belief isn't knowledge, because it is merely accidentally true. Moreover, we can dig down a bit and say something about the aspect of Tony's belief in virtue of which it is only accidentally true. The belief is only accidentally true despite the presence of a potentially knowledge-yielding relation (gaining testimony about) between Tony and the object of her belief (Orwell). What makes her belief merely accidentally true is that the hyperintensional profile of her belief makes it accidental that the belief is about someone she stands in this relation to. The violation of *Accord* really does explain the failure to know.

Accord is the hyperintensional face of the non-accidental truth requirement on knowledge. It's not the face we first think of when we think of that requirement but it is a face the requirement indisputably has. The next section uses *Accord* to explain why some knowledge is essentially *de se*.

(II) Accord and de se knowledge

I repeat *Cognitive* from above.

> *Cognitive*: Ruth lacks any non-trivial identity knowledge concerning herself. She is in a particular mental state: a state of pain. She attends to this pain via nociception. As a result, something cognitive happens. She forms a belief whose truth-condition is: <Ruth, is in pain>.

I stipulate that we are dealing with a version of *Cognitive* in which the belief Ruth forms constitutes nociceptive knowledge. If so, the belief must satisfy *Accord*. So we can try to use *Accord* to reverse-engineer its hyperintensional profile.

I'll use '*t* is in pain' as a way of referring to Ruth's belief. I'll first consider the following proposal: '*t*' picks out Ruth because Ruth has the property of *being Dever's favorite author*. No doubt this is a ridiculous proposal. I doubt anyone in the literature on nociceptive knowledge has ever given the property of being Dever's favorite author any serious attention. I still think it is illuminating to say explicitly why the proposal can't be right. That's because I am going to argue that the problem is that the proposal violates *Accord*. I further argue that the only way to remove the violation is to accept that the hyperintensional profile of '*t*' is as follows: '*t*' picks Ruth out because she is the subject of the state in which '*t*' figures. As noted earlier, this is to accept that the hyperintensional profile of '*t*' is the very hyperintensional profile that *de se* states are standardly assumed to have.

That permits an argument in support of the standard assumption about the hyperintensional profile of *de se* states. We have a strong intuition that, if the belief Ruth forms in *Cognitive* is nociceptive knowledge, it will be a *de se* belief about Ruth – i.e. one she would express with 'I'. I am just about to give an argument based on *Accord* to show that, if the belief Ruth forms in *Cognitive* is nociceptive knowledge, it will be about Ruth because she is its subject. These two together imply that a *de se* belief of Ruth's is about

her because she is its subject. That is, they imply exactly what the standard view about *de se* beliefs says.

With the standard view in place, we can give our explanation of why *de se* belief is essential for nociceptive knowledge. A belief with the hyperintensional profile *de se* beliefs are standardly taken to have is essential for nociceptive knowledge because only a belief with that hyperintensional profile is compatible with *Accord*. A *de se* belief is essential for nociceptive knowledge because – the standard view about *de se* beliefs being correct – *de se* beliefs are the beliefs with the hyperintensional profile *de se* beliefs are standardly taken to have.

I turn now to the argument that Ruth's *'t'*-belief needs to be about her because she is its subject if it is to be compatible with *Accord*.

Consider first the *Dever's favorite author* proposal about *'t'*. On this proposal, Ruth's *'t'*-belief picks out someone because they have the property of *being Dever's favorite author*. This makes it accidental that the belief is about a person whose pain Ruth is related to via nociception. The relevant accident is that Ruth is related to Dever's favorite author via nociception. So, Ruth's belief isn't knowledge. This is despite the fact that she stands in a potentially knowledge-yielding relation (nociception) to the person who the belief is in fact about. In this, she resembles Tony, who stands in a potentially knowledge yielding relation (gaining testimony about) to the person who her 'Orwell$_{mypatient}$' belief is in fact about. In both cases, a potentially knowledge-yielding relation obtains between the subject and the object of their belief. In both cases, the subject gains true belief but not knowledge. This is so in Tony's case, in fact. It is so in Ruth's case, on the supposition that *'t'* picks her out because she has the property of *being Dever's favorite author*. But remember – we stipulated that this is a case in which Ruth's belief *is* knowledge. We can conclude by *reductio* that the proposal about *'t'* is incorrect.

Hooking onto Ruth because she has the property of *being Dever's favorite author* is just one hyperintensional profile *'t'* could have. There are many others, corresponding to the many properties Ruth alone has. Here are three more of those:

> *Being the author of 'White Queen Psychology and other essays'.*

> *Being the fusion of a certain egg, E, and a certain sperm, S.*

> *Being picked out in a baptism that is the origin of a name-using practice involving a certain name pronounced 'Ruth Millikan'.*

Who someone's favorite author is depends on accidents of formation. Perhaps Dever could easily have liked Eric Hobsbawm better than he liked Ruth Millikan, if he'd signed up to a History rather than a Philosophy survey course at the age of 18. In some sense, it is far less of an accident – much more part of who Ruth is – that Ruth is the author of *White Queen Psychology and other essays*. It took a mind of a certain cast to write that book, and Ruth's mind had that cast. To the maximal extent possible, it is part of who Ruth is that she is the fusion of *E* and *S*. That fusion is her origin. If origins are essential, it is her essence. On one view about the individuation of proper names, proper names have their references essentially. If that view is right, only Ruth could have been picked out in a baptism that is the origin of a name-using practice involving the relevant name pronounced 'Ruth Millikan'.

Each of the three specified properties corresponds to a different proposal about how '*t*' picks Ruth out. Does any make for a more plausible proposal, one that stands a better chance of not violating *Accord*? No. The problem is that the notion of accidentalness on which Ruth's having one of these properties is less of an accident than having the property of being Dever's favorite author is purely metaphysical. For example, suppose Ruth were to respond to nociception by forming the belief 'The fusion of *E* and *S* is in pain'. It would still be an accident that the person her belief is about is someone she stands in the relation of nociception to. Perhaps, if we like, we can trace this back to the possibility of 'E' hooking onto a different egg, one that wasn't the egg Ruth came from, or 'S' onto a different sperm. But, the main point is that, epistemically, it can be accidental that a person has the metaphysically essential properties they have just as, epistemically, it can be accidental that a number (e.g. 2485) has the metaphysically essential properties it has (e.g. being the product of 71 and 35).

There is a pattern in the failure of the proposals we've considered: the lack of a connection between nociception and the hyperintensional profile '*t*' is being proposed to have.

That connection seems restored on the proposal that '*t*' has the hyperintensional proposal that *de se* states are standardly taken to have. On that proposal, '*t*' picks out Ruth because she has the following property:

> Being the subject of the state in which '*t*' figures.

Suppose one wanted to say that, on this proposal as well, *Accord* is violated. One would have to argue that the hyperintensional profile of '*t*' – picking out Ruth because she is *the subject of the belief* – makes it accidental that it is about a person whose pain Ruth is related to via nociception. But, on the face of it, that is false. Nociception – the normal way of observing pain that does not involve observing pain-behavior – is a relation that a subject can only stand in to herself and her own pains. One can easily know about other peoples' pains, but not via nociception. So there really does seem to be a connection between nociception and the hyperintensional profile of Ruth's belief on the *de se* proposal about '*t*', in contrast to all the non *de se* proposals.

The key fact about nociception that did work in this argument is the fact that it is a knowledge-yielding relation one can only stand in to oneself and one's own properties. I'll summarise this by saying it is a *single-object relation*.[15] The explanation of why nociception yields essentially *de se* knowledge, if it works, should generalize to any other single-object relations. Focusing on bodily examples, Evans lists 'our proprioceptive sense, our sense of balance, of heat and cold, and of pressure' (1982: 220). The list plausibly also includes our way of knowing about our current thoughts, and our current actions. *Accord* explains why the knowledge single-object relations provide us with is essentially constituted by *de se* beliefs.

I end this section by considering an objection to my explanation. It emerges from an influential point first made by Sydney Shoemaker.[16]

---

[15] 'Single-object faculty' is used in Martin (1995) and O' Brien (2007).

[16] See Shoemaker (1968). Shoemaker himself uses his point to defend a view about which judgments are immune to error through misidentification (IEM).

I emphasized that *Accord* is crucial to my treatment of *Cognitive*. But so, it turns out, is something else – the assumption that nociception is a single-object relation. This assumption might be contested. Perhaps there are metaphysically possible cases in which Ruth is sensitive to someone else's pain in a way that is similar to the way she is sensitive to her own pain in *Cognitive* (i.e. she doesn't have to observe any pain behavior). The same knowledge-yielding relation as is present in *Cognitive* is arguably also present in these cases. In *Cognitive*, Ruth stands in that relation to her own pain. In these cases, she stands in the same relation to someone else's pain. On this view, the knowledge-yielding relation I have been calling nociception is, as Shoemaker might put it, only *de facto* a single-object relation. In our world, and in worlds at all like it, the only person whose pains Ruth can nociceive is Ruth. But in far-flung worlds a subject could nociceive other peoples' pains.[17]

My reply is that it is enough for my argument for 'single object relation' to mean *de facto* single-object relation.

Suppose for the sake of argument that nociception, the knowledge-yielding relation in *Cognitive,* is merely a *de facto* single object relation.

That doesn't make the 'Dever's favourite author' proposal about *'t'* look *any better*. It remains the case that it is an accident that Ruth's 'Dever's favorite author' belief is about someone she can nociceive. Suppose Dever suddenly switches his favoritism to Hobsbawm, as he could easily do. It isn't as though this switch will somehow trigger a reorganization of the universe such that, as soon as the change occurs, Ruth will stop nociceiving Ruth (Dever's *earlier* favorite author) and start nociceiving Hobsbawm (Dever's *later* favorite author). There is still a disconnect between the relation that yields knowledge and the belief's hyperintensional profile.

Just as importantly, supposing that nociception is merely *de facto* a single-object relation doesn't make the *de se* proposal about *'t'* look *any worse*. There might be far flung cases in which Ruth stands in the relation of nociception to the pains of someone other than Ruth. But, precisely because they are far flung, they don't make it accidental that it is herself she stands in that relation to. So her *'t'*-belief will still be knowledge, on the *de se* proposal. One can have knowledge of one's environment on the basis of perceptual experiences even if there are far flung cases in which one is envatted and matching perceptual experiences lead one to form false beliefs. The remote possibility of cases in which one is envatted doesn't make it accidental that when it looks to one as though there is a chair with three legs in front of one, there really is. Similarly, Ruth's *de se* belief is nociceptive knowledge, despite the far-flung possibility of her belief being about one thing (herself), and the thing she is in a position to gain knowledge of (the person she is hooked up to via nociception) being something else.

Summary: The non-accidental truth requirement on knowledge doesn't have anything to do with the *de se* in particular. Neither does *Accord*. But in conjunction with the fact that some of the relations we use to gain knowledge are single-object relations, *Accord* explains why sometimes, *de se* belief is essential for knowledge. In the next section, I argue that parallel non-accidentality considerations about action explain a connection between action and *de se* intention.

---

[17] If you don't have this worry for pains, e.g. because they're mental entities, run through the same worry for proprioception.

(III) Action and de se intention.

How non-accidentality considerations explain the connection between action and *de se* intention is sensitive to what one's view of intentional action (hereafter, action) is. I'll first motivate the connection on what I call the *knowledge view* of action. I'll then switch to what I call the *causal view* of action.

Action and de se intention 1: the knowledge view of action

Suppose you are dropping crumbs while eating cake, and someone brings the fact to your attention. You observe that you didn't *know* that you were dropping crumbs. On the face of it, the force of the reply is that your dropping the crumbs wasn't intentional. Partly motivated by this kind of case, some philosophers of action accept the view that intentionally φ–ing always involves knowledge of the fact that one is φ-ing and, moreover, this knowledge – 'practical knowledge' – is constituted by an intention you have that is directed at φ–ing. This package is what I am calling the *knowledge view*. If knowledge is always constituted by belief, then the *knowledge view* implies that intentions are a kind of belief, which is what Velleman's (1989) version of the *knowledge view* says. If intentions are not beliefs, then the *knowledge view* implies that knowledge is not always constituted by belief, which is what Anscombe (1963), Setiya (2008) and Campbell's (2018) versions say.

Suppose Ruth intentionally φ's. Given the *knowledge view*, she will know that she is φ–ing. It might be misleading to say that there is a *faculty* on the basis of which she knows. On the *knowledge view* the very intention on which she acts constitutes her knowledge of acting, so it isn't as though the state that constitutes her knowledge tracks some independently existing entity, which might be read into talk of a 'faculty'. But even on the *knowledge view* there is a relation to her own actions in virtue of which Ruth is in a position to have practical knowledge of them – it might just be the *authoring* relation. That relation is a single-object relation – the cases in which she stands in it to anyone else's actions, if they are any, are very remote. So *Accord* implies that the knowledge this relation makes available to her must be constituted by a *de se* state. According to the *knowledge view,* the knowledge is constituted by the intention she has that is directed at φ–ing. So this intention must be a *de se* state. *Accord* and the *knowledge view* jointly imply that Ruth's intentionally φ–ing involves a *de se* intention directed at φ–ing.

All the advocates of the *knowledge view* I've mentioned – Anscombe, Campbell, Setiya, Velleman – take for granted that practical knowledge is *de se*. This is natural enough. The intuition we have about intentionally dropping crumbs is that, if someone is doing it, they will have knowledge they would express with the sentence 'I am dropping crumbs'. But even if it is obvious that practical knowledge of our own actions is *de se*, there is an explanatory question about why it must be. *Accord*, together with the point that the relation one stands in to one's actions is a single-object relation, answers that question. It explains why practical knowledge must be *de se*. If the *knowledge view* is correct, this converts into an explanation of why intentional action always involves *de se* intention, the connecting thought being that practical knowledge is constituted by intention. If the knowledge constituted is *de se*, the constituting intention must be too.

Action and de se intention 2: the causal view of action

Perhaps intentional action is action that is *caused in the right kind of way* by an intention.[18] And perhaps it is possible for an intention to cause action in the right kind of way without the agent having knowledge of what they are doing. Is rejecting *the knowledge view* and accepting this *causal view* a way of resisting the need for *de se* intention? I claim it isn't, on the grounds that without *de se* intention, the following non-accidentality condition on intentional action would be violated:

> *Non-accidental satisfaction:* If one intentionally acts, then there is an intention one acts on whose satisfaction is not accidental.

That there is *a* reading on which this condition is true is built into the *causal view*. If a climber's intentionally dropping another climber involves a dropping of the other climber that is caused, in the right kind of way, by an intention directed at dropping the other climber then intentionally dropping another climber will involve an intention whose satisfaction is in one sense not accidental – the intention *causes* the dropping, and the dropping is the thing the satisfaction of the intention requires. The thing that needs to be shown is that someone who accepts the *causal view* should say that the condition is true on a reading that implies that, whenever there is intentional action, there will be *de se* intention.

I repeat *Conative* from above:

> *Conative*: Ruth lacks any non-trivial identity knowledge concerning herself. She has just moved a step backward. This resulted from something conative happening. She formed an intention whose satisfaction-condition is <Ruth, takes a step backward>.

Some remarks about time are in order. *Conative* need not be a case in which Ruth formed an intention at some earlier point, and stored it up for control of future action. Much intentional action is spontaneous. The only intention she has may be an intention formed as she begins to act. We should also assume that the intention has an appropriate temporal component – Ruth would express it 'now' rather than '3pm' for example.

As noted above, we seem to have a strong intuition that if *Conative* is a case in which Ruth intentionally takes a step backward then the intention she formed will be one that she would express with 'I take a step backward' rather than e.g. 'Ruth takes a step backward'. I want to explain why this is so by reference to a plausible non-accidental satisfaction requirement on intentional action.

The strategy is parallel to the strategy used in the previous section in relation to *Cognitive*. I use the sentence '*t* takes a step backward' to refer to Ruth's intention. I begin by looking at the perhaps ridiculous proposal that '*t*' picks Ruth out because she is *Dever's favorite author*. I'll argue that on this proposal the satisfaction of the intention is implausibly accidental. I'll then argue that, as was the case in *Cognitive*, and for a structurally identical reason, to get a better proposal we need to accept that '*t*' is *de se*. Finally, I'll argue that the point generalizes from the action involved in *Conative* to actions across the board.

---

[18] See Davidson (1980) for a classic statement of the *causal view*.

Suppose the *Dever's favorite author* hypothesis about '*t*' is correct. Ruth's '*t* takes a step backward'-intention is non-accidentally satisfied, in that it is an early point in a causal chain that leads up to the very thing that its satisfaction in fact requires – a certain movement of Ruth's body. (Similarly, even on the *Dever's favorite author* hypothesis about Ruth's '*t* is in pain' belief, that belief was part of a causal chain an earlier point in which was the thing that its truth in fact requires, Ruth's being in pain.) But the intention is merely accidentally satisfied in that the later point in the causal chain that its satisfaction in fact involves could easily have been irrelevant to its satisfaction. The movement of Ruth's body would have been irrelevant if Hobsbawm had been Dever's favorite author. In that case, the satisfaction of an intention with the proposed hyperintensional profile would have required a movement of Hobsbawm's body, not Ruth's.

Taking a step backward is comfortably within Ruth's behavioral range. So, on the correct proposal about '*t*', we should be able to explain why her '*t* takes a step backward' intention gets satisfied just by noting that Ruth formed it, and nothing funny has happened (e.g. Ruth's body is not suddenly paralyzed, nobody grips Ruth's legs just as she is about to move them).

The *Dever's favorite author* hypothesis about '*t*' is incompatible with this. In order to explain why the intention gets satisfied, on that hypothesis about it, the fact that Ruth is Dever's favorite author needs to be mentioned. Why? The crucial point is that there is one body, her own, that Ruth's intention has its most immediate impact on. It is not an accident that, if Ruth forms the intention to take a step backward, the next thing that happens is that her own body moves in such a way as to take a step backward. By contrast, it is accidental that, if Ruth forms an intention to move a step backward, the next thing that happens is that Dever's favorite author's body moves in such a way as to take a step backward. Granted that Ruth *is* Dever's favorite author, when Ruth's body moves, Dever's favorite author's body moves. But, it isn't as though, if Dever suddenly came to prefer Hobsbawm over Millikan, as could easily happen, the universe would reorganize itself such that, now, Ruth's intentions have their immediate impact not on Ruth's body but on Hobsbawm's.

By contrast, on the *de se* hypothesis about '*t*', it isn't accidental that the intention causes the very thing that its satisfaction requires. Single-object relations, as I defined them, were epistemic relations. The immediate impact relation is a motor relation. But it shares a structural property with single-object relations. It is a relation that, at least *de facto,* each of us only stands in to herself. So, if Ruth has an intention that hooks onto her *because she is the subject of the intention*, and that has its impact via the immediate impact relation, it's not accidental that the person the intention is about – the person movement of whose leg would constitute the satisfaction of the intention – is the very person whose body the intention impacts on.

Suppose that explanation of why Ruth's intention in *Conative* needs to be *de se* is accepted. How far does the point generalize? Can we conclude that intentional action *always* depends on *de se* intention?

Nothing I said about *Conative* seems sensitive to the predicative component of the intention – 'takes a step backward'. It doesn't make a difference if we switch to a case where Ruth acts not on a '*t* takes a step backward' intention but on '*t* takes a step forward' or '*t* does a pirouette' intention. The immediate impact relation seems equally relevant to these intentions, so the case that the intention acted on needs to be *de se* is the same. The point doesn't even seem special to the intentions involved in bodily, as

opposed to mental, actions. Suppose Ruth forms an intention to visualize the first three rooms of the National Gallery. The next thing that happens is that somebody's stream of consciousness becomes populated with visual imaginings of the first three rooms of the National Gallery. It isn't an accident that it is Ruth's stream of consciousness in which this happens (even if there are far flung cases in which it is someone else's). On the face of it, the immediate impact relation is relevant for *all* intentional actions. That is the initial case for thinking that there is the same need for *de se* intention whenever one intentionally does anything: intentionally φ–ing always involves a *de se* intention that one φ (a.k.a. an intention to φ).[19] The case is strengthened by considered an objection in the next section.

Summary: Much about intentional action is controversial. The *knowledge view* and the *causal view* are fairly different pictures of the nature of a subject's relationship to her own actions. But, on either view, there is a strong case that we need *de se* intentions whenever we intentionally do anything. Non-accidentality considerations can explain why.

(IV) Comparison with other positions: a reply to the skeptics

Castañeda, Lewis and Perry made the idea of the *de se* and of essential indexicality famous. Subsequently, there has been a bit of a backlash against their idea – two of the three papers I'll discuss in this section have the word 'myth' in their title. The interest of the current paper does not depend on how well-motivated this backlash is. Cases like *Cognitive* and *Conative* raise explanatory questions that demand answers and that haven't already received them. Even so, it's natural to wonder whether non-accidentality considerations – relied on in the previous three sections of this paper to try to answer the explanatory questions *Cognitive* and *Conative* raise – might *also* provide a plausible reply to the authors who constitute the anti-essential indexical backlash. I'll argue they can.

The main authors of interest here are: Ruth Millikan, Herman Cappelen and Josh Dever (C&D hereafter), and Ofra Magidor. Between the three of them (counting C&D as one), I think we can identify two main views. One view is explored by Millikan and C&D but not Magidor. The other is explored by C&D and Magidor, but not Millikan.

C&D characterize the overarching position they accept as follows:

> There is no philosophically distinctive role to be played by perspectivality in the explanation of action, inquiry, or perception. (2013: 2)

The two views C&D appeal in support of this position are both present in their discussion of a passage they quote from Perry. Perry writes:

> Consider a transaction with a fax machine. To press certain buttons on it, I have to move my fingers a certain distance and direction from me. It isn't enough to know where the buttons were relative to one another, or where the fax machine was in the building or room. I had to know where these things were relative to me. (1998: 87)

---

[19] Some argue that the intention to φ and the intention that one φ are two different, albeit both *de se,* intentions (see Recanati 2007). I disagree (see Morgan forthcoming). But the disagreement is orthogonal to the current discussion. Even if there are two intentions, each has the same satisfaction-conditions and in each case the satisfaction-conditions involve the subject because they are the subject. I've argued that Ruth must have *an* intention that has both of these properties. She requires a *de se* intention. She may not require either subspecies of *de se* intention in particular, if there really are two subspecies.

C&D's comment on the passage is as follows:

> Millikan is right, and importantly so, in pointing out that there's not even the beginning of an argument for anything properly labeled "essential indexicality" here. If anything is shown, it has to do with something like what Millikan calls an "active self-name" and it need not be indexical. However, we think that even the claim about the "active self-name" (or "*de se* concept", if you prefer that label) is unsupported. Why think that the agent had to represent herself *at all* in order to push the button? (2013: 45)

The first part of the comment (before 'However') is motivated by the view that C&D share with Millikan – I'll call it the *self-name* view. The second part of the comment (after 'However') is motivated by the view C&D share with Magidor – I'll call it the *worldly states* view. This section considers each in turn.

 (1) *The self-name view: Millikan and C &D*:

Suppose we accept that *de se* states have some distinctive property – e.g. being essential to knowledge yielded by single-object relations, being involved in any intentional action at all. Does anything interesting follow about *indexical* states? Only if *de se* states are indexical states. But that might be contested. The *word* used to express *de se* states – 'I' – is indexical. But perhaps indexicality is just a property of some words that theorists have incautiously projected onto the thoughts those words express. *De se* thoughts could be thoughts with a characteristic linguistic expression (the indexical 'I'), perhaps also some kind of special relation to knowledge and action (gestured at by Millikan's phrase '*active self-name*'), but not themselves be indexical. Instead, they involve names.

> The *self-name* view: *A de se* state of *S*'s involves a name of *S* and is not indexical.

C&D express approval of the *self-name* view, but they don't provide their own discussion of it. Hence, I'll focus on Millikan's presentation. She writes:

> My mental "I", my @"RM", is not an indexical. More reasonably it is a (Millian) name for me, your "I", which may well have quite a different mental shape, is a (Millian) name for you (1990:732).

She also supplies definitions for '(Millian) name' and 'indexical'. A (Millian) name is a sign 'about the semantics of which nothing can be said beyond that it is a name with such and such a referent' (1990: 732 fn 9). An indexical is a sign 'that has no constant referent, no referent qua sign type' (1990:724).

I've argued that non-accidentality considerations provide reason to accept that the hyperintensional profile of a *de se* state is that it is about someone *because they are its subject*. But is there any reason to think that this is incompatible with Millikan's position? Couldn't someone think the following:

Millikan is correct that Ruth's *de se* state involves a (Millian) name for Ruth. That is just to say that the whole story about the semantics of (the singular component of) the state is:

It is about Ruth

But taking that view about the state's semantics in no way constrains our options in saying what its hyperintensional profile is (or forces us to say it doesn't have one). We still want to know: 'Why is the state about Ruth?'. And we can still answer: 'It is about Ruth because Ruth is the subject of it'. Given Millikan's view, this answer won't describe the state's semantics (as it would be if the state were indexical). But it can still be the correct answer and, in particular, a correct description of its metasemantics.[20]

This irenic suggestion doesn't work. To bring out why not, it helps to consider a possible 'no fact of the matter' view about the issue Millikan raises.

Ruth's token of the word 'I' is a token of a (linguistic) indexical sign. That, by Millikan's definition of 'indexical', is because it is a token of a sign that has *no constant referent*: when other English speakers produce the same sound, with a *different referent*, that counts as them using the *same sign*.

Does a *de se* state of Ruth's involve a token of a (mental) indexical sign? That, by Millikan's definition of 'indexical', depends on whether it is a token of a sign that has *no constant referent*. Different peoples' *de se* states clearly have *different referents*. So the issue of whether *de se* states are indexical turns on whether different peoples' *de se* states involve the *same sign*.

The relevant 'no fact of the matter' view is the view that whether we say different peoples' *de se* states involve the *same sign* – hence whether we say they are indexical – is up to us. We can group intentional tokens into sign types in whatever way we please.

Millikan must reject this view. Her denial of mental indexicality is intended to bust a myth, not to be an optional reformulation of agreed facts. There is also a plausible basis for rejecting it. We should acknowledge single signs corresponding to natural sets of intentional tokens (e.g. all tokens of the word 'I') but not to gerrymandered sets (e.g. all tokens of the word 'I', the first token in this paragraph of the word 'myth', and Ruth's most intensely held *de se* belief).

However, characterizing the issue in this way makes it clear that there is a *prima facie* case for accepting that there is a single sign corresponding to different peoples' *de se* states. As sets of intentional entities go, the set of *de se* states seems highly natural. It is the set of: all the states that Millikan is trying to bust a myth about; all the states that bear a 'special connection' to the first-person pronoun; all the states that involve what Millikan calls 'active self-names'. If Millikan nevertheless denies the set corresponds to a single sign, there needs to be some difference between, say, my *de se* states and Ruth's *de se* states to justify that denial. Difference in hyperintensional profile could play this role. Suppose that Ruth's *de se* states hooked onto Ruth for one reason and mine hooked onto me for some completely different reason. Then, Millikan could say, the things different subjects' *de se* states have in common – e.g. linguistic expression – mask underlying differences – in hyperintensional profile – that justify the refusal to accept they all

---

[20] Millikan mentions in a footnote (1990:732 fn 9) that she wants to give a teleosemantic account of how (Millian) names across the board get their referents.

correspond to a single sign. But that way of justifying the refusal is not available if a *de se* state is about its subject because they are its subject – in that case, my *de se* state is about me for the very same reason that Ruth's *de se* state is about her. There is no other way I can see of justifying the refusal. So, I conclude, the irenic suggestion doesn't work. If it's true that a *de se* state is about its subject because they are its subject, then it's also true that different peoples' *de se* states involve the *same sign*. This sign clearly lacks a constant reference. So it is an indexical and not a (Millian) name.[21]

Summary: We should reject the *self-name* view. The next section argues that we should reject the other main view that anti-essential indexical authors have appealed to.

(2) The worldly states view: C&D and Magidor

In (III), I offered an explanation of why Ruth's <Ruth, takes a step backward> intention in *Conative* must be *de se*. I then argued that the point was not sensitive to the *predicative* component of the intention – that that component was 'takes a step backward' rather than e.g. 'does a pirouette' or 'visually imagines the first three rooms in the National Gallery'. What about the *singular* component? The argument surely did rely on the fact that the intention had satisfaction-conditions that involved the agent. If an intention doesn't – if it is *worldly* rather than *self-directed* – then it can't be *de se*. Strictly speaking, all my argument did was show that, *if* an action is the result of a self-directed intention, non-accidentality considerations give us reason to think the intention must be *de se*. To conclude that every action involves *de se* intention, we need the premise that every action involves self-directed intention. The *worldly states* view implies that this premise is false.

> The *worldly states* view: Some intentional actions do not involve any self-directed states at all.

In this section, I'll say, first, what the intuitive case that intentional action always involves self-directed intention is and, then, why the guiding analogy defenders of the *worldly states* view appeal to in order to defend their contrary claim is unpersuasive.

---

[21]C&D use the word 'indexical' (which Millikan does) and the word 'perspectival' (which Millikan doesn't). Might the two words mean something different? Might someone be persuaded that *de se* states are indexical but still insist that *de se* states are not perspectival?

I don't give separate consideration to the view that *de se* states are not perspectival for two reasons.

First, it's natural to understand 'indexical' and 'perspectival' as synonyms. For example, Adrian Moore, in his book length study of perspectival representation, offers the following as an initial gloss on the notion: 'The content of a perspectival representation depends not only on its type, but also on the point of view from which it is produced'.(1997:11). This is very close to Millikan's gloss on 'indexical'. C &D themselves freely alternate between 'indexical' and 'perspectival'. E.g. the heading of their section 1.2. is 'Our Target: The Almost Universally Accepted View that Indexicality is Philosophically Deep' while the first line of that section is 'This book is an extended exploration and defense of the view that perspectivality is philosophically shallow' (see 2013: 2).

Second, suppose someone tells us 'perspectival' doesn't mean indexical. That alone doesn't make the view that *de se* states are not perspectival one that needs to be engaged with. We would first need to hear: what different thing 'perspectival' means; why *de se* states being indexical doesn't guarantee their also being perspectival; what the initial case for thinking they aren't perspectival is; and why their not being perspectival should be regarded as news. So I take it to be reasonable to think the only view in this area that needs to be engaged with is Millikan's denial that *de se* states are indexical.

We've already seen C&D's discussion of Perry's fax machine case and their claim that there is no reason to think that the agent 'had to represent herself *at all* in order to push the button'. So they think that pushing the button on a fax machine is an example of the kind of action the *worldly states* view claims is possible.[22]

Magidor offers the following case:

> Suppose that I believe that people in Africa are starving, that when money is donated to Oxfam this prevents people from starving, and that preventing people from starving is a good thing. Suppose that as a result of holding these beliefs, I go ahead and donate money to Oxfam. This seems like a clear case of an intentional action. Note that at least as I have described the case, no first-personal beliefs were involved…. It is thus far from clear that so called *de se* attitudes play the unique or distinctive role with respect to action that defenders of the myth commonly attribute to them (2015:259). [23]

Magidor's initial idea is that, since the subject need not have any self-directed beliefs at all, they need not, in particular, have any *de se* self-directed belief. Perhaps that is right. But her ultimate conclusion – after the ellipses – concerns *attitudes* in general rather than *beliefs* in particular. That conclusion wouldn't be supported by the case (the only one Magidor discusses) if the subject of the case didn't need a self-directed belief but did need a self-directed intention, and it had to be *de se*. So, Magidor is committed to denying that the subject needs a self-directed intention. There surely will be *some* intention behind their action (It isn't in general true that merely *believing something to be good* causes one to act in such a way as to bring it about. There are too many conflicting believed-to-be-good things for that to be so.). If the case supports the *worldly states* view, this intention will have to be a worldly intention – e.g. the intention *that money is donated to Oxfam*.

In the example of action I discussed – *Conative*, where Ruth's goal involved Ruth taking a step backward – the agent was a constitutive part of the goal. In the examples of action C&D and Magidor discuss – e.g. a case in which the goal is that a fax message is sent, or one in which the goal is that Oxfam is donated to – the agent isn't part of the goal. Does this difference make a difference to the need for a self-directed intention?

There's an intuitive case that it doesn't. Even in the case of actions of mine in which I am not involved as part of the goal, I am involved as the instrument. If I do something then, of course, *I* do something. E.g., suppose Ofra intentionally brings it about that Oxfam is donated to. There will be some more specific means by which she pulls this off. E.g. suppose we are told she does it *by donating to Oxfam herself*. Aren't we now in a position to say that Ofra will have formed an intention *to donate to Oxfam?* Worldly intentions seem possible. But they seem to lead to action only via self-directed intentions.

The guiding analogy

The defenders of the *worldly states* view resist this. They appeal to an analogy. Here is how Magidor puts it:

---

[22] They discuss a variety of other cases. E.g. they suggest that someone (e.g. a publisher) might act on the worldly intention 'Kripke releases his unpublished papers', without having any self-directed intention. (2013:51).

[23] See Magidor (2015:259).

> Suppose in order to lift my hand, I must first shift my shoulder blade. It does not seem that I cannot intentionally lift my hand without also having the belief that in order to lift my hand I need to shift my shoulder blade: even intentional actions are ultimately achieved via other actions, ones that are more basic or direct, rather than being the result of some process of thought. In a similar manner, reaching into my pocket in order to donate money might be such a more basic or direct action (2015:259).

Sometimes one does one thing in a way that involves doing something else where one has no thought – e.g. no belief, no intention – directed at the second thing. The familiar case Magidor points to is: one lifts one's hand in a way that involves shifting one's shoulder blade but one has no thought directed at the shoulder blade shift. One doesn't *believe* anything like 'Shifting this shoulder-blade thusly is needed to lift my hand', or *intend* anything like 'This shoulder blade moves thusly'.[24]

The following is what needs to be possible if the Oxfam case is to undermine the claim that action always depends on self-directed intention: Ofra forms a worldly intention that Oxfam is donated to. As a result of this intention, she intentionally does something. This is despite the fact that she lacks any self-directed intention. In particular, she lacks an intention *to donate money* (which would be self-directed even without the 'reaching into my pocket' part). Ofra brings it about that Oxfam is donated to via donating herself (as opposed to, for example, convincing someone else to donate). But – and here is the guiding analogy – Ofra's donating is a non-intentional action via which she does something else intentionally (bringing it about that Oxfam is donated to), just like her shifting her shoulder blade is a non-intentional action via which she does something else intentionally (lifting her hand).

I'll argue that this self-directed intention free route to action is not possible in two stages. First (i), I'll argue that even if we were to accept that the route is possible, that wouldn't make a difference to whether Ofra needs a self-directed conative state that hooks onto her because she is its subject. On either route to action – self-directed intention involving or self-directed intention free – there is such a state. The outstanding controversy is whether this state has to be a person-level intention. I'll argue that it does – (ii).

(i) Magidor draws an analogy between the case she is interested in and a familiar case. But she doesn't say much about the familiar case. It helps to say something. When I lift my hand by shifting my shoulder blade I don't have an intention directed at the shoulder blade shift. But a sub-personal conative state – a motor command – certainly is involved. It isn't magic that my shoulder blade does the very thing it needs to do if my intention to lift my hand is to be satisfied.[25]

A motor command whose satisfaction-conditions are <This shoulder-blade, shifts thusly> is arguably not *self-directed*. A shoulder-blade is at best a small part of a self. But, if Magidor were right that the Oxfam case might work in the same way as the hand-lifting case, the satisfaction-conditions of the corresponding motor command in the Oxfam case would have to be: <Ofra, donates to Oxfam>. Such a motor command would be self-directed. Moreover, it would have its impact via the immediate impact relation. So, the case based on non-accidentality considerations that it would have to be about Ofra

---

[24] C&D appeal to the same point at (2013:43). To avoid repetition, I focus on Magidor's discussion.

[25] See Jeannerod (2006: ch 1) for discussion of the role of motor commands in intentional action.

because she is its subject is the same case, discussed in section III, as if it had been a self-directed intention. That case isn't sensitive to the person-level/subpersonal distinction.

(ii) I've just argued that, even if the analogy with the hand-lifting case worked, and Ofra had no self-directed intention, she would still have a self-directed conative state. More particularly, she would have a conative state with the satisfaction-conditions <Ofra, donates to Oxfam> that hooks onto Ofra because she is its subject. [26] The outstanding controversy is whether this conative state must be a (person-level) intention as opposed to a (subpersonal) motor command. I'll argue that it must be an intention.

Suppose Ofra has just intentionally lifted her hand and we ask her *how* she lifted her hand (e.g. did it involve the shoulder blade moving, or does the action get directed just involving parts of her arm below the level of her elbow). It's entirely possible that she may have nothing very informative to say. 'I just did it' or 'I just did it: whatever needed to happen with my body did happen' or 'I just did it: I'm not a physiologist'.

Suppose Ofra has just brought it about that Oxfam is donated to by donating herself and we ask her *how* she brought it about that Oxfam is donated to (e.g. did it involve her donating herself, did it involve her persuading someone else to donate?). If Ofra's self-directed conative state is subpersonal, it ought to be possible that she exhibits a comparable level of ignorance: 'I just brought that about, who knows how' or 'I just brought it about: some relevant means will have been selected; I can't say which'. That really doesn't seem believable. Ofra's focus may, as she donated, have been entirely on Oxfam's need and not on her own role in catering to it. That she catered to that need by donating herself may not have involved any protracted deliberation, and other ways of catering to the same need – e.g. via persuading someone else to donate – may not have occurred to her. All the same, Ofra has a conative state whose satisfaction-conditions are <Ofra, donates to Oxfam>. And we seem to be confident that this is a state that Ofra could in principle access via introspection. Being such as to be in principle accessible by introspection is a marker of being a person level state as opposed to a subpersonal one. [27] Another marker of being a person level state is being poised to play a role in reasoning. Again, Ofra's conative state seems poised to play a role in reasoning. If Ofra is in this state then she will not pursue further deliberations about how to bring it about that Oxfam is donated to. But, if the state were subpersonal, there's no reason why being in it should arrest her person level deliberation on that question.

Summary: Intention is the most challenging state-type for the *worldly states* view. It's plausible that we can form worldly intentions. The difficulty is to see how they could lead to action otherwise than by leading to self-directed intentions (the hand-lifting analogy does not make this easier to see, I have argued). If they can't, then any case of action, whether or not its ultimate goal is worldly, will involve self-directed intention. Hence, what applies to *Conative* applies to intentional actions across the board.

---

[26] Is this already incompatible with the *worldly states* view's treatment of the case? That depends on whether the view involves denying the need for *person level* self-directed states, or the need for *any* self-directed states. Magidor doesn't make the distinction. C&D do. They deny the need for states that represent the subject whether these are 'conscious or subpersonal' (see 2013:43).

[27] See Block (1995).

Conclusion

Thought about the self can be mind-boggling. Methodologically, it makes sense to approach the topic armed with general principles that have been supported using examples from a different domain. In this paper, the general principles are non-accidentality conditions on knowledge and intentional action. It is a fact that we have a special epistemic perspective on ourselves – even if it turns out there are far flung cases where we have the same perspective on someone else. It is a fact that we have a special motor perspective on ourselves – even if it turns out there are far flung cases where we have the same perspective on someone else. We exploit the epistemic perspective whenever we gain knowledge on the basis of single-object relations, and we exploit the motor perspective whenever we intentionally do anything. Non-accidentality conditions on knowledge and on intentional action explain why exploiting these perspectives always involve *de se* states.

References:

Anscombe, G. E. M (1963). *Intention.* 2nd ed. Oxford: Blackwell.

Block, N (1995) 'On a confusion about a function of consciousness' *Behavioral and Brain Sciences* 18, 227-287.

Becker, K., & Black, T. (eds.). (2012). *The Sensitivity Principle in Epistemology.* Cambridge: Cambridge University Press.

Campbell, J (1994). *Past, Space and Self.* Oxford: Clarendon Press.

Campbell, J (1992). 'Immunity to Error through Misidentification and the Meaning of a Referring Term', *Philosophical Topics* 26 (1999), 89-104.

Campbell, L (2018). 'An epistemology for practical knowledge', *Canadian Journal of* Philosophy, 48 (2):159-177.

Cappelen, H and Dever, J (2013). *The Inessential Indexical: on the Philosophical Insignificance of Perspective and the First Person.* Oxford: Oxford University Press.

Carter, A, Gordon, E. and Jarvis, B (eds) (forthcoming) *Knowledge First Approaches to Epistemology and Mind.* Oxford University Press.

Castañeda, H-N (1967). 'Omniscience and Indexical Reference'. *Journal of Philosophy* 64 (7): 2013-210.

Frege, G (1966). 'The Thought: A Logical Inquiry', in P.F. Strawson (ed.) *Philosophical Logic.* Oxford: Oxford University Press.

Goldman, A. (1976). 'Discrimination and Perceptual Knowledge'. *The Southern Journal of Philosophy,* 73: 771-91.

Evans, G. (1982). *The Varieties of Reference.* Oxford: Clarendon Press.

Hawthorne, J. & Gendler, T (2005). "The Real Guide to Fake Barns." *Philosophical Studies* 124: 331-352.

Heck, R. K. (1995). 'The Sense of Communication', *Mind* 104 (413): 79 – 106 (originally published under the name 'Heck, R.').

Jeannerod, M (2006). *Motor Cognition: what actions tell the self.* Oxford: Oxford University Press.

Kaplan, D. 'Demonstratives'. In Joseph Almog, John Perry & Howard Wettstein (eds.), *Themes From Kaplan.* Oxford University Press. pp. 481-563 (1989)

Lewis, D (1979). 'Attitudes *De Dicto* and *De Se'*, *The Philosophical Review* 88: 513-543.

Martin, M.G.F (1997). 'Self-Observation'. *European Journal of Philosophy* 5, 119-40.

Magidor, O (2015). 'The Myth of the De Se', *Philosophical Perspectives* 29 (1): 249-283.

Millikan, R (1990). 'The Myth of the Essential Indexical', in *Noûs* 24 (5): 723-734.

Millikan, R (2001), 'The Myth of Mental Indexicals' (revised and expanded version of 'The Myth of the Essential Indexical') pp. 163-77 in Brook and DeVidi (eds.) *Self-Awareness and Self-Reference* John Benjamins Publishing Company.

Morgan, D (forthcoming) in Ball and Schuringa (eds) *The Act and Object of Judgment* Routledge.

Moore, A (1997). *Points of View*, Oxford: Oxford University Press.

O' Brien, L (2007). *Self-Knowing Agents*, Oxford: Oxford University Press.

Peacocke, C (2008) *Truly Understood*. Oxford: Oxford University Press.

Perry, J (1979). 'The Problem of the Essential Indexical', in *Noûs* 13 (1): 3-21.

Pritchard, D (2005). *Epistemic Luck*, Oxford: Oxford University Press.

Pritchard, D (2012). 'In defense of modest anti-luck epistemology', in Becker, K., & Black, T. (Eds.).

Prosser, S & Recanati, F (eds.) (2012). *Immunity to Error Through Misidentification: New Essays*. Cambridge University Press.

Recanati, F (2007). *Perspectival Thought: a plea for moderate relativism*. Oxford: OUP.

Russell, B. (1948). *Human Knowledge: its Scope and its Limits*. London: Allen & Unwin.

Setiya, K (2008). 'Practical Knowledge', *Ethics* 118 (3): 388-409.

Shoemaker, Sydney (1968). "Self-Reference and Self-Awareness." *The Journal of Philosophy* 65 (19): 555–67

Sosa, E. (1999). 'How to Defeat Opposition to Moore'. *Philosophical Perspectives* 13: 141-54.

Unger, P. (1968). 'An Analysis of Factual Knowledge', *Journal of Philosophy* 65, 157-170.

Velleman, J. D. 1989. *Practical Refection*. Princeton, NJ: Princeton University Press.

Williamson, T. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.

Williamson, T. (forthcoming). 'Acting on Knowledge'. In Carter, A, Gordon, E. and Jarvis, B (eds).