

Immunity to error through misidentification: what does it tell us about the de se?

Daniel Morgan

in S. Prosser and F. Recanati (eds.) *Immunity to Error Through Misidentification: New Essays*. Cambridge: Cambridge University Press, 2012: 104-124.

One role for immunity to error through misidentification in a philosophical discussion is as an *object of investigation* in its own right. One begins effortlessly enough with some examples, such as the following seminal one of Wittgenstein's.¹ After an accident, feeling a pain in my arm, I might see a clearly broken arm by my side, assume that this arm is my own arm, and on that basis judge "My arm is broken" or "I have a broken arm". If the broken arm is in fact my neighbour's, I will have made a somewhat limited kind of error that it is fairly natural to call an "error through misidentification". By contrast, if I judge "I have tooth-ache", in the normal way, my judgment seems immune to that kind of error.

Such examples help convey at least a rough grasp of what the phenomenon of immunity to error through misidentification is. The hard part of the discussion is to go on to provide a rigorous *definition* of that phenomenon – a set of necessary and sufficient conditions that is couched in general terms, and so does not rely, in the way the last paragraph did rely, on ostending particular examples of the phenomenon. A possible addition and refinement is a taxonomy of various different sub-types of immunity to error through misidentification, one which illuminates by carving the phenomenon at its natural joints.

¹ See Wittgenstein 1958.

A different role for immunity to error through misidentification in a philosophical discussion is as a *source of illumination* on immediately adjacent philosophical topics. One begins effortlessly enough, conveying at least a rough grasp of what the phenomenon of immunity to error through misidentification is by means of a few examples. One assumes only as much about that topic as seems relatively clear and uncontroversial. The hard part is then to use what one has assumed to gain leverage on something else.

An attractive feature of discussions of immunity to error through misidentification in which the second role predominates is their optimistic flavor. Instead of there being one more thing that needs to be done – provide a rigorous definition of immunity to error through misidentification – there is a new tool that can be used to get other things done. Moreover, it does seem fair to assume, as such an approach must, that philosophers working in the area have at least a rough, working grasp of what immunity to error through misidentification is. After all, if they did not, they would struggle to come up with their own *examples* to illustrate the phenomenon, which in fact they do not. And they would have little basis on which to set about deciding what the *right* definition of immunity to error through misidentification is. So it does seem reasonable to ask whether what is clear and uncontroversial about immunity to error through misidentification is also substantial enough to be usefully deployed in an account of something else.

James Higginbotham and François Recanati approach immunity to error through misidentification primarily as a source of illumination.² For each of them, the topic to

² See Higginbotham 2003 and Recanati 2007.

be illuminated is the same: *de se* thought.³ Each in effect asks: what must the nature of *de se* thought be such that it can exhibit the pattern of immunity and vulnerability to error through misidentification relative to the first person which we know that it does exhibit. Thus, the presumed need for an account of *de se* thought to explain immunity to error through misidentification relative to the first-person is recruited as a novel constraint that can be used to adjudicate between competing accounts of the *de se*. Higginbotham argues that his “reflexivist” theory of *de se* thought is the key to explaining immunity to error through misidentification relative to the first person, and should be accepted because it does that explanatory job. Recanati argues that his distinction between two kinds of *de se* thoughts – “implicit” and “explicit” – is the key to explaining immunity to error through misidentification relative to the first person, and should be accepted because it does that explanatory job.

In this chapter, I give reasons to disbelieve Higginbotham and Recanati’s theories. But I also undermine their shared motivational strategy. I do so by arguing that the best explanation of immunity to error through misidentification, whenever it arises and not just in the *de se* case, is what I call *The Simple Explanation*. *The Simple Explanation* just consists in the observation that a judgment will be immune to error through *misidentification* when it is not based on an *identification*. *The Simple Explanation* may not appear very exciting – or perhaps even very explanatory. However, what it says does at least seem true. Moreover, under scrutiny, it turns out that neither Higginbotham’s nor Recanati’s theory goes beyond it – neither explains anything *The Simple Explanation* leaves unexplained. Moreover, *The Simple Explanation* does not itself rule out any candidate theory of *de se* thought. So the

³ “*De se* Thought” is just another phrase that means the same as “First-Person Thinking”. In this chapter, I follow Higginbotham and Recanati in using the phrase “*De se* Thought”.

attempt to use immunity to error through misidentification to gain leverage on the topic of *de se* thought fails.

I discuss Higginbotham first and then Recanati since to some extent Recanati's position is a reaction to Higginbotham's. Before discussing either though, there are a few terminological preliminaries to be dealt with.

This chapter discusses immunity to error through misidentification *relative to the first-person* more than any other kind of error through misidentification. For convenience, I use "fp-immunity" as an abbreviation (and "fp-immune" as an abbreviation for the corresponding adjective).

Higginbotham and Recanati mostly speak of "thoughts" as exhibiting or failing to exhibit immunity to error through misidentification. It is clear that both beliefs and judgments can exhibit immunity to error through misidentification. I take it therefore that "thought" is intended to apply indifferently to beliefs and to judgments. Higginbotham assumes that thoughts are a kind of *state*. I take it that his notion of a *state* is a fairly relaxed one on which even judgments, more often thought as a kind of event, count as states. So he is using "state" in the broad way that I used "episode" in other chapters of this thesis. In this chapter, I follow Higginbotham in referring to judgments as states.

In this paper, I use phrases like 'S's judgment "I hear trumpets"'. This phrase refers to the judgment that would be expressed by an utterance of the sentence "I hear trumpets" produced by S. I assume that S's judgment "I hear trumpets" is different

from S's judgment "S hears trumpets". To make exposition of Higginbotham and Recanati's positions as simple as possible, I shall follow them in also assuming that the reason these two judgments differ is that they have different contents. However, nothing substantial turns on this. Nothing changes if, for example, we instead assume that the two judgments differ because they involve apprehending the same content under different modes of presentation.

6.1 Reflexivism and FP- Immunity to Error Through Misidentification.

Higginbotham's reflexivist account of *de se* thoughts is a distinctive version of the familiar Fregean idea that to think a *de se* thought is to think about oneself using a special first person mode of presentation. The distinctiveness lies in the account Higginbotham gives of the special mode of presentation. According to reflexivism, being in a state that involves thinking about oneself using the first person mode of presentation is being in a reflexive state, i.e. a state that involves thinking about oneself *as the subject of that very state*.

As a preliminary to an assessment of this account, I run through some of the data any account of any kind of immunity to error through misidentification needs to respect, concentrating on a feature – which I shall call basis-relativity – that reflexivism seems to run into trouble with.

There are different *bases* on which I might judge “I hear trumpets”.⁴ For example, I might base that judgment on an auditory perception I am having. If I do, the judgment seems to be fp-immune. But what if I had made the same judgment on the basis of an inference from the following judgments: “The person in the third row hears trumpets” and “I am the person in the third row” (we can imagine that I have come to know both of these premises through testimony)? At least when made on this kind of inferential basis my judgment “I hear trumpets” does not seem fp-immune. If the first premise of the inference had been true, but the second premise had been false, the error I made would have been an error through misidentification.⁵

Once made, the point that immunity to error through misidentification in general, and fp-immunity in particular, is *basis-relative* seems irrefutable. But, as we shall now see, basis-relativity is a problem for reflexivism.

Higginbotham’s introduction of reflexivism, and his assertion that it explains fp-immunity has a slightly blink-and-you-miss-it character. He asks:

What is the reason for immunity to error through misidentification in the case of thinking, on the basis of a present perception, “I hear trumpets”? I will assume that it is this: when I am in the relevant perceptual state, what I think is *the subject of that state hears trumpets*. Hence, there can be no question of my *identifying* myself as the subject of the state (2003: 13).

⁴ For the first appearance of the “I hear trumpets” example that I, following Higginbotham following Campbell, discuss, see Campbell 1999.

⁵ One might quibble with this on the grounds that “the person in the third row” (unlike, for example, “that man”) is not a genuine singular term and that, as a result, my judgment “I am the person in the third row” does not count as a genuine *identification*. As I noted at the beginning of this paper, one kind of discussion in which immunity to error through misidentification appears homes in on the question of how exactly the key notions in the area – including especially the notion of an identification – ought to be defined, and what exactly their extension is. I do not wish to become embroiled in that kind of discussion in this paper. I continue to use my judgment “I am the person in the third row” as an example of an identification, but anyone who has doubts about this is encouraged to substitute a different example (e.g. my judgment “I am that man”).

I begin by addressing a fairly rudimentary question. The passage mentions what seem to be two different states – an (auditory) perception and a judgment. Which of these do the words “*that state*” in the fourth line refer to? It is far more natural to read them as referring to the perception. But, if so, there is a question about why the *de se* judgment should be thought to count as *reflexive*. The content of the judgment is that the subject of *that state* (i.e. the perception) is hearing trumpets. If the perception is distinct from the judgment, then the judgment does have a content that refers to *a* state of the same subject, but it does not have a content that refers to *itself*. So it is not a reflexive judgment.

My guess is that Higginbotham would say that one and the same token-state is both perception and judgment, so that the judgment, in having a content that refers to the perception, thereby has a content that refers to itself. Perhaps this is a plausible view about the metaphysics of states. It is, in any case, the most plausible interpretation of the passage that occurs to me. However, importantly, nothing in this paper depends on this way of interpreting the passage. Things do not change if we instead suppose that the perception and the judgment are distinct, and that “that state” refers to the judgment; or that the perception and the judgment are distinct, and that “that state” refers to the perception. The reason these matters need not detain us is that the most straightforward and compelling objection to reflexivism – what I call the basis-relativity objection – operates at a fairly high level of abstraction. It trades only on the fact that Higginbotham’s reflexivism is, as he himself puts it, “a solution in terms of logical form”.

The basis-relativity objection does not even require us to have any very definite ideas about what logical forms are. It is enough if we can assume that any two tokens of the same judgment will have the same logical form. If so, then the mere fact that two tokens of the same judgment are made on different bases will not cause them to have different logical forms. So logical form is not basis-relative. But immunity to error through misidentification in general, and fp-immunity in particular, is basis-relative. So reflexivism's *explanans* and its *explanandum* seem mismatched. I flesh this objection out below as a trilemma, again taking my "I hear trumpets" judgment as my example of a *de se* judgment.

My judgment "I hear trumpets" has the same logical form, regardless of what basis it is made on. (i) If the logical form of that judgment, whatever it is, guarantees that it will be fp-immune, then we will have to deny, implausibly, that there is any basis on which that judgment can be made relative to which it *is not* fp-immune. (ii) If the logical form of that judgment, whatever it is, guarantees that it will *not* be fp-immune, then we will have to deny, implausibly, that there is any basis on which that judgment can be made relative to which it *is* fp-immune. (iii) If the logical form of that judgment, whatever it is, neither guarantees that it will be fp-immune, nor guarantees that it will not be, then the question of why fp-immunity occurs will not have been answered "in terms of logical form". On the contrary, the true explanation of fp-immunity will have to appeal to something – e.g. facts about the basis on which the judgment is made – that can vary between token-judgments whose logical form is the same. From the point of view of explaining fp-immunity, reflexivism, *or any other thesis about logical form one might care to advance*, then becomes an idle wheel.

The point about immunity to error through misidentification being basis-relative is pretty well-established.⁶ It would be surprising if such a well-established point should refute reflexivism, and do so in such a straightforward way. How might the reflexivist respond?

6.2 Saving Reflexivism

Since the basis-relativity objection to reflexivism is a trilemma, there are three possible responses to it, corresponding to the trilemma's three horns. I cannot say which of these horns Higginbotham might go for, since he nowhere explicitly considers the basis-relativity objection.⁷ The three possible responses are, in any case, as follows.

One response would be to challenge the “idle wheel” point in (iii). The reflexivist could say that the explanation of why a given *de se* judgment exhibits fp-immunity will have to advert to facts about the basis on which that judgment is made, but *also* to the fact that that judgment is a reflexive judgment. Merely mentioning facts about the basis on which the basis is made does not give as good an explanation. So the hypothesis about logical form is not a mere idle wheel. Rather, it is part – and if so

⁶ This is not to deny that coming to see this point constituted an important advance in our understanding of immunity to error through misidentification. Evans (1982) points out that Wittgenstein's having missed it shows up in his assumption that his distinction between uses of the first person “as subject” and “as object” lines up with a distinction between different kinds of *sentence*.

⁷ I note though that Higginbotham in the passage quoted above *does* specify the basis on which the “I hear trumpets” judgment he uses as an example is made. This strongly suggests (what anyway seems highly likely) that he is well aware that fp-immunity *is* basis-relative – for if fp-immunity were not basis-relative, there would be no reason for Higginbotham to complicate his example by mentioning facts about the basis on which the judgment is made. For me, this makes Higginbotham's paper exegetically rather puzzling. Has he considered the basis-relativity objection to reflexivism and decided on what his response to it will be? If he has, why does he nowhere say what that response is?

then, undeniably, a surprising and interesting part – of an explanation of fp-immunity that also pays heed to basis-relativity.

In one way, this option is highly desirable, since it does not force the reflexivist to make any implausible denials. However, left thus schematic, there is no reason at all to think it achievable. It is not as though we have some general reason to think that, by hook or by crook, reflexivism must play a role in explaining fp-immunity, even in advance of hearing the details about how exactly it does that.

Another option is to accept the “implausible denial” in (ii) and say that the logical form of *de se* judgments guarantees that they will *not* exhibit fp-immunity. It is clear that this has no attraction for the reflexivist. It is implausible, and it also severs the connection between reflexivity and fp-immunity.

The only option left is to accept the “implausible denial” in (i) and say that *de se* judgments *are* guaranteed to exhibit fp-immunity. On the face of it, this seems hopeless, since it clashes with the apparent datum that fp-immunity is basis-relative. However, Recanati helpfully suggests a position Higginbotham might take to make this option less implausible (he also tentatively suggests that Higginbotham does in fact adopt that position). Recanati’s suggestion, as I understand it, is for the reflexivist to advance a revisionary view about which token judgments count as *de se*. In particular, token judgments that are fp-immune – like a token judgment of “I hear trumpets” or “My legs are crossed” that is *based directly on an experience* – will count as *de se*. But token judgments that are not fp-immune – like a token judgment of “I hear trumpets” or “My legs are crossed” that is *made on some suitably indirect*

basis – will not count as *de se*, or at least will not count as “primarily or exclusively” *de se*.

A number of things make this revisionary view difficult to accept. If the judgment “My legs are crossed” does not involve the same exclusively *de se* way of thinking of oneself regardless of what basis it is made on, it is hard to see why the phrase ‘the judgment “My legs are crossed”’, which pointedly includes a token of the first person, is a natural way of referring to that judgment, regardless of what basis it has been made on. And it is hard to see why it is natural to *ascribe* that judgment, regardless of what basis one imagines it has been made on, using a sentence like “He judges that *he himself* has crossed legs”. Perhaps it could be suggested that these things only seem natural because almost everyone has made an error about which states belong together as involving the same distinctive way of thinking of a person. But then the question is why everyone should be making that error.

A further difficulty can be brought out by looking at inferences like the following one:

I hear trumpets (based directly on an experience)

My legs are crossed (made on some more indirect basis relative to which it is not fp-immune)

So, somebody whose legs are crossed is hearing trumpets

If the two premises really did involve *different* ways of thinking of a person – the first an exclusively *de se* way and the second some not exclusively *de se* way – it seems

there would be an epistemic risk, which the subject would have to do something to rule out, that the two different ways of thinking the subject employs are in fact ways of thinking of two different people. But there does not seem to be any such epistemic risk in this inference. This is just what we would expect if we thought that both premises involve exactly the same – *de se* – way of thinking of a person, despite the fact that one is made on a basis relative to which it is fp-immune and the other is not.

Perhaps it could be replied that although the inference involves two different ways of thinking of a person, the special circumstance that both of them are at least “somewhat *de se*” means that there is no epistemic risk of their being ways of thinking of the different people. But this reply faces problems of its own. If the two ways of thinking are epistemically guaranteed to be ways of thinking of the same person, in what sense are they really *different* ways of thinking of a person? Whatever difference there is between them seems to lack cognitive significance. Perhaps it is a bit like the difference between “horse” and “steed”, or “dog” and “cur” – a difference in *tone*. But it is hard to see how appealing to that kind of difference could be relevant here. After all, the difference between *de se* thoughts and non *de se* thoughts is manifestly a cognitively significant difference, rather than a mere difference in tone.

For these reasons, I think the revisionary view is not very plausible. Still, we should at least ask if there’s anything that might make it seem plausible. Recanati’s sketch of the position is as follows:

On Higginbotham’s account as I understand it, failures of immunity in first person thoughts are...traced to the fact that the subject does not primarily or exclusively think of the person the thought is about as $\sigma(e)$ – i.e. as the subject of that state (*DM*) –, but in some other way. The use of the first person in such cases is justified by the additional belief that the object identified in that other

way is identical to $\sigma(e)$. So a *de se* belief is involved, and *it* exhibits the property of immunity which derives from employment of the reflexive concept $\sigma(e)$, but it is only one element in a complex belief state which involves also non- *de se* elements (such as *the person in the mirror*). If the speaker only thought of himself, i.e. $\sigma(e)$, as having his legs crossed, his self-ascription would be immune to error through misidentification; but the subject who sees himself in the mirror also thinks of himself as *the person in the mirror*, and that is the source of the failure (2007: 178-9).

I tentatively suggest that an ambiguity in “belief-state” (unconnected to the familiar type-token ambiguity) may make the view sketched in this passage look more defensible than it actually is. In one sense, a belief-state is just a state of holding a particular belief – e.g. the state of believing that today is Tuesday. Call this a *unit* belief-state. In a different sense, one’s belief-state *at a time* is a composite of all the unit belief-states one has at that time. Call this a *total* belief-state.

It is clear that anybody’s total belief-state, at any time, is going to be massively complex, and in particular that anybody’s belief-state will involve many different ways of thinking about oneself. The revisionary view says that when one believes “My legs are crossed” on the basis of an inference from premises like “The person in the mirror’s legs are crossed” and “I am the person in the mirror”, the belief-state one ends up is complex, and involves both a *de se* element and the non *de se* element “the person in the mirror”. In one sense of “belief-state” – the total belief-state sense – this is true. But the revisionary view is clearly a claim about unit belief-states – it is a claim about unit belief-states that seem to be *de se* despite not being fp-immune.

Moreover, the fact that one’s *total* belief-state is complex is no reason to think that the *unit* belief-state of believing “My legs are crossed” on the basis of the inference mentioned above is complex. After all, one’s *total* belief-state would still have been

complex even if one had at that moment instead been judging “My legs are crossed” *on the basis of an experience*. Total belief-states just are complex. Still, if one equivocates on “belief-state”, the revisionary view may gain some undeserved plausibility through confusion with the correct observation that, when one judges “My legs are crossed” on the basis of the inference mentioned above, one’s total belief-state is complex, and involves both *de se* and non *de se* elements.

Sometimes one’s inferential route to a *de se* judgment involves thinking about oneself in a non *de se* way – e.g. as the person in the mirror. This does not provide any reason to think the judgment itself is not “exclusively” *de se*. To think it does is to confuse a fact about the *basis* on which a judgment is made with a fact about what its *content* is. So we should not accept the revisionary view about which states count as *de se*. The revisionary view was the only developed response to the basis-relativity objection to Higginbotham’s argument for reflexivism. So we have no reason to accept reflexivism.

6.3 Explaining Immunity to Error Through Misidentification: The Simple Explanation Vs The Reflexivist Explanation.

Does not accepting reflexivism involve sacrificing anything? Does it, in particular, involve sacrificing an explanation of fp-immunity (the view’s alleged capacity to explain that phenomenon was the main point offered in its favor)? I shall address this question by first considering an explanation of immunity to error through misidentification in general which is entirely neutral with respect to the various competing theories of the *de se*.

Judgments exhibit immunity to error through misidentification when they are not based on identifications. (*The Simple Explanation*).

No one should deny the truth of *The Simple Explanation*.⁸⁹ Certainly Higginbotham does not. Higginbotham even seems to accept that *The Simple Explanation* is *part of the explanation* of immunity to error through misidentification. He says:

It is obvious that immunity to error through misidentification must arise, not because identification is infallible, but because there is no question of identification at all (2003: 21).

However, Higginbotham cannot regard *The Simple Explanation* as the whole explanation of immunity to error through misidentification. For he also appeals to reflexivism, and he says that *it* is the explanation of fp-immunity. Presumably, then, he would say that *The Simple Explanation* takes for granted something that in fact can, and must, be explained.

What *The Simple Explanation* most obviously takes for granted is the fact that some judgments are not based on any identification. So, in deciding whether reflexivism, or any other explanation of a given judgment's fp-immunity, goes beyond *The Simple*

⁸ Perhaps *The Simple Explanation* should be seen as a special case of a more general principle: judgments (or, more generally, actions) are immune to error (or, more generally, miscarriage) through ϕ -ing, when they do not involve any ϕ -ing. Another special case of the same principle would be: attempts to find one's way home are immune to error through misuse of a compass, when they do not involve the use of a compass.

⁹ How much ought to be said about the notion of an identification? I say as little as possible. An identification is a kind of belief. But the literature on the topic brings out that many beliefs are *borderline* cases for being identifications. E.g., is my "I am the man in the third row" belief an identification or does the fact that "the man in the third row" is a description rather than a genuine singular term mean that it isn't one? Quite plausibly, *The Simple Explanation* has different precisifications, corresponding to different precisifications of "identification" and "misidentification". Importantly though, so long as the precisifications of "identification" and "misidentification" are coordinated, *The Simple Explanation* will still come out true on all precisifications.

Explanation, the question we need to ask is this: does the explanation *explain* the fact that the judgment in question is not based on any identification? This question turns out to be quite an effective dialectical instrument.

So, does the fact that some judgment which is not based on an identification is also reflexive explain *why* that judgment is not based on an identification? On the face of it, it does not. *For reflexive judgments can be based on identifications, just like any other kind of judgment.* To see this, set to one side any possible connection between *de se* judgments and reflexive judgments. Just suppose I perform the following inference whose conclusion is an explicitly reflexive judgment.

The person in the mirror has crossed legs.

The subject of this state is the person in the mirror.

The subject of this state has crossed legs.¹⁰

Despite being reflexive, the conclusion of this inference *is* based on an identification. So some reflexive judgments *are* based on identifications, even if others are not. So merely pointing out that a given judgment that is not based on an identification is also reflexive does not on its own do anything to explain *why* that judgment is not based on an identification.

¹⁰ Perhaps it could be argued that this inference needs to be supplemented with an identify premise to the effect that the subject of the first reflexive state (the second premise) is the same as the subject of the second reflexive state (the conclusion). However, this is not a promising line for the reflexivist to take. On the reflexivist view, the inference is a more explicit version of a corresponding *de se* inference of the form "The g is F. I am the g. So, I am F". And, as Campbell (1994) convincingly shows, it is not plausible that such inferences need to be supplemented with the any identity premise of the form "I am I".

Recanati suggests a refinement to reflexivism that might appear to help with this problem. The refinement is for the reflexivist to insist that the judgments he is interested in, and wishes to associate *de se* judgments with, are only those reflexive judgments in which the concept “the subject of this state” is used “as it were, *attributively* (2007: 178)”.¹¹ Recanati suggests that *attributive* reflexive judgments *will* be guaranteed not to be based on an identification. The reflexive judgment that is the conclusion of the inference we have just looked at is not a counter-example since it is reflexive but it is not attributive.

I agree that a judgment’s being an attributive reflexive judgment *guarantees* that it will not be based on an identification. Still, I do not think that the former fact can be what *explains* the latter fact. As the notion of attributiveness is usually explained, and as Recanati himself explains it, it is definitional that a descriptive judgment is *attributive* just in case it is not based on any identification of the object that satisfies the judgment’s descriptive concept (e.g. an identification like “Jones is Smith’s murder”, in the example of Donnellan’s Recanati mentions, or like “The subject of this state is the person in the mirror” in the inference above).¹² If so, then “being an attributive reflexive judgment” and “being a reflexive judgment not based on an identification” do not pick out distinct properties, of which the instantiation of one might explain the instantiation of the other. They are just different names for the same property.

¹¹ I take it that the “as it were” is a gesture to the fact that attributiveness is usually predicated of uses of descriptions in language and not, as here, of occurrences of descriptive concepts in thought.

¹² See Donnellan 1966.

The upshot is that the original conclusion stands. Sacrificing reflexivism does not mean sacrificing any explanation of fp-immunity.

6. 4 Recanati, Implicit *De Se* Thoughts and FP-Immunity.

In rejecting reflexivism, I am following Recanati.¹³ Recanati's own theory is that *de se* thoughts divide up into two importantly different sub-classes. One sub-class is the class of *explicit de se* thoughts. These are analyzed as attitudes to classical propositions – entities that have truth-values relative to worlds but not relative to anything else. What sets such thoughts apart from non *de se* thoughts is that they involve a special first person mode of presentation (this is the analysis Frege and Higginbotham provide of *all de se* attitudes).

The other sub-class of *de se* thoughts is the class of *implicit de se* attitudes. These are analyzed as attitudes to person-relative propositions (this is the analysis David Lewis provides of *all de se* attitudes).¹⁴ Person-relative propositions do not contain any mode of presentation of a person, first personal or otherwise. But they are evaluated relative to persons.¹⁵

¹³ Recanati's reasons for rejecting reflexivism are, I think, less clear-cut than the ones I have given. They depend on how exactly one describes what is going on in the kind of alien-thought-insertion delusion suffered by some schizophrenics. See 2007: 180-188.

¹⁴ See Lewis 1979.

¹⁵ The very different treatment of the two kinds of *de se* attitudes means that Recanati's theory is committed to the existence of entities posited by very different theories. It will be objectionable to you if you have qualms about person-relative propositions. But it will *also* be objectionable to you if you have qualms about the notion of a special first person mode of presentation. Since many fans of the first of these two kinds of theoretical entity are primarily motivated by suspicion of the second of them (and *vice versa*), this already rules out quite a few potential adherents to Recanati's theory. For example, it rules out Lewis, one of whose motivations for introducing person-relative propositions is precisely so as to be able to capture what is distinctive about *de se* attitudes *without* committing to the existence of a special first person

Recanati says that implicit *de se* thoughts are thoughts that do not involve any concept or mode of presentation of the self “in their explicit content” or, equivalently, “on the *lekton* side”. He puts things this way because, in addition to the notion of explicit content, he has a second notion of content: the notion of complete content. Intuitively, there must be some sense in which when I have an implicit *de se* thought, whose explicit content is just a person-relative proposition, my thought is still about, or at least concerns, me. After all, the person-relative proposition does end up getting evaluated relative to me, rather than relative to Stevie Wonder, and the proposal would not be believable if this did not happen. Recanati captures this datum by saying that despite the fact that an implicit *de se* thought’s explicit content is a person-relative proposition, its ‘complete’ content is a classical proposition, which *will* be about some one individual.¹⁶

Recanati offers the same positive motivation for Recanati’s view as Higginbotham offered for Higginbotham’s: the view is supposed to provide an explanation of fp-immunity. More specifically, Recanati’s idea is that fp-immunity can ultimately be traced back and explained in terms of implicit *de se* thoughts – so long of course as

mode of presentation. And it also rules out Frege, who would have said that no attitudes are correctly analyzed as attitudes to person-relative propositions.

On the other hand, this might just be thought of as a reflection of the fact that Recanati’s view is very original. Moreover, it should also be pointed out in its favour that *if* Recanati’s view is correct, there may be pay-offs in other areas in philosophy. For example, there is a fairly lively, but inconclusive, debate about whether considerations in the philosophy of language – e.g. data about how we use phrases like “says that” and “agrees that” – force us to accept that some attitudes are attitudes to something other than classical propositions (see Cappelen and Hawthorne 2008). It would certainly be of interest to participants in this debate if it could be shown that, however exactly the linguistic data pans out, we have to accept that some attitudes are attitudes to person-relative propositions in order to explain fp-immunity, a phenomenon studied by philosophers of mind.

¹⁶ So, perhaps somewhat confusingly, the adjective “explicit” turns up twice in the theory. Explicit *de se thoughts* are a kind of *de se* thought. Explicit *content* is a kind of content, a kind of content that all thoughts, including implicit *de se* thoughts, have. Implicit *de se* thoughts have person-relative propositions as their explicit contents.

there are such things as implicit *de se* thoughts. If so, then the need to explain fp-immunity gives us a good reason – an inference-to-the-best-explanation type reason – to think that there are such things as implicit *de se* thoughts:

The distinction between implicit and explicit *de se* thoughts accounts for the observation that some, but not all, of our *de se* thoughts have the property of immunity to error through misidentification. When, looking at a mirror and seeing my legs crossed, I think ‘My legs are crossed’, I entertain a *de se* thought – a thought about myself in which I am thinking of myself ‘in the first person’ – yet the thought is *not* immune to error through misidentification... immunity to error through misidentification follows from the fact that (only) implicit *de se* thoughts are identification-free, since they do not involve the concept of self on the *lekton* side (2007: 177).

Is it true that accepting that there are such things as implicit *de se* thoughts allows us to give an explanation of fp-immunity that goes beyond *The Simple Explanation*? Any explanation that goes beyond *The Simple Explanation* will have to explain what *The Simple Explanation* takes for granted – the fact that some judgments are not based on any identification. Recanati’s attempt to explain fp-immunity appeals to the putative fact that some judgments are correctly analyzed as attitudes to person-relative propositions. So, in assessing Recanati’s explanation, the key question we need to focus on is this one:

Does the fact that certain *de se* judgments are correctly analyzed as attitudes to person-relative propositions explain the fact that those judgments are not based on any identification?¹⁷

This is the passage that I understand to contain Recanati’s argument for thinking the answer to that question is ‘yes’:

¹⁷ If the answer to this question is “yes”, there is still a further question about how the apparent fp-immunity of some *explicit de se* thoughts is to be explained. Since Recanati has told me his answer to this question has changed since Recanati 2007 I set this question aside.

When the belief that my legs are crossed is gained through experiencing my own body from inside, no such mistake (i.e. a mistake of identification – *DM*) is possible.

The explanation of this phenomenon is straightforward, and it involves the distinction between mode and content once again. Information gained on the *proprioceptive/ kinaesthetic mode* – henceforth to be called the *internal mode* – can only be about one's own body... So the (explicit) content of the bodily experience is not a complete subject-predicate proposition involving a certain person (myself) and a property (being cross-legged or having one's legs crossed). The content is a relativized proposition, true at a person, and the internal mode determines the person relative to which that relativized content is evaluated: myself...The content of the judgment, therefore, is not the proposition that *my* legs are crossed, strictly speaking...Since I am not explicitly 'identified' as the person whose legs are crossed, I cannot be misidentified. The judgment, as Evans says, is identification-free (Evans 1982: 180-1). The identity of the person of whom the property holds is secured by the architecture of the system, which guarantees that "internal" information can only concern the subject's own body (2007: 147).

There seem to be two distinct notions of identification in play in this passage. First, there is the notion Recanati has in mind when he defines implicit *de se* thoughts as thoughts that do not involve a concept of the self on the *lekton* side. Call this identification-1.

A judgment involves an identification-1 just in case it is correctly analyzed as an attitude to a classical proposition that involves a concept of the self.

Second, there is the notion of identification Evans has in mind when *he* says that some judgments are "identification-free". Evans accepts the Fregean view that *all de se* judgments are correctly analyzed as attitudes to classical proposition that involves a concept of the self – so he thinks that *no de se* judgments are free of identifications-1. The notion of identification Evans has in mind is the notion that appears in the

Simple Explanation. Identifications in this sense are a kind of belief – identity-beliefs like “I am that man”.

A judgment involves an identification-2 just in case it is not based on an identity-belief.¹⁸

The Simple Explanation takes it for granted that some judgments do not involve any identifications-2. The question of whether Recanati’s explanation of fp-immunity explains what *The Simple Explanation* takes for granted thus boils down to the following question:

Does the fact that certain *de se* judgments do not involve any identifications-1 explain the fact that they do not involve any identifications-2?

Recanati does not explicitly attempt to answer this question. In what follows, I set out what I take to be the best case for answering this question with a “yes” and then say why I think the best case is still unconvincing. I focus on the judgment “My legs are crossed” (based on a proprioceptive experience), which is one Recanati’s examples of an implicit *de se* judgment.

Recanati’s thinks that the proprioceptive experience the judgment is based on is to be analyzed as an attitude to a person-relative proposition – the person-relative

¹⁸ It is no objection to Recanati that his passage contains two different notions of identification. On the contrary, if the claim that some *de se* judgments do not involve any identification is to provide an explanation of immunity to misidentification that goes beyond *The Simple Explanation*, that claim cannot simply invoke the notion of identification which *The Simple Explanation* invokes. On the other hand, it isn’t totally clear how one would rewrite the passage replacing “identification” (and cognate expressions like ‘identified’) with either “identification-1” or “identification-2”.

proposition “x’s legs are crossed”. He thinks that the judgment itself is to be analyzed as an attitude to the very same person-relative proposition. So, on his account, one and the same person-relative proposition is both the content of the experience *and* the content of the judgment based on the experience. Surely this is enough to explain why the judgment is justified (as it intuitively is). So it explains why the judgment is justified, despite not being based on identification-2. So it explains why the judgment does not *need* to be based on an identification-2. So it goes at least some way to explaining why the judgment *is* not based on an identification-2. So it goes at least some way to explaining what *The Simple Explanation* takes for granted.

There are a number of steps one might challenge here but the one I wish to focus on is the very first. In the background, we have a very plausible idea: that when a judgment simply involves taking an experience *at face value* this fact about the judgment can explain why it is justified. Recanati’s account does seem to capture the fact that the judgment simply involves taking the experience it is based on at face value. For it posits an *identity* in content between the experience and the judgment. The problem is, as we shall now see, that there are other accounts of the transition from proprioceptive experience to judgment that seem equally well-placed to capture the same fact. So accepting what is distinctive about Recanati’s account does not seem necessary in order to explain why the judgment we are looking at is justified. So even if all the other steps mentioned in the previous paragraph are in order, Recanati’s theory does help us to explain anything *The Simple Explanation* takes for granted.

I set aside theories on which experiences lacks contents – since it might be argued that such theories actually do have trouble respecting the fact that some judgments simply

take the experiences they are based on at face value. Consider instead someone who doesn't believe in implicit *de se* thoughts and says that the reason the judgment we are discussing simply involves taking the experience it is based on at face value is that both the judgment and the experience are attitudes to the very same *classical* proposition.¹⁹ This treatment of the transition captures the fact that the judgment simply takes the experience at face value in the very same way that Recanati's treatment does – by positing an identity in content between the experience and the judgment. The only difference between them – about what sort of theoretical entity the shared content is – does not seem relevant to the question of why the judgment is justified.

Perhaps we should understand Recanati as arguing instead for the following weaker, conditional conclusion: if the experience is an attitude to a person-relative proposition, then the judgment must also be an attitude to a person-relative proposition. This makes the case for implicit *de se* thoughts depend on a highly contentious view in the philosophy of perception/proprioception. But it does not deprive it of all interest. There are some philosophers who think that some experiences are attitudes to person-relative propositions, but that other kinds of attitude, e.g. judgments, are not. Showing this “hybrid” view to be untenable would definitely be a worthwhile contribution. But I do not see any way to show that.

The problem is that Recanati himself acknowledges – and, as we have seen, needs to acknowledge – that states have complete contents, as well as explicit contents.

Someone who takes the hybrid view can say that the reason the judgment counts as

¹⁹ As far as I can tell the person-relative/classical distinction among contents crosscuts the nonconceptual/conceptual distinction among contents. So we can imagine this person being a believer in nonconceptual content, or a disbeliever.

taking the experience it is based on at face value is that, although their *explicit* contents are different, their *complete* contents are the same. It is hard to see what can be said against such a position. That taking experience at face value is a source of justification is a pretty commonsensical and intuitive idea. It is hard to believe that the intuitions that support it are philosophically nuanced enough to enforce a demanding conception of taking an experience at face value, on which this requires match in explicit content, over an undemanding conception, on which match in complete content is enough.

I think the claim that we should accept the theory that there are two categories of *de se* judgments – implicit and explicit – *because* doing so provides an explanation of fp-immunity that wouldn't otherwise be available has been refuted. Is there any reason though actually to disbelieve that theory?

There is. It is controversial whether the *explicit* contents of proprioceptive experiences are person-relative propositions, and it is controversial whether the *explicit* contents of *de se* judgments based on proprioceptive experiences are person-relative propositions. But, as already noted, it is not controversial whether the *complete* contents of those experiences, and those judgments, are person-relative classical propositions. Everyone accepts that they are not. This imposes an explanatory obligation on anyone who claims that the explicit contents of proprioceptive experiences, or judgments based on those proprioceptive experiences, are person-relative propositions. They must explain how the discrepancy between explicit and complete content is accounted for. Recanati is gratifyingly upfront about this obligation, and he attempts to fulfill it. The key notion in his explanation of the

discrepancy is the notion of a *mode*, which Recanati introduces as the psychological analogue of an illocutionary force.

The problem is that when we examine the details of this explanation, we see that it *only* applies to *experiences*. It does not extend to the *de se judgments* that are based on those experiences. Since we have no way of explaining how the discrepancy between the explicit content and the complete content of a *de se* judgment arises, we should assume that none does arise. In that case, since the complete content of any *de se* judgment will be a classical content, the explicit content of any *de se* judgment must also be a classical content. What follows fleshes this argument by explaining why Recanati's explanation applies only to experiences, and not also to judgments.

Suppose the explicit content of some experience I have is the person-relative proposition "x's legs are crossed". Why might it nevertheless be the case that the experience's complete content is the classical proposition that my legs are crossed? Recanati's answer lies with the fact that my experience is an experience of a particular type – the proprioceptive type – and that experiences of that type involve what he calls "the internal mode". Information gained using the internal mode is only ever about the subject gaining the information. We can, Recanati thinks, appeal to this fact about the internal mode to explain why the complete content of the experience is a classical proposition about me, even though its explicit content is only a person-relative proposition. As Recanati puts it, the fact that the experience is about me is guaranteed by the "architecture" of the system.

Suppose the explicit content of some judgment I make is the person-relative proposition that “x’s legs are crossed”. Why might it nevertheless be the case that the judgment’s complete content is a classical proposition about me? The answer cannot lie with the fact that my judgment is an experience of a particular type. My judgment is not an experience of any type; it is a judgment. My judgment does involve a particular mode: the mode of judging. But there is nothing especially “internal” about that mode. It certainly isn’t the case that, whenever one judges about anything, one judges about oneself. So, it seems we cannot appeal to the mode my judgment involves to explain the putative discrepancy between its explicit and complete content. Since there is no other plausible explanation in the offing, the null hypothesis – that no discrepancy between explicit and complete content exists – wins by default. In that case, the fact that the complete content of the judgment is a classical content guarantees that its explicit content must be a classical content as well. In that case, there are no implicit *de se* judgments.

Could one block this line of argument by insisting that there is a special mode – the “Judgment Based on Proprioception” mode – that, really, *this* is the mode of my judgment, and that it accounts for a discrepancy between explicit and complete content? This is a disturbingly ad hoc way of individuating modes. More importantly, it is also incompatible with Recanati’s account of modes as the analogues, at the psychological level, of illocutionary forces:

Searle emphasized the similarity between speech acts and mental states such as belief. Both have a propositional component, and both have a dual structure, with the ‘psychological mode’ corresponding, on the side of mental states, to the illocutionary force on the side of speech acts. To assert that *p* is to perform a speech act whose content is the proposition that *p* and whose force is that of an assertion. Likewise, to believe that *p* is to be in an intentional state whose content is the proposition that *p* and whose

psychological mode is that of belief. Now, because of that dual aspect, it is all too easy to ‘misplace’ some of the information carried by a mental state, by ascribing to the content of the state information that is actually carried by its mode (2007: 128).

Regardless of what basis my utterance of the sentence “My legs are crossed” is made on, its illocutionary force is the same – it has the illocutionary force of an assertion (unless, for example, I pronounce it with a distinctive interrogative tone). If the notion of a mode is the analogue, at the psychological level, of the notion of an illocutionary force, then something similar ought to be true of the mode of my judgment that my legs are crossed. The mode of that judgment is the same regardless of what basis it is made on. If the notions of force and mode really are analogues, then modes cannot be sliced more finely than forces. There cannot, in particular, be a special mode corresponding to judgments made on a particular basis.

One might try to avoid this result by severing the connection between modes and illocutionary forces. However, this seems unattractive. The only handle we have on what “modes” are is that they are supposed to be psychological analogues of illocutionary forces. Once that handle is lost, the notion of a mode has been evacuated of all its content. In that case, it will not be able to perform the explanatory role it is intended to in accounting for a discrepancy between explicit and complete content (or any other explanatory role).

Conclusion

Explaining fp-immunity does not require us to accept that *de se* thoughts are *reflexive* (as Higginbotham claims). Neither does it require us to accept that some *de se*

thoughts are *implicit* (as Recanati claims). The most attractive explanation of fp-immunity – *The Simple Explanation* – imposes no tight constraints on an account of the *de se*.

It may also be worth stepping back a little and asking whether there is any good reason to *expect* an explanation of fp-immunity, or of any other kind of immunity to error through misidentification, that goes beyond *The Simple Explanation*. What is it about immunity to error through misidentification that makes it look as though it is crying out for some non-obvious explanation? Very often, one's sense of what *needs* to be further explained, and what can be taken for granted, will partly depend on one's sense of what it might be *possible* to further explain. Is there something about fp-immunity that makes it look as though it might be possible to provide a deeper explanation of it than *The Simple Explanation* provides?

