What is the Problem of *De Se* Attitudes?*

Dilip Ninan
`dilip.ninan@tufts.edu`

1 Introduction

1.1 Skepticism and exceptionalism

It is widely thought that *de se* attitudes (*aka indexical* or *self-locating* atti-
tudes) are special in some way, and that their distinctive features require some
alteration to (once) standard philosophical views about propositional attitudes.
This sort of claim was first made in contemporary philosophy in the 1960s and
1970s by, among others, Castañeda, Perry, and Lewis.[1] This work has proven
influential, and the idea that *de se* attitudes pose a challenge to theories of
attitudes is now the received view.

But despite widespread agreement that there is a problem of *de se* attitudes
('the problem of the essential indexical' to use Perry's term), the literature on
these topics has been less than completely clear on just what that problem
is supposed to be. More specifically, what is unclear is what the *distinctive*
problem of *de se* attitudes is, what problem such attitudes raise over and above
other more familiar problems facing theories of propositional attitudes (e.g.
Frege's Puzzle).

Of course, one reason it might be difficult to unearth a distinctive problem
in this vicinity is that, *contra* the received view, there simply is no distinctive
problem of *de se* attitudes. This is the view of the *de se skeptic*, the philosopher
who flouts the Perry-Lewis orthodoxy and holds that any problem raised by
de se attitudes is really just instance of a more general problem. This in turn
suggests that any problem raised by *de se* attitudes can be resolved by whatever

[1]Castañeda (1966, 1967, 1968), Perry (1977, 1979), Lewis (1979, 1983a, 1986). See also
Geach (1957), Anscombe (1975), Evans (1981), Stalnaker (1981), and Kaplan (1989).

resources are needed to resolve the more general issue. According to this line of thought, there is nothing especially problematic about the *de se*, and so no reason to modify otherwise satisfactory theories of propositional attitudes to accommodate it.[2] Some representative skeptical remarks give the flavor of this view:

> "...attitudes *de se* are simply attitudes *de* their owners." (Boer and Lycan 1980, 432)

> "The received view is that indexical belief presents a *special* problem for the traditional picture of belief, and Perry's arguments show us what this problem is... I deny that the problem has anything special to do with indexical belief narrowly construed. Perry's problem arises quite generally for what is commonly called singular belief." (Spencer 2007, 179-180)

> "...the considerations coming out of the Perry-Lewis tradition give us no reason to change our theory of content." (Cappelen and Dever 2013, 16)

> "...the received view is that there is something particularly problematic about first person thoughts, commonly known as '*de se*'... I think that the received view is a myth..." (Devitt 2013, 133)

> "...the *de se*/*de re* distinction can be satisfactorily accounted for in a representational framework that keeps to the standard conceptions of proposition and belief." (Douven 2013, 288)

> "...the category of *de se* attitudes (if there is one) does not... require any special amendment of our general account of propositional attitudes." (Magidor ming, 44)

Such skepticism might seem surprising: surely there is *something* distinctive about *de se* thoughts, some way in which they constitute a theoretically important class of psychological states. But one should distinguish two claims (cf. Magidor ming, 1): (i) that *de se* attitudes have certain distinctive *features* that differentiate them from non-*de se* attitudes; and (ii) that *de se* attitudes raise a distinctive *problem* for theories of propositional attitudes. Even if the first point is granted, the second does not follow (though presumably the first follows from the second). And it is the second that is at issue here. So even if

---

[2]See Boer and Lycan (1980), Stalnaker (1981), Millikan (1990), Tiffany (2000), Spencer (2007), Cappelen and Dever (2013), Devitt (2013), Douven (2013), and Magidor (ming). It may be that not all of these authors have in mind precisely what I mean here by "*de se* skepticism." Be that as it may, I take myself to be focussing on the central issue in the vicinity.

de se attitudes are distinctive in terms of their epistemology, or in the role they play in motivating action, such facts would not automatically mean they create a distinctive problem for theories of propositional attitudes.[3]

Even with that qualification, one might still think that the *de se* skeptic can be easily answered by simply trotting out some of the standard examples from the literature, examples that purport to illustrate what is problematic about *de se* attitudes. Didn't Perry and friends show us what the problem was by pointing to cases in which amnesiacs wander around libraries, crazy people think they are long dead philosophers, and so on and so forth?

Unfortunately for the *de se exceptionalist*—the philosopher who believes that *de se* attitudes do raise a distinctive problem—such cases often provide grist for the skeptic's mill. Take, for example, a case discussed by both Perry and Lewis: the case of mad Heimson, a mad man who believes that he is David Hume (he has a belief he could express by saying, "*I* am David Hume"). Lewis (1979, 141-142) takes this example to pose a problem for the theory that the contents of attitudes are sets of possible worlds. One problem posed by this case, according to Lewis, is that there are no possible worlds in which Heimson is identical to Hume.[4] Thus, the set of possible worlds in which Heimson is Hume is the empty set, a proposition which clearly doesn't capture what Heimson believes when he believes that he is Hume.[5]

But as the skeptic will be quick to point out, this problem is clearly an instance of a more general one. For Heimson's supportive wife might also believe that Heimson is Hume. The case of Heimson's wife, then, appears to raise precisely the same problem for the possible worlds view that the case of mad Heimson raises, even though her belief is not (or not obviously) *de se*. The possible worlds theory does seem to face a *prima facie* problem here, but it is one that can be seen without looking at *de se* attitudes.

There is more to say about all of this, some of which will get said in due course. But this should give you some sense of the challenge that *de se* skepticism poses to the received view. The main aim of this paper is to clarify that challenge and to see whether it can be answered. I think that it can—I think that *de se* attitudes do raise a distinctive problem for theories of propositional attitudes. But the skeptic is right that many of the arguments that one finds in the literature fail to demonstrate this.

The remainder of this section covers some preliminary matters and sets the terms for the debate. §2 sets out the *prima facie* case in favor of *de se* skep-

---

[3]So there are certain skeptical theses that I won't be discussing here. For example, Cappelen and Dever (2013, Ch. 3) argue at length that *de se* attitudes are not essential to agency (as some have thought). But that issue is of interest to us here only to the extent that it bears on the question of whether *de se* attitudes pose a special problem for traditional theories of attitudes. The reason for looking for a *problem* raised by *de se* attitudes is that Perry and Lewis, for example, attempted to motivate novel theories of propositional attitudes as a response to that (putative) problem.

[4]Or for the counterpart theorist: no worlds in which they have a common counterpart.

[5]Lewis also uses this case to mount a second argument against the possible worlds theory; I discuss this in §2.2.

Note that I use the terms "proposition" and "content" interchangeably.

ticism. §3 is the heart of the paper, where I argue for the exceptionalist thesis that there really is a distinctive problem of *de se* attitudes. Contents have standardly been thought to play a role in the explanation of action and a role in the characterization of inter-personal cognitive relations like agreement. *De se* attitudes reveal that no single object can play both roles. In §4, I examine how various well-known theories of the *de se* can be seen as responding to this problem.

1.2 Radical *de se* skepticism

One reason it might be difficult to unearth a distinctive problem associated with *de se* attitudes is that there is no such thing as a '*de se* attitude.' The *radical de se skeptic* holds that philosophers have not succeeded in using terms like "*de se* attitude" and "indexical attitude" with any clear meaning. The radical skeptic holds something like the following:

> "[The expressions "*de se attitude*" and "self-locating belief"] denote nothing—there is no such thing as essential indexicality, irreducibly *de se* attitudes, or self-locating attitudes... The brilliant papers by Lewis and Perry lead people to believe in, and theorize about, an illusion." (Cappelen and Dever 2013, 3)

This sort of view might seem surprising at first, even absurd. But there is a way of making sense of this kind of radical skepticism. For it might be that the notion of a *de se* attitude is a term that only makes sense in the context of a certain theory. If so, then if one rejects the theory, one might end up rejecting the whole notion of a *de se* attitude. According to this line of thought, *de se* thoughts are the phlogiston of philosophical theories of propositional attitudes.

For example, we might introduce the notion of a *de se* attitude this way: "A *de se* attitude is a thought about oneself when one thinks of oneself via a special, first-person mode of presentation that no one else has access to." If we regard this as a definitional truth about *de se* attitudes, it then becomes a substantive question whether there are any *de se* attitudes. For it is a substantive question whether there are any 'modes of presentation' in Frege's sense. And even if we grant the existence of such entities, it is a substantive question whether there are modes of presentation that are *private* in the manner suggested. The same goes, *mutatis mutandis*, for a definition of an 'indexical thought' according to which such thoughts are identified with 'sentences in the language of thought that contain mental indexicals.' For it is again a substantive question whether there is a language of thought; and even if we grant this, it is a substantive question whether such a language contains indexicals in any interesting sense.[6] A similar problem would beset a definition of *de se* attitudes in functionalist terms, for that would again tie the notion of a *de se* attitude to a particular theoretical framework that one might reject. If our introduction of the notion

---

[6] One of the central themes of Millikan (1990) is that there is no such thing as a 'mental indexical.'

of a *de se* attitude takes any of these forms, the radical skeptic has his foot in the door.

But we needn't—and shouldn't—introduce the notion in any of these ways. I suggest that a more fruitful way to proceed is to characterize the target notion by using examples of the phenomenon we are interested in investigating:

- Mad Heimson has a belief he could express by saying, "I am David Hume." This belief of Heimson's is a *de se* belief. In contrast, Heimson's belief that the man with the red overcoat is David Hume is not a *de se* belief, even if Heimson is the man with the red overcoat.

- Perry (1977, 21-22) considers the case of the amnesiac Rudolph Lingens who is lost in the Stanford library. If Lingens says, "Rudolph Lingens has been to Paris; I have not," we can infer that he does not believe *de se* that he has been to Paris, even though he does believe that Rudolph Lingens has.

- Normally, if I sincerely say, "I want to eat some ice-cream" I would be reporting one of my *de se* desires.

We can characterize the notion of a *de se* attitude ostensively: by pointing to examples of such attitudes, and contrasting them with examples of non-*de se* attitudes. The standard examples in the literature help to bring out the phenomenon we are interested in discussing.[7]

In addition to providing examples of *de se* attitudes, we can also offer rough heuristics for detecting them. For example: if an agent has a belief that he or she could express by uttering a sentence containing a first-person pronoun like "I," "me," or "my," then that belief is probably a *de se* belief. Characterizing the notion using examples and rough rules-of-thumb allows us to remain relatively neutral on the underlying nature of *de se* attitudes, and thus avoids building a controversial philosophical thesis into the very definition of the notion.[8]

It should be noted that these examples and heuristics are intended only to help us get an initial grip on the object of inquiry; no particular example or heuristic should be regarded as sacrosanct. It might be that some of the cases that we initially class as *de se* beliefs turn out not to be (and vice-versa) in the light of more advanced theorizing. For example, it might be that our best theory of psychological attitudes is a functionalist one, and that the functionalist account of *de se* attitudes doesn't line up perfectly with our pre-theoretical judgments about what counts as a *de se* attitude. There are difficult questions here about how much divergence we should allow between initial judgments and theoretical verdicts, but that is an issue we can postpone for the moment.

---

[7]I should note that several *de se* skeptics accept this way of characterizing the notion of a *de se* attitudes. See, for example, Spencer (2007, 179) and Devitt (2013, 1). Douven (2013, 2) even writes that an "informal characterization of the two types of belief [*de se* and *de re*]... suffices perfectly. No one has ever complained that the *de se*/*de re* distinction is vague or ambiguous or otherwise hard to make."

[8]Cf. Williamson (1994, 200-201) on the definition of "vagueness."

It's hard to see how one could object to the term "*de se* attitude" if it is characterized in this way, for it carries with it little substantive philosophical baggage. One can still use the term to say all manner of deflationary things about *de se* attitudes: that *de se* attitudes do not constitute a philosophically or psychologically interesting class; that there is nothing distinctive about the content of *de se* states (there is no such thing as a '*de se* content'); and that *de se* attitudes do not create any special problem for otherwise adequate theories of propositional attitudes.

According to this characterization, the term "*de se* attitude" applies specifically to *first-person* thoughts, thoughts one could express or report using the first-person pronoun. But the term is sometimes in a somewhat broader way according to which it can be used interchangeably with "indexical attitude." As a rough, first-pass characterization, we can say indexical attitudes are those that can be expressed or reported using an indexical like "I," "now," or "here." Maybe this class should be broadened to include other context-sensitive expressions.[9] In any case, we can put this matter aside for the moment. For simplicity, I will mostly focus on first-person attitudes in what follows. I do think much of what I say in this essay can be extended to temporally indexical thoughts, but I leave that extension as a task for future work.[10]

1.3 The doctrine of propositions

I noted earlier that Lewis's discussion of the Hume-Heimson case doesn't make clear what, if anything, is peculiarly problematic about *de se* attitudes. But there is also a second limitation of Lewis's discussion. For even if Lewis's argument had succeeded in showing that *de se* attitudes raised a distinctive problem for the possible worlds theory, he would still only have established a somewhat limited result, viz. that the possible worlds theory was inadequate for handling *de se* attitudes. This would not be without interest, but the general impression one gets from the literature is that the underlying problem of *de se* attitudes

---

[9] But then again maybe such an expansion wouldn't affect matters if all indexical attitudes can be understood as combinations of "I"/"now" attitudes (cf. Perry 1979, 46).

[10] One potential downside of my characterization of *de se* and indexical attitudes is that it may leave the impression that the notion is language-dependent in some way, or that it only applies to agents who speak a language. But in fact my characterization is fully compatible with ways of theorizing about the underlying nature of *de se* attitudes that would justify attributing such attitudes to non-linguistic creatures. For example, one might hold that a *de se* attitude is, at bottom, an attitude with a certain type of functional role, i.e. an attitude with a certain causal profile vis-à-vis stimuli, other attitudes, and action. On such an approach, *one* aspect of the functional role of a given *de se* belief may be a disposition to make certain utterances containing a first-person pronoun. But linguistic behavior of this sort need play no special role in the overall characterization of the belief's functional role.

It is also worth nothing that, for all I've said, it may turn out that there is no theoretically deep connection between indexical *attitudes* as I've characterized them and indexical *expressions* in natural language (e.g. "I," "now," "here"). All we can say at the moment is that indexical attitudes are often expressed by utterances of sentences containing indexical expressions, but his could turn out to be a superficial, contingent feature of such attitudes. That said, I suspect there is some theoretical connection to be made: perhaps a speaker/believer can only express his or her indexical beliefs in a language that contains indexical expressions.

is somewhat broader than that. But how broad is it? What set of assumptions about propositions and attitudes is the problem of *de se* attitudes a problem for? The debate between the skeptic and exceptionalist concerns whether or not *de se* attitudes raise a distinctive problem for otherwise satisfactory theories of attitudes. But then we don't really understand what the debate is about unless we have some sense of what is to count as an 'otherwise satisfactory theory of attitudes.'

Perry (1979, 36–37) tells us that the 'problem of the essential indexical' is a problem for "a traditional way of thinking of belief," a way of thinking that he calls the *doctrine of propositions*. This is a good place to start, since the initial statement of it presupposes little about the nature of propositions. Perry's doctrine consists of three claims, which I render as follows:

(1) TWO-PLACE RELATION

An attitude relation (believing, desiring, etc.) is a two-place relation between an agent and a content.

An attitude (type) like the belief that snow is white is a mental state that consists in having a certain relational property, i.e. the property of standing in the relation of believing to the content that snow is white.[11]

(2) FREGE'S CONSTRAINT

Contents are assigned to attitudes in a manner that accommodates Frege cases. So if a rational agent could have a belief he could express by saying, "Hesperus is bright" without having a belief he could express by saying, "Phosphorus is bright," these two beliefs have different contents.[12]

(3) ABSOLUTENESS

The contents of attitudes are *absolute*, i.e. contents do not vary in truth-value across individuals or times.[13]

Note that this doctrine is largely neutral on the nature of propositions. For all we've said, propositions might be complexes of Fregean senses, or structured propositions containing objects and relations, or sets of possible worlds. So if we could find a problem for this doctrine, it would presumably be a problem for

---

[11]As Fodor (1987, 11) writes, "...the canonical way of picking out an attitude is to say (a) what sort of attitude it is (a belief, a desire, a hunch, or whatever); and (b) what the content of the attitude is..."

[12]Perry (1979, 37) puts this by saying that "propositions must be individuated in a more "fine-grained" way than is provided by truth-value or the notion of truth-conditions." But this would render the doctrine of propositions incompatible with (e.g.) the possible worlds theory of attitudes, something I don't wish to do at the outset. The issue is not so much how fine-grained contents are, but how a theory *assigns* contents to attitudes.

[13]Note that 'absoluteness' in the relevant sense is compatible with the idea that propositions vary in truth-value across possible worlds. I will mostly abstract away from this point in what follows. The idea that propositions have absolute truth-values in this sense is fairly standard; see, for example, Frege (1956), Cappelen and Hawthorne (2009), and McGrath (2014), among many others.

a wide range of approaches to propositions and attitudes, not merely a problem for this or that idiosyncratic approach to these issues.

Note also that the distinctive problem of *de se* attitudes will presumably not be a problem for all theories one can find in the literature, since many prominent contemporary theories of attitudes reject one or more of the above assumptions. This should come as no surprise, since many contemporary theories were constructed in part to accommodate *de se* attitudes. One wouldn't expect a theory to founder on the distinctive problem of *de se* attitudes if the theory was motivated by at least a tacit recognition of that problem.

One last preliminary point. I will sometimes say (have already said) things like this:

> Rudolph Lingens believes that Rudolph Lingens has been to Paris, but does not believe *de se* that *he* has been to Paris.

I assume that, by making such utterances, I can convey two pieces information. First, the information that Rudolph Lingens has a belief with content p, a belief that he could express by saying, "Rudolph Lingens has been to Paris". And second, the information that Lingens does not have a belief with content q, a belief that – had Lingens had it – he would have been able to express by saying, "I have been to Paris." So I am assuming that p and q are different propositions (this is in keeping with thesis (2) above). Note that I am not assuming that the sentence displayed above is literally true in any context, only that it can be used to convey true information.[14]

2 The skeptical challenge

2.1 Famous cases

Our remit is find a problem for the doctrine of propositions, a problem that is specific to the *de se*. To that end, consider two well-known examples from the literature on these topics:

> *The messy shopper*
>
> Perry once followed a trail of sugar along a supermarket floor, looking for the shopper with the torn sack to tell him that he was making a mess. With each trip around the store, the trail became thicker, but there was no sign of the messy shopper. Finally, Perry realized that *he* was the shopper with the torn sack that he was trying to catch. Having realized this, Perry of course stopped following the trail and turned the torn sack upright. (cf. Perry 1979, 33)
>
> *Rudolph Lingens*

---

[14]I use the terms "proposition" and "content" interchangeably. I assume that propositions/contents are the objects of attitudes, but I assume little about their nature.

The amnesiac Rudolph Lingens is lost in the Stanford library. Lingens's amnesia is severe, and he has forgotten who he is. After reading a biography of Rudolph Lingens, he has a belief he could express by saying, "Rudolph Lingens has been to Paris." But at the same time, he also has a belief that he could express by saying, "I have never been to Paris." (cf. Perry 1979, 21–22)

There is, I think, a general reason to be skeptical that either of these cases will reveal what the distinctive problem of *de se* attitudes is (assuming that there is one). For both cases look like instances of something more general; both cases look like Frege cases. Thus, it is easy to provide non-*de se* cases that look structurally identical to these. For example, Cappelen and Dever offer the case of 'messy Superman':

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

The *de se* skeptic then asks: are we really to believe that Perry's messy shopper case poses a problem for the doctrine of propositions that is *distinct* from the problem posed by the messy Superman case?

Similarly, we might imagine that Gustav Lauben knows of a man called "Rudolph Lingens" and also knows that there is an amnesiac wandering through the Stanford library. Failing to realize that 'they' are one and the same, Lauben might (for whatever reason) come to believe that Lingens has been to Paris, while the amnesiac in the library has not. The *de se* skeptic then asks: what problem does the original Lingens case raise that isn't also raised by the Lauben case?

So far this is just a *challenge*. It is open to the *de se* exceptionalist to respond by arguing that the *de se* cases pose a problem for the doctrine of propositions, whereas the non-*de se* cases do not. Or she could argue that both types of cases pose problems for the doctrine, but that the problems are distinct. But given that the two types of case look so similar, it would somewhat surprising if this strategy succeeded.

We can bolster this point by looking at how some have tried to use the *de se* cases to create trouble for the doctrine of propositions. For example, Feit (2008, 28ff.) uses the messy shopper case to argue as follows. If what Perry learned when he learned that he was the one making the mess is an absolute proposition, then it is either a singular proposition or a descriptive proposition. But neither type of proposition appears to fit the bill. So what he learned cannot be an absolute proposition.

Filling this argument in requires us to say why the content of Perry's belief can be neither a singular nor a descriptive proposition. Start with the former point. Consider the singular proposition that Perry is making a mess. Depending on your view of propositions, this might be the set of worlds in which Perry is making a mess, or a structured proposition that contains Perry as a constituent. But as Feit points out, on most views about what it takes to believe a singular proposition, Perry might believe this proposition without believing *de se* that he was making a mess. Perry might, for example, see a reflection of the messy shopper in a mirror, fail to realize that he is seeing himself, and come to believe that (as he would put) *that man is making a mess.* He would thereby believe the singular proposition that he was making a mess, but would still fail to believe *de se* that he was making a mess. So what Perry learned cannot be that singular proposition.

Feit also points out—as Perry himself did—that what Perry comes to believe cannot be a descriptive proposition. Let F be some qualitative property that Perry alone possesses. Could Perry's coming to believe that he was making a mess be his coming to believe that the F is making a mess? There are many reasons for thinking the answer to this is "no." Here is one. Suppose Perry thinks he might be in a 'reduplication universe,' i.e. a universe in which the same set of qualitative properties are distributed in the same pattern across two distinct regions of space. Then there is no qualitative property F that Perry believes is uniquely instantiated. It follows that there is no qualitative property F such Perry believes that the F is making a mess. But all of this is consistent with Perry's believing that *he* is making a mess.

Now one question about this argument is whether it succeeds. There is room for doubt. Neo-Fregeans like Evans (1981) and McDowell (1984) might respond by arguing that there are multiple kinds of singular propositions to the effect that Perry is making a mess, and one of these is such that Perry can only believe it if he has a *de se* belief to the effect that he is making a mess. But even if we put such worries aside, and grant that the argument *does* reveal a problem with the doctrine of propositions, it isn't clear whether the argument is of much use to the *de se* exceptionalist. I say this because, as many *de se* skeptics have pointed out, it isn't clear that the trouble here is specific to the *de se* (Boer and Lycan 1980, 450-453; Tiffany 2000, 38-41; Spencer 2007, 183-184; Cappelen and Dever 2013, §3.1).

To see this, note that we can mimic Feit's argument using the messy Superman case. The narrator of the case learns that Superman is the one making the mess. If what she learned is an absolute proposition, then it is either a singular proposition or a descriptive proposition. But it is neither. So what she learned is not an absolute proposition.

It is not the singular proposition that x is making a mess, where x is Superman. For the narrator already believed this proposition before she believed that *Superman* was making a mess. She believed that singular proposition in virtue of believing that Clark Kent was making a mess.

Suppose F is some qualitative property that Superman possesses uniquely. Could the narrator's learning that Superman was making a mess be her coming

to believe that the F was making a mess? It seems not. For suppose that, like Perry, the narrator believes that she might be in a reduplication universe. Then there is no qualitative property F that the narrator believes is uniquely instantiated. It follows that there is no qualitative property F such the narrator believes that the F is making a mess. All that is consistent with her believing that Superman is making a mess.

If Feit's argument is sound, then it seems likely that the 'Superman' argument is sound as well. And if that is correct, the Feit's argument doesn't reveal what the distinctive problem of *de se* attitudes is.

Let us consider one final case before moving on:

> *The two gods*
>
> "Consider the case of the two gods. They inhabit a certain possible world, and they know exactly what world it is. Therefore they know every proposition that is true at their world. Insofar as knowledge is a propositional attitude, they are omniscient. Still I can imagine them to suffer ignorance: neither one knows which of the two he is." (Lewis 1979, 139)

I think it is a bit harder for the skeptic to construct a non-*de se* case that has the very same structure as this case does. But even so, the skeptic can say two things about this case.

First, it isn't clear whether Lewis's case succeeds in showing what he takes it to show. Stalnaker (1981, 144–145) attempts to defend the possible worlds theory from Lewis's objection. But even if Stalnaker's response fails, it is a further question whether Lewis's case poses a problem for the more general doctrine of propositions (cf. Magidor ming, 11–13). To see why it might not, consider a more theory-neutral description of Lewis's case. Consider the propositions that can, at least in principle, be expressed by sentences that do not contain indexicals. These propositions belong to a certain class, and what Lewis is imagining is a situation in which each god knows all the true propositions in that class, but fails to know some other truths. The advocate of the doctrine of propositions will conclude from Lewis's case that there are true propositions lying outside this class, such as the content of the belief that, were he to have it, the god on the tallest mountain could express by saying, "I am on the tallest mountain." That might show that propositions are not sets of possible worlds, but it isn't obvious that it reveals a flaw in the the doctrine of propositions, which isn't committed to a possible worlds conception of propositions.

Second, even if it is less clear how to construct a non-*de se* version of Lewis's case, it is still a Frege case (or a pair of Frege cases, one for each god) (Cappelen and Dever 2013, §5.4). The god on the tallest mountain knows that the god on the tallest mountain is on the tallest mountain, but doesn't know that *he* is on the tallest mountain. The god's predicament isn't obviously different in kind from that of the mortal who knows that Hesperus is a planet but doesn't know that Phosphorus is. What problem does the case of the ignorant gods pose that doesn't arise in the case of the ignorant mortal?

These considerations suggest that there is no clear route from these sorts of cases to the distinctive problem of *de se* attitudes. The *de se* exceptionalist

should seek that problem elsewhere.[15]

2.2 Internalism and exceptionalism

I have already mentioned Lewis's discussion of the Hume-Heimson example, and considered one argument that Lewis made using that example. Although, that particular argument is of doubtful utility to the *de se* exceptionalist, Lewis also uses that example to mount a second argument against the possible worlds theory. The *de se* exceptionalist might see more promise in this second argument. For although Lewis again takes the target of the argument to be the possible worlds theory, the argument, if successful, would actually cause trouble for the doctrine of propositions. For the real target of Lewis's argument is Absolute-ness, the claim that the contents of attitudes do not vary in truth-value between individuals.

Recall that mad Heimson believes *de se* that he is Hume. Hume presumably also believes *de se* that he is Hume. Lewis then argues as follows:

> "The second problem arises when we ask why Heimson is wrong. He believes he is Hume. Hume believed that too. Hume was right. If Hume believed he was Hume by believing a proposition, that proposition was true. Heimson believes just what Hume did. But Hume and Heimson are worldmates. Any proposition true for Hume is likewise true for Heimson. So Heimson, like Hume, believes he is Hume by believing a true proposition. So he's right. But he's not right. He's wrong, because he believes he's Hume and he isn't.
>
> There are two ways out. (1) Heimson does not, after all, believe what Hume did. Or (2) Heimson does believe what Hume did, but Heimson believes falsely what Hume believed truly." (Lewis 1979, 142)

Lewis then proceeds to argue against the first option as follows:

> "...Heimson may have got his head into perfect match with Hume's in every way that is at all relevant to what he believes. If nevertheless Heimson and Hume do not believe alike, then *beliefs ain't in the head*! They depend partly on something else, so that if your head is in a certain state and you're Hume you believe one thing, but if your head is in that same state and you're Heimson you believe something else. Not good. The main purpose of assigning objects of attitudes is, I take it, to characterize states of the head; to specify their causal roles with respect to behavior, stimuli, and one another. If the assignment of objects depends partly on something besides the state of the head, it will not serve this purpose. The states it

---

[15]I should note that, as a *de se* exceptionalist, it would be fine with me if I were wrong about all this. If *de se* Frege cases do pose a distinctive problem for the doctrine of propositions, then that just gives me a further reason to embrace *de se* exceptionalism, i.e. a reason in addition to the one given in §3 of this paper.

characterizes will not be the occupants of the causal roles."[16] (Lewis 1979, 142-143)

Having ruled out the first option, Lewis notes that his favored theory can make sense of the second. On Lewis's view, the contents of mental states like beliefs are *properties*, things that vary in truth-value over possible worlds, times, and individuals. Hume and Heimson both believe ('self-ascribe') the property of being Hume. So they believe alike. And yet Hume believes truly, since he has the property of being Hume, while Heimson believes falsely, since he lacks the property of being Hume.[17]

As I said, one nice feature of this argument, at least from the point of view of the *de se* exceptionalist, is that it attacks a fairly general assumption about the nature of attitude content, viz. that contents are absolute, things that do not vary in truth-value between individuals. But although the *target* of the argument might be quite general, I suspect that the *appeal* of the argument will be somewhat limited. For the argument clearly depends on a controversial internalist premise that many will reject.

I will return to that point in a moment, but first I want to consider whether Lewis's argument succeeds on its own terms: I want to consider whether Lewis has given the *internalist* a reason to be a *de se* exceptionalist. For even if one accepts Lewis's argument, one might deny that it reveals something special about *de se* attitudes. One might deny this because one might maintain that the argument can be re-run without bringing *de se* attitudes into the picture.

For consider Oscar, here on Earth, where H_2O fills the oceans and falls from the sky, and Twin Oscar, on far-away Twin Earth, where XYZ fills the oceans and falls from the sky (Putnam 1973). We can suppose that Oscar's head and Twin Oscar's head are in 'perfect match' in every way that is at all relevant to what they believe. Suppose that Oscar has a belief he could express in English by saying, "Water contains hydrogen," and that Twin Oscar has a belief could express in Twin English by saying, "Water contains hydrogen." Given Lewis's internalist premise, it would seem to follow that the content of Oscar's belief is identical to the content of Twin Oscar's belief. And yet Oscar's belief is true, while Twin Oscar's belief is false (assuming XYZ contains no hydrogen and that "hydrogen" in Twin English refers to hydrogen). Thus the content they believe cannot be an absolute proposition.[18]

We have here an internalist argument against taking belief contents to be absolute propositions, and the argument invokes beliefs concerning natural kinds

---

[16]What is it for Heimson to get his head into 'perfect match' with Hume's in the relevant sense? Lewis doesn't say, but I assume that it requires them to be in the same brain states or in the same (narrow) functional states.

[17]The view that the contents of attitudes are properties is also defended in Loar (1976) and in Chisholm (1981).

Lewis took properties to be sets of *possibilia*, 'world-bound time-slices.' His view is closely associated with (and often rendered as) the view that the contents of attitudes are sets of *centered worlds*, triples consisting of a world, a time, and an individual.

[18]Spencer (2007, 189ff.) makes a similar point in response to an argument she attributes to Perry. This is her main argument for thinking that indexical attitudes pose no distinctive problem for theories of propositional attitudes.

rather than indexical beliefs. We could also run a similar argument using beliefs expressible with proper names instead. So perhaps the problem Lewis is pointing out is not peculiar to indexical attitudes after all, since we can re-create the problem without invoking such attitudes.

One response to this response is to say that, according to the sorts of internalist theories under discussion, attitudes expressible using natural kind terms and proper names are, in fact, indexical after all. For internalists of this sort often claim that Oscar could express his belief by saying something like, "The colorless, odorless liquid that falls from the sky and fills the lakes *on my planet* contains hydrogen."[19] Twin Oscar could express his belief by uttering a similar sentence of Twin English. The same goes, *mutatis mutandis*, for beliefs expressible using proper names: Oscar's belief that Obama is tall is a belief he could express by saying, "The person *I* know under the name "Barack Obama" is tall." Thus, the fact that beliefs expressible using natural kind terms and proper names pose a similar problem for absolute propositions does not necessarily show that the problem isn't, at root, a problem concerning *de se* belief. For it might be that the former really are *de se* beliefs after all. If this is right, then perhaps Lewis *has* shown that indexical attitudes raise a distinctive problem for standard theories of propositional attitudes; it just turns out that more beliefs count as *de se* than one might have expected.

But even if this is correct, I suspect that many *de se* skeptics will simply reject Lewis's argument by rejecting the internalist premise on which it is based, as Cappelen and Dever (2013, 87 n. 2) and Magidor (ming, 23–25) do.[20] Lewis's argument obviously depends on the premise that if two belief states have the same (narrow) causal role, then they have the same content. Thus, the *de se* skeptic might simply respond by rejecting Lewis's 'same causal role, same content' principle, and so deny that Lewis's argument shows that the contents of attitudes cannot be characterized by absolute propositions.

Note that it would be somewhat strange if it turned out that Lewis's argument was the main or only argument for *de se* exceptionalism. For then the distinctive problem of *de se* attitudes would only really be a problem for the doctrine of propositions *if* we assume internalism about mental content. But it would be surprising to find out that the issue of *de se* exceptionalism was tied so closely to the outcome of the internalism-externalism debate. For while internalism is not all that popular these days, the view that there is a special problem of *de se* attitudes is 'the received view' (as Spencer and Devitt put it). Perhaps this state of affairs simply reveals the extent of our confusion about this topic. But before we settle on that unflattering conclusion, I think we ought to persist a little longer in our search for a problem with a broader reach.

---

[19]Lewis's treatment of *de re* belief suggests he would have been sympathetic to this approach; see Lewis (1979, §XIII). For other relevant discussion, see Putnam (1973, 710), Pollock (1990, 111-112) Jackson (1998, 50), Chalmers (2002, 156), and Brown (2011).

[20]See also Stanley (2011, 95), who makes this response on behalf of the Fregean.

3 The problem of *de se* attitudes

Despite the preceding discussion, I think that *de se* attitudes *do* raise a distinctive problem for standard ways of thinking about propositions and attitudes, a problem that is independent of internalism about content. But to appreciate the problem, we need to expand Perry's doctrine of propositions a bit. More specifically, we need consider two of the theoretical roles that the contents of attitudes are normally thought to play: one in the characterization of agreement, the other in the explanation of action. A problem arises because, given certain features of *de se* attitudes, the contents of attitudes cannot play both roles.

I will begin this section by adding two claims about propositions and attitudes to the original doctrine of propositions; I call the resulting set of claims *the expanded doctrine*. After explaining the expanded doctrine (§3.1), I will argue that *de se* attitudes pose a problem for this doctrine (§§3.2–3.3), and then explain why I think the problem is peculiar to the *de se* (§3.4).

3.1 Expanding the doctrine of propositions

The contents of attitudes are often thought to play a role in characterizing certain interpersonal activities and relations, like communication, agreement, and disagreement. I focus here on agreement. Cappelen and Hawthorne (2009) distinguish between two senses of "agree." They write:

> "The verb 'agree' has a use according to which it picks out a state of some plurality of individuals—where some individuals agree that P if they all believe the proposition that P. There is also a different use according to which it denotes an activity, where agreeing that P is the endpoint of a debate, argument, discussion, or negotiation. On this use, 'agreeing that P' marks an event... The latter use is interactive: it requires that the agents who agree or disagree interact in some way... However, the former use is perfectly applicable to interaction-free pairs of individuals so long as there is some view about the world that they share." (Cappelen and Hawthorne 2009, 60)

We focus here on the state sense of "agree."

Given the doctrine of propositions, a natural account of agreement (in the state sense) would involve the following claims:

(4) AGREEMENT

Agreement is a two-place relation between a group of individuals and a content.

A group of individuals stands in the agreement relation to a content p iff all the members of the group believe p.

Let B_p^x and B_q^y be token beliefs belonging to x and y respectively. Then x and y agree on something in virtue of x's having token belief B_p^x and y's

having token belief B_q^y iff $p = q$.[21]

AGREEMENT only concerns beliefs. But we will also need a notion of agreement that applies to desires. Natural language is less suited to expressing the relevant notion of 'desire-agreement,' but I think there is an intuitive notion here, one that can be explained by way of an example. Suppose that you and I both want Elizabeth Warren to run for President. Then it seems like there is a sense in which we agree: we agree on how we'd like things to be, at least in one respect. We'd like things to be such that Elizabeth Warren runs for President. It is natural to understand this type of agreement in terms of shared content as well. So we can say the following:

> Let D_p^x and D_q^y be token desires belonging to x and y respectively. Then x and y agree on how they'd like things to be (in one respect) in virtue of x's having token desire D_p^x and y's have token desire D_q^y iff $p = q$.

Let us use "AGREEMENT" to refer to the above claims concerning belief along with this claim about desire.

AGREEMENT may be part of the rationale behind ABSOLUTENESS. For suppose that propositions are not absolute, so that some proposition p might be true for you but false for me. Now if AGREEMENT is true, then if you and I both believe p, we ought to count as agreeing on p. But if we agree on something, how can you be right while I am wrong (MacFarlane 2014, 126)? Thus, if AGREEMENT is true, it seems that propositions must be absolute.[22] (Of course, there may be other rationales for ABSOLUTENESS in addition to this one.)

In addition to serving as the objects of agreement, another role the contents of attitudes are often thought to play concerns everyday psychological explanations of behavior. Sally went to the zoo because she wanted to see a panda and believed that the easiest way for her to see a panda was to go to the zoo. Given the doctrine of propositions, we can think of the *explanans* here as consisting of two facts: (i) the fact that Sally has a desire with a certain content q, viz. *that she sees a panda*, and (ii) the fact that Sally has a belief with a certain content p, viz. *that the easiest way for her to see an panda was to go to the zoo*. We

---

[21]A note on notation. B the is relation of believing and D the relation of desiring. If A is an attitude relation, p a content, and x an agent, then A_p is the attitude type that is individuated by A and p, and A_p^x is a token of that type that belongs to x.

(Although the idea that there are token attitudes is fairly standard in the philosophy of mind, some have questioned it; see, for example, Steward (1997, 105–134) and Williamson (2000, 40). But I think the assumption is hard to avoid. When I say that John's belief that Hillary will win is unjustified, I am not attributing a property to the relevant attitude type. For Sue might have a belief of the same type, but her belief might be justified. It seems that "John's belief" refers to something that is specific to John, and I am calling this thing a "token belief.")

[22]Recent work on relativism in the philosophy of language has tried to make sense of the idea of disagreements in which both parties are right (Kölbel 2002, 2004, Lasersohn 2005). Thus, relativists may also want to countenance cases of agreement in which exactly one party is right. But even if this project succeeds, it would constitute something of a revision of standard theoretical approaches to propositions and attitudes, which is enough for my purposes.

explain why Sally does what she does by attributing certain attitudes to her, attitudes that are individuated by an attitude relation and a content.

Of course, understanding *how it is* that bearing an attitude relation to an abstract proposition could play a role in explaining one's behavior would require us to say more about what propositions and attitudes are. I think a promising answer to this question involves saying that the assignment of propositions to attitudes is, at least in a part, a way of saying something about the causal roles of those attitudes vis-à-vis perception, other attitudes, and behavior. The reason causal claims can be inferred from content attributions is that the *semantic* relations between propositions (e.g. entailment) mirror the *causal* relations between attitudes, in something like the way mathematical relations between numbers mirror physical relations between quantities.[23] But we needn't get into that story here; for our purposes, it suffices to note that attitudes are routinely cited in ordinary explanations of action.

A good explanation ought to be something that generalizes, something that leads to by-and-large correct predictions in new cases. If we know that Sally went to the zoo because of her beliefs and desires, then if we learn that Sam has those same beliefs and desires, then it is reasonable for us to expect that Sam too will go to the zoo, so long as other things are equal. Given the doctrine of propositions, Sam 'has the same beliefs and desires' as Sally does just in case he believes and desires the same (relevant) propositions that she does. We can record the above thought as follows:

(5) Explanation

Suppose the fact that x performed action α is explained by the fact that x has beliefs $B^x_{p_1}, ..., B^x_{p_n}$ and desires $D^x_{q_1}, ..., D^x_{q_k}$.

Then, if y has beliefs $B^y_{p_1}, ..., B^y_{p_n}$ and desires $D^y_{q_1}, ..., D^y_{q_k}$, then, other things being equal, y will also perform α.

I will sometimes express the idea behind Explanation a bit loosely, by saying that if two agents have all the same (relevant) beliefs and desires, then, other things being equal, they will behave in the same way.

It is notoriously difficult to explain what it is for 'other things to be equal' in claims like this. But to give you some sense for what the clause includes, I take it that other things are not equal if:

- y is suffering from psychological deficiencies such as weakness of will; or

- y, for whatever reason, is unable to perform action α; or

- y is able to perform α, but performing α is not the most natural, normal, or efficient way for y to reach y's goal.

But, of course, this list is not exhaustive.

---

[23]For relevant discussion, see Churchland (1979, 105), Stalnaker (1984, 8–9), Fodor (1987, 12–14), and Fodor (1990, 13–15), among others.

I take it that the general idea behind Explanation is not specific to psychological explanation *per se*. If you tell me that particle X behaved in a certain way because it has a certain spin and a certain charge, then I will assume that if another particle Y of the same type has the same spin and charge, then, other things being equal, it will behave just as X did. If not, then I will assume that there is some relevant difference between X and Y that explains why they behaved differently. Whatever this difference is, it will presumably point to a factor that was left out of your initial explanation of why X behaved the way that it did.[24]

3.2 The problem

The *expanded doctrine* consists of Perry's initial doctrine of propositions (1)-(3) together with Agreement and Explanation. Note that, like (1)-(3), both Agreement and Explanation are largely neutral on the nature of contents. For all they say, contents might be Fregean propositions, Russellian propositions, or sets of possible worlds. If we can find a problem for the expanded doctrine, then we have found a problem for anyone who hopes that the contents of attitudes can play the two theoretical roles laid down by those two principles, regardless of their other views about the nature of propositions.

Here is the problem as I see it. Agreement would seem to entail if two agents agree on how things are and on how they would like them to be, then those agents have the same beliefs and desires, i.e. they believe and desire the same propositions. And Explanation tells us that if two agents have the same beliefs and desires, then, others things being equal, they will behave in the same way. But then it follows from those two claims that if two agents agree on how things are and on how they would like them to be, then, other things being equal, they will behave in the same way. But, as I shall now argue, this final claim is false. If my argument is sound, it follows that Agreement and Explanation are inconsistent, and that the expanded doctrine is false.

The problem arises in connection with a type of example introduced by Perry and further discussed by Stalnaker.[25] Stalnaker (1999, 19–21) takes such cases to show that *de se* attitudes pose a special problem for the possible worlds theory of content. Here it is shown that examples of this sort pose a problem for a more general class of theories, namely any theory that encompasses the expanded doctrine.

Stalnaker (1999, 20–21) describes the relevant examples as cases involving "different believers who know the ways that they are differently situated, but who, in one sense, do not differ in what they believe." To see what Stalnaker has in mind, imagine that you and I are walking in the woods when a bear begins to chase me. Now suppose that you and I agree on all of the relevant facts of

---

[24]Of course, these claims would need to be qualified to allow for the possibility of probabilistic explanations. But even if psychological explanations are probabilistic, I don't think that this complication will affect the substance of our discussion and so I shall ignore it in what follows.

[25]See the 'bear attack' example—presently discussed at length—from Perry (1977, 23).

the case, and we also agree on how we'd like the situation to turn out. For example, suppose that we agree that I'm being chased a bear. That is, suppose I believe *de se* that I'm being chased by a bear and that you believe '*de te*' that I am being chased by a bear, i.e. you have a belief you could express to me by saying, "You are being chased by a bear." Now if, as I am assuming, we count as agreeing that I am being chased by a bear in virtue of my having this *de se* belief and your having the corresponding *de te* belief, then it would seem to follow from AGREEMENT that these beliefs have the same content. So if we let p_1 be the content of my *de se* belief, p_1 is also the content of your *de te* belief. Letting 'm' denote me and 'y' you, this means that I have a token belief $B_{p_1}^m$ and you have a token belief $B_{p_1}^y$.

The same goes for any other matters on which we agree. Suppose, for example, that we agree that I will be mauled unless I curl up into a ball and play dead, i.e. I believe *de se* that I will be mauled unless I curl up into a ball, and you believe *de te* that I will be mauled unless I curl up into a ball. Then if we let p_2 be the content of my belief, it is likewise the content of yours; so I have a token belief $B_{p_2}^m$ and you have a token belief $B_{p_2}^y$. And suppose that we agree that you are not being chased by a bear, i.e. I believe *de te* that you are not being chased by a bear, and you believe *de se* that you are not being chased by a bear; let p_3 be the object of our agreement on this point. So I have $B_{p_3}^m$ and you have $B_{p_3}^y$. Furthermore, suppose that we both desire that I not be mauled—I desire *de se* that I not be mauled and you desire *de te* that I not be mauled—and let q_1 be the object our shared desire. Then I have desire $D_{q_1}^m$ and you have $D_{q_1}^y$.

Now focus on me for a moment. Given the intuitive description of what I believe and desire, we would naturally predict that, other things being equal, I would curl up into a ball. For I believe that I will be mauled unless I curl up into a ball, and I desire that I not be mauled. Assuming I have no other relevant beliefs and desires, and assuming all else is equal, I will presumably curl up into a ball and play dead. So I let us suppose that this indeed what I do.

Now it is natural to suppose that what *explains* my behavior is just what led us to *predict* that behavior in the first place, viz. the fact that I have the beliefs and desires that I do. I curl up into a ball because of the following facts: the fact that I have a *de se* belief to the effect that I'm being chased by a bear, the fact that I have a *de se* belief to the effect that I will be mauled unless I curl up into a ball, and the fact that I have a *de se* desire to the effect that I not be mauled. In other words, I curl up into a ball because I have beliefs $B_{p_1}^m$ and $B_{p_2}^m$ and desire $D_{q_1}^m$.

But now we have a problem. For we have:

(a) The fact that I curled up into a ball is explained by the fact that I have beliefs $B_{p_1}^m$ and $B_{p_2}^m$ and desire $D_{q_1}^m$.

And we have:

(b) You have beliefs $B_{p_1}^y$ and $B_{p_2}^y$ and desire $D_{q_1}^y$.

But from (a), (b), and EXPLANATION it follows that, so long as other things are equal, *you too will curl up into a ball.* But this prediction seems false. For given what you believe and desire, we simply would not expect that, other things being equal, *you* would curl up into a ball (Perry 1977, 23). For you believe that *I* am being chased by a bear, not that *you* are. Given what you believe and desire we would expect you to call for help, or to attempt to distract the bear, or to shout instructions to me—in short, we would expect you to behave in ways characteristic of someone who believes that *her hiking partner* is being chased by a bear, and not in ways characteristic of someone who believes that *she* is being chased by a bear.

Given the agreement facts in this scenario, AGREEMENT tells us that you and I have all the same beliefs and desires, i.e. we believe and desire all of the same relevant propositions. But given EXPLANATION, it then follows from this that you and I will, other things being equal, behave in the same way. We can even stipulate that all else *is* equal—you are able to curl up into a ball, are not suffering from weakness of will, are not suffering from a sudden bout of irrationality, etc.—in which case it follows from our assumptions that you and I will behave in the same way. That prediction is false. So it appears that either AGREEMENT or EXPLANATION is false. If that appearance is correct, the expanded doctrine must be false.

3.3   Objections and replies

But is that appearance correct? I consider two objections.

Objection 1. There is a sense in which the explanation of my behavior offered above is incomplete. For example, suppose I have a very strong standing desire never to curl up into a ball. If I were to have such a desire, I could have beliefs $B_{p_1}^m$ and $B_{p_2}^m$ and desire $D_{q_1}^m$ without those attitudes leading me to curl up into a ball. So perhaps the complete intentional explanation of my action would need to cite further facts, facts about other beliefs and desires that I have and facts concerning beliefs and desires that I do *not* have. So one could try to resist the foregoing argument by maintaining that, given that you and I behave differently, there must be some difference in our attitudes that went unmentioned.

Reply. While I agree that there is some sense in which the above explanation is incomplete, the basic thought behind the objection strikes me as wrongheaded. Given AGREEMENT, if there is a proposition I believe/desire that you do not, then there is something on which we fail to agree. The objector is thus insisting that the difference in our behavior must be traceable to some matter on which we fail to agree. Either we have slightly different pictures of the objective situation in which we find ourselves, or we have slightly different views about how we'd like that situation to turn out. But this just seems like a mistake. We don't behave differently because there is some aspect of our objective situation about which we fail to agree. The difference that explains our differential behavior is a difference concerning who, in that objective situation, we each take ourselves to be: I take myself to be the one who is being chased by a bear, you

take yourself to be the one whose hiking partner is being chased by a bear. It is this difference in what we believe *de se* that explains our differential behavior. And this difference is consistent with us agreeing on all of the relevant facts about our objective situation.

Objection 2. Some of the discussion in Cappelen and Dever (2013, §3.10) suggests another response. It seems to me that Cappelen and Dever would grant that, in the bear attack scenario, you and I do believe and desire all the same relevant propositions. They also appear to accept EXPLANATION, the assumption that if two agents have all the same belief and desires, they will, other things being equal, behave in the same way (Cappelen and Dever 2013, 52). But how can they maintain both of these claims, given the foregoing argument? Their strategy is to invoke the "other things being equal" clause: you and I can be expected to behave in the same way only if *all else is equal*. And they deny that, in this case, other things really are equal.

We can re-state EXPLANATION as follows:

> Suppose the fact that x performed action α is explained by the fact that x has beliefs $B^x_{p_1}, ..., B^x_{p_n}$ and desires $D^x_{q_1}, ..., D^x_{q_k}$.
>
> Then, if y has also beliefs $B^y_{p_1}, ..., B^y_{p_n}$ and desires $D^y_{q_1}, ..., D^y_{q_k}$, then if C is true, y will also perform α.

where C denotes a conjunction of claims capturing the content of the "all else equal" clause. If any one conjunct of C is false, then x and y might have the same beliefs and desires even while x performs α and y does not. As I noted above, it is a familiar point that clauses like C are difficult to specify in any simple and informative way. But I also noted that C would usually be thought to include the claim that α is an action that y is able to perform. Your lifting a two hundred pound bag of sand might be explained by your beliefs and desires; I might have those same beliefs and desires, but fail to behave as you do because I am too weak to lift such a heavy object.

Cappelen and Dever's strategy is to suggest that the reason you and I behave differently in the bear attack scenario is that different actions are available to each of us. Consider, for example, 'the action *that DN curls up*.'[26] This is an action that is open to me, but not to you, i.e. an action I am able to perform but you are not. And consider 'the action *that you call for help*'. This is an action that is open to you, but not to me. Cappelen and Dever hold that I perform 'the action *that DN curls up*' and that my performing this action is explicable in terms of my beliefs and desires. But since you are unable to perform this action, the "all else equal" clause of the relevant instance of EXPLANATION is triggered, i.e. one of the conjuncts in that clause is false. This means we can maintain that EXPLANATION is true, despite the fact that you and I have the same beliefs and desires but behave differently.

---

[26]This is Cappelen and Dever's way of describing actions; see Cappelen and Dever (2013, 52).

Reply. A notable feature of Cappelen and Dever's proposal is the use of *agent-specific* action types such as 'the action *that DN curls up*'. But do Cappelen and Dever also allow into their ontology *agent-neutral* actions, such as the (much more familiar) action of curling up? Unlike the former, the latter is an action that agents other than me can perform. Cappelen and Dever certainly write in various places as if they do accept action types of this sort (e.g. Cappelen and Dever 2013, 47). And it is a good thing too, since the idea that no two agents can perform the same action is absurd.

But if they do accept the existence of agent-neutral action types, then the argument against the conjunction AGREEMENT and EXPLANATION can simply proceed as before.[27] For all Cappelen and Dever have done is to explain how the following can all be true: EXPLANATION is true; my performing 'the action that *DN curls up*' is explained by my beliefs and desires; you and I have the same relevant beliefs and desires; you do not perform 'the action that *DN curls up*.' Note that here the relevant instance of EXPLANATION is this:

> Suppose the fact that DN performs 'the action that *DN curls up*' is explained by the fact that he has beliefs $B_{p_1}^m$ and $B_{p_2}^m$ and desire $D_{q_1}^m$.
>
> Then if you have beliefs $B_{p_1}^y$ and $B_{p_2}^y$ and desire $D_{q_1}^y$, then if C_1 is true, you too will perform 'the action that *DN curls up*'.

Note that the consequent of this '*suppose..., then...*' discourse is true simply because C_1 is false. C_1 is false because it entails the false claim that you are able to perform 'the action that *DN curls up*'.

But if we have at our disposal the agent-neutral action of curling up, we get *another* instance of EXPLANATION:

> Suppose the fact that DN performs the action of curling up is explained by the fact that he has beliefs $B_{p_1}^m$ and $B_{p_2}^m$ and desire $D_{q_1}^m$.
>
> Then if you have beliefs $B_{p_1}^y$ and $B_{p_2}^y$ and desire $D_{q_1}^y$, then if C_2 is true, you too will perform the action of curling up.

Since you *can* perform the action of curling up, there is no parallel objection to the truth of C_2. And so our argument proceeds as before: we have a scenario in which I perform *the action of curling up* because of my beliefs and desires; you have those same beliefs and desires; but you don't curl up *despite the fact that you are able to perform this action*. If all those things are true, this instance of EXPLANATION must be false.

3.4   Is the problem specific to the *de se*?

I think the foregoing constitutes a genuine problem for the expanded doctrine. And I think this problem constitutes a substantial challenge to *de se* skepticism. For it is not at all obvious how one might construct a nonindexical case

---

[27]Thanks to Eric Swanson for discussion on this point.

that would pose a similar challenge to the conjunction of AGREEEMENT and EXPLANATION. In what follows, I want to bolster this point, by exploring some considerations that suggest that the problem here really is specific to the *de se*.

I begin with two definitions. Consider two agents x and y, x possessing belief B_p^x and y possessing belief B_q^y. Let us say that B_p^x and B_q^y are *agreement-similar* iff x and y count as agreeing on something in virtue of x's possessing B_p^x and y's possessing B_q^y. And let us say that B_p^x and B_q^y are *functionally similar* iff B_p^x and B_q^y play the same causal role vis-à-vis other attitudes and action.

Let me pause here to comment on this notion of functional similarity, since it is slightly different from the notion of *sameness of functional role*. Although sameness functional role implies functional similarity in my sense, the reverse implication may not hold. The functional role of an attitude has a *backward-looking* aspect, which concerns the perceptions and attitudes that can cause it, and a *forward-looking* aspect, which concerns what it, together with other attitudes, can cause. We can think of the functional role of an attitude A as determining a pair of functions, a *backward function* and a *forward function*. The backward function is a function from perceptions and attitudes to A; the forward function is a function from A and other attitudes to actions (and perhaps further attitudes). Functional similarity in my sense is *sameness of forward function*, rather than *sameness of functional role*, i.e. sameness of backward and forward function. It is the former notion that is relevant here, since EXPLANATION concerns only what *actions* an attitude can help to explain.

Two more definitions. Let us say that a pair of token beliefs (b, b') is a *de se pair* iff either b or b' (or both) is a *de se* belief. Let us say that a pair of token beliefs (b, b') is a *non-de se pair* iff neither b nor b' is a *de se* belief.

I shall shortly argue for the following two claims:

(a) If AGREEMENT is true, then if two beliefs are agreement-similar, they have the same content.

(b) If EXPLANATION is true, then if two beliefs are functionally distinct, then they differ in content.

Given these claims, it follows that if we can find a pair of beliefs (b, b') that are agreement-similar but functionally distinct, then either AGREEMENT or EXPLANATION is false. I think our bear attack case furnishes us with a pair of beliefs of that sort. But what is equally important is this: we can vindicate *de se* exceptionalism if we can show that *the only such pairs are de se pairs*. For if only *de se* pairs can be agreement-similar without being functionally similar, then the present problem arises specifically in connection with *de se* attitudes. I now proceed to argue for these claims.

The argument for (a) is straightforward. By the definition of agreement-similarity, if two beliefs are agreement-similar, then the possessors of those beliefs count on agreeing on something in virtue of having those beliefs. And AGREEMENT tells us that if two agents count on agreeing something in virtue of one's having belief B_p^x and the other's having belief B_q^y, then those beliefs

have the same content, i.e. $p = q$. So if Agreement is true, then if two beliefs are agreement-similar, they have the same content.

To see why (b) is true, suppose otherwise: suppose that Explanation is true, that x's belief B_p^x is functionally distinct from y's belief B_q^y, but that $p = q$. Now I take it that to say that B_p^x and B_q^y are functionally distinct just means the following: that there are attitude types $A_{r_1}, ..., A_{r_n}$ and an action α such that x's possessing $B_p^x, A_{r_1}^x, ..., A_{r_n}^x$ would cause x to perform α; but y's possessing $B_q^y, A_{r_1}^y, ..., A_{r_n}^y$ would not cause y to perform α. So suppose that x possesses attitudes $B_p^x, A_{r_1}^x, ..., A_{r_n}^x$ and, as a result, performs action α, and that y possesses attitudes $B_q^y, A_{r_1}^y, ..., A_{r_n}^y$ and does not perform action α. Now presumably x's performing α is explained by x's having $B_p^x, A_{r_1}^x, ..., A_{r_n}^x$, since it is these attitudes which cause x to perform α. Now since $p = q$, $B_q^y = B_p^y$. Thus, y also possesses attitudes $B_p^y, A_{r_1}^y, ..., A_{r_n}^y$. Assuming each of the A_{r_i} is either a belief or a desire, it follows from Explanation that, so long as all else is equal, y will perform action α. But if we stipulate that all else *is* equal, we reach a contradiction, since we were supposing that y did not perform α. So if Explanation is true, B_p^x and B_q^y must differ in content after all—p must be distinct from q.

Now I shall argue that there are *de se* pairs (b, b') where b and b' are agreement-similar but not functionally similar. For example: my *de se* belief that I will be mauled unless I curl up into a ball and your corresponding *de te* belief. These beliefs are agreement-similar: we count as agreeing that I will be mauled unless I curl up into a ball in virtue of our having these beliefs. But these beliefs are not functionally similar. My belief combines with my *de se* desire that I not be mauled to cause me to curl up in to a ball. Let's assume, as seems natural, that you too have a *de se* desire that you not be mauled. My *de se* desire is presumably either type-identical with your *de se* desire or it is type-identical with your *de te* desire that I not be mauled. But your *de te* belief that I will be mauled unless I curl up into a ball does not combine with either of these desires to cause you to curl up into a ball, even though all else is equal. So my *de se* belief that I will be mauled unless I curl up into a ball and your *de te* belief that I will be mauled unless I curl up into a ball are not functionally similar.

If Agreement is true, it follows from (a) that these beliefs have the same content, since they are agreement-similar. But if Explanation is true, it follows from (b) that they do not have the same content, since they are functionally distinct. Since it cannot be that they both have, and do not have, the same content, either Agreement or Explanation is false.

This is essentially the same argument we gave in §3.2 in a slightly different guise. But this way of setting things up allows us to ask whether non-*de se* attitudes pose a similar problem. And, at least *prima facie*, it seems that they don't. For if the elements of a non-*de se* pair are agreement-similar, there is a case to be made that they will also be functionally similar. Suppose, for example, that I believe that Mark Twain is buried in Elmira, i.e. I have a belief that I could express by saying, "Mark Twain is buried in Elmira." What would it be for you to have a belief that is agreement-similar to mine? Presumably, it

would be for you to have a belief that you could express by uttering the same sentence. But here it is quite plausible that our beliefs are also functionally similar. If, for example, I want to visit Mark Twain's grave, my belief might cause me to seek routes to Elmira. But if you too want to visit Mark Twain's grave, your belief would likely have similar sorts of effects on your behavior.

This suggests a potential asymmetry between *de se* beliefs and non-*de se* beliefs:

- There are *de se* pairs (b, b') such that b and b' are agreement-similar, but not functionally similar (and vice-versa).

- For all non-*de se* pairs (b, b'), if b and b' are agreement-similar, they are functionally similar.

If these claims are correct, then the problem for the expanded doctrine that I have been exploring is specific to the *de se* case. Since they seem correct to me, it seems to me that the *de se* exceptionalist is right: *de se* attitudes really do raise a distinctive problem for an otherwise attractive picture of propositions and attitudes.

4 Theories of *de se* attitudes

In §2.1, we noted that many of the famous examples discussed in the literature on *de se* attitudes appear to be Frege cases. I left it open that such cases might pose a problem for the original doctrine of propositions. But (following my skeptical friends) I noted that it seemed doubtful that focussing on *de se* Frege cases would reveal a problem specific to the *de se*.

Focussing on Frege cases has another downside, which is that it is somewhat unclear what the connection is between *de se* Frege cases and the specific sorts of theories of *de se* attitudes that we find in the literature. For example, Heimson might believe *de se* that he is Hume without believing that the man with the red overcoat is Hume (where Heimson is, in fact, the man with the red overcoat). Why would this case lead anyone to reject ABSOLUTENESS, as Lewis does? Why would it lead Perry to reject the idea that believing is a two-place relation to between an agent and a content? Or consider Frege's thesis that the content of a *de se* attitude is *private*, i.e. something that only the owner of the attitude can entertain. Frege's claim was about *de se* (or perhaps indexical) attitudes specifically, not about 'singular' attitudes in general. But why would the Heimson case lead Frege to this surprising doctrine, if the structurally similar case of Heimson's wife (cf. §1.1) would not?

An advantage of the approach taken here is that it can help us to see why we might find in the literature views that have these specific sorts of features. For example, I will presently argue that it is a short step from EXPLANATION to a thesis that is incompatible with the *conjunction* of ABSOLUTENESS and the thesis all contents are *public* or *shareable*. Thus, it is not surprising that we should find in the literature a view like Lewis's, which denies ABSOLUTENESS, or a view like Frege's, which denies the shareability of *de se* contents. Similarly, the

conflict between AGREEMENT and EXPLANATION suggests that no single type of entity can play two theoretical roles traditionally associated with the notion of content. It is not surprising, then, to find theories like Perry's, which assign *two* entities to each attitude, one corresponding to each of these theoretical roles.

My interest here is not primarily historical. I am not suggesting that any of these theorists really arrived at his or her view by reacting to the problem of *de se* attitudes as I have formulated it. Nevertheless, I think it will prove illuminating to show that these disparate theories *can be seen* as reacting to our problem. My aim here is to provide something of a 'rational reconstruction' of the debate over these issues, with the goal of getting clearer on the space of theoretical options.

4.1   Frege and Lewis

As I just noted, one might react to the problem of *de se* attitudes by rejecting the idea that attitude relations are two-place relations between agents and contents. Let us put aside this idea for a moment, and consider the (in some ways) more conservative reaction of retaining theses (1) and (2) of the expanded doctrine. A philosopher who takes this option must then reject either EXPLANATION or AGREEMENT; let us consider philosopher who rejects AGREEEMENT in order to retain EXPLANATION.

What sorts of theories is such a philosopher in a position to accept? Here I will argue that a philosopher of this sort is likely to be led either to a theory like Lewis's or to a theory like Frege's. My argument will proceed in two steps. First, I will argue that it is a short step from EXPLANATION to a principle that I will call "DE SE CONTENT." I will then show that this latter principle conflicts with the conjunction of ABSOLUTENESS and the claim that contents are shareable.

But before I embark on that argumentative journey, it will be good to say a bit more about the idea that contents are shareable. Let us set out that idea as follows:

> PUBLICITY
>
> Generally speaking, the contents of attitudes are *public* or *shareable*, i.e. if an agent x can entertain a content p, then so can any other agent y.

The "generally speaking" qualification is intended to accommodate the fact that there may be some restrictions on just who can entertain which contents. Someone (e.g. a small child) who 'lacks certain concepts' (as we say) may be unable to entertain certain propositions. And if there are acquaintance requirements on having singular or *de re* beliefs, then someone who lacks acquaintance with a particular individual may not be able to entertain certain singular propositions concerning that individual. Neither of these qualifications will matter in what follows.

One motivation for PUBLICITY is that contents must be public if they are to serve as the objects of communication, agreement, and disagreement (cf.

26

Frege 1956, 301-302). Take our principle AGREEMENT; it is consequence of this principle that if x and y agree on something, then there is a content p that both believe. This content p, then, is something that both x and y can entertain—it is shareable, at least in the sense that both x and y can believe it. (Of course, there may be other motivations for PUBLICITY in addition to this one.)[28]

As I mentioned above, Frege famously denied that the contents of *de se* attitudes were public:

> Now everyone is presented to himself in a particular and primitive way, in which he is presented to no-one else. So, when Dr. Lauben thinks that he has been wounded, he will probably take as a basis this primitive way in which he is presented to himself. And only Dr. Lauben himself can grasp thoughts determined in this way. (Frege 1956, 298)

(Frege's 'thoughts' are our contents or propositions.) If PUBLICITY is taken to entail the shareability of the contents of *de se* attitudes, then Frege denied PUBLICITY.[29]

Now for the promised two-part argument. First part: I begin by tracing a route from EXPLANATION to the aforementioned principle (viz. 'DE SE CONTENT'). In §3.4, I argued for the following claim:

(b) If EXPLANATION is true, then if two beliefs are functionally distinct, then they differ content.

In that section, I also argued that my *de se* belief that I will be mauled unless I curl up into a ball is functionally distinct from your corresponding *de te* belief. Now it is natural to suppose that if you accept that claim, then you will also accept a more general claim, something along these lines:

(c) If B_p^x is a token *de se* belief to the effect that x is F, and B_q^y is *not* a token *de se* belief to the effect that y is F, then B_p^x and B_q^y are functionally distinct.

Assuming EXPLANATION, claims (b) and (c) entail the following principle:

DE SE CONTENT

If B_p^x is a token *de se* belief to the effect that x is F, and B_q^y is *not* a token *de se* belief to the effect that y is F, then B_p^x and B_q^y differ in content, i.e. $p \neq q$.

Note that this is equivalent to the following:

---

[28]The Stanford Encyclopedia of Philosophy defines propositions as "the sharable objects of the attitudes and the primary bearers of truth and falsity" (McGrath 2014). If truth and falsity are taken to be monadic properties of propositions, this definition encompasses both PUBLICITY and ABSOLUTENESS.

[29]For developments of Frege's view, see Evans (1981), Peacocke (1981), McDowell (1984), Forbes (1987), Heck (2002), Kripke (2008), and Stanley (2011, Ch.3). For other views that reject PUBLICITY for *de se*-related reasons, see Chisholm (1976), Schiffer (1978), Markie (1988), Hanks (2013), and Ninan (2012, 2013).

> If B_p^x is a token *de se* belief to the effect that x is F, and B_p^x and B_q^y have the same content (so $q = p$), then B_q^y is a token *de se* belief to the effect that y is F.

I use both formulations in what follows. What DE SE CONTENT essentially says is that the property of being a *de se* belief to the effect that one is F is a property of a belief that is reflected in, or built into, the content of that belief. No one else can believe that content unless she also believes *de se* that *she* is F.

We now show that DE SE CONTENT, ABSOLUTENESS, and PUBLICITY are mutually incompatible. To see this, first note that if ABSOLUTENESS is true, then it looks like the following principle is also true:

(*) If x has a belief b with content p, then b is true iff p is true.

Now suppose that I have a *de se* belief to the effect that I am being chased by a bear. As before, we let p_1 be the content of my belief, and so $B_{p_1}^m$ is my belief with content p_1. If PUBLICITY holds, then we would expect that agents other than me—like you, for example—could believe p_1 too.[30] So let's suppose that you do believe p_1, and so $B_{p_1}^y$ is your belief with content p_1. Given DE SE CONTENT, it follows that $B_{p_1}^y$ is a *de se* belief to the effect that you are being chased by a bear.

Now note: our beliefs have different truth-conditions, despite having the same content.[31] My belief $B_{p_1}^m$ is a *de se* belief to the effect that I'm being chased by a bear, so my belief $B_{p_1}^m$ is true iff I'm being chased by a bear. Your belief $B_{p_1}^y$ is a *de se* belief to the effect that you're being chased by a bear, so your belief $B_{p_1}^y$ is true iff you're being chased by a bear. And now we have a problem. For suppose we add the following facts to our scenario: I am being chased by a bear and you are not. Then the truth-condition of my belief $B_{p_1}^m$ is satisfied, while the truth-condition of your belief $B_{p_1}^y$ is not. So my belief $B_{p_1}^m$ is true, while your belief $B_{p_1}^y$ is false. From (*) and the fact that my belief with content p_1 is true, it follows that p_1 is true. From (*) and the fact that your belief with content p_1 is not true, it follows that p_1 is not true. So p_1 is both true and not true. Contradiction. It follows that one of DE SE CONTENT, ABSOLUTENESS, and PUBLICITY must be false.

If one accepts EXPLANATION, one is likely to also accept DE SE CONTENT, given the plausibility of (b) and (c). And if one accepts *that*, one must reject either ABSOLUTENESS or PUBLICITY. From this perspective, it is no accident that one finds in the literature a theory like Lewis's that denies ABSOLUTENESS. Nor is it a surprise to find a theory like Frege's that denies PUBLICITY. If one responds to the problem of *de se* attitudes by hewing to the traditional idea that believing is a two-place relation between agents and contents, then it seems as though one must give up either AGREEMENT or EXPLANATION. If one resolves

---

[30]We can assume that you meet any conceptual/acquaintance requirements for entertaining p_1.

[31]Two token beliefs b, b' have the same truth-conditions (in the relevant sense) iff in any world w in which both b and b' exist, b is true in w iff b' is true in w.

that dilemma in favor of Explanation, then one is very likely to also accept De Se Content. And in light of what we've just seen, this means either rejecting Absoluteness (à la Lewis) or Publicity (à la Frege).

4.2 Perry

In §3.4, I argued for these two claims:

(a) If Agreement is true, then if two beliefs are agreement-similar, then they have the same content.

(b) If Explanation is true, then if two beliefs are functionally distinct, then they differ in content.

I then argued that there are pairs of token beliefs (b, b') such that b and b' are agreement-similar but functionally distinct. Such as: my *de se* belief that I will be mauled unless I curl up into a ball and your corresponding *de te* belief. Suppose p_2 is the content of my *de se* belief. If we say that p_2 is also the content of your belief, we capture the fact that our beliefs are agreement-similar, but not the fact that they are functionally distinct. If, on the other hand, we say that p_2 is *not* the content of your belief, we have the reverse problem: we capture the fact that our beliefs are functionally distinct, but not the fact that they are agreement-similar. No single entity can play both of the theoretical roles that the expanded doctrine assigns to the notion of content.

That suggests a natural (and well-known) solution: assign *two* content-like entities to my (and every) belief. Entity 1 can play the explanation-of-action role; let us call this the *guise* of the belief. Entity 2 can play the agreement role; let us continue to call this the *content* of the belief. This, at least in broad outline, is the sort of view that Perry advocated in his early papers on these topics (Perry 1977, 1979). The idea that *de se* attitudes require us to assign two content-like entities, rather than just one, to each attitude was one of Perry's many insights into these issues.

If guises are to play the explanation-of-action role, then the principle concerning the role of attitudes in the explanation of action will make essential reference to guises. So this approach will reject our principle Explanation in favor of a principle according to which two agents who believe and desire under the same guises will (other things being equal) act in the same way. To formulate that principle, let us use $B^x_{\gamma,p}$ to denote a token belief of x's with guise γ and content p and $D^x_{\delta,q}$ to denote a token desire of x's with guise δ and content q. Using this notation, our amended principle is as follows:

> Guise Explanation
>
> Suppose the fact that x performed action A is explained by the fact that x has beliefs $B^x_{\gamma_1,p_1},...,B^x_{\gamma_n,p_n}$ and desires $D^x_{\delta_1,q_1},...,D^x_{\delta_k,q_k}$.
>
> Then, if y has beliefs $B^y_{\gamma_1,r_1},...,B^y_{\gamma_n,r_n}$ and desires $D^y_{\delta_1,s_1},...,D^y_{\delta_k,s_k}$, then, other things being equal, y will also perform A.

We could then combine this principle with Agreement–or rather, with a slightly-reformulated version of Agreement, since Agreement assumes something that is now being rejected, viz. that believing is a two-place relation. The amended principle might look something like this:

Agreement*

Agreement is a two-place relation between a group of individuals and a content.

A group of individuals stands in the agreement relation to a content p iff for every member x of the group, there is a guise γ such that x believes p under γ.

Let $B^x_{\gamma,p}$ and $B^y_{\delta,q}$ be token beliefs belonging to x and y respectively. Then x and y agree on something in virtue of x's having token belief $B^x_{\gamma,p}$ and y's having token belief $B^y_{\delta,q}$ iff $p = q$.

Unlike Agreement and Explanation, Agreement* and Guise Explanation appear to be compatible. To appreciate this, it might help to think about how this approach applies to the bear attack scenario. Consider my *de se* belief b that I will be mauled unless I curl up into a ball, and your corresponding *de te* belief b'. Since these beliefs are functionally distinct, they will be associated with different *guises*. On Perry's approach, the guise of a belief is the property–or "relativized proposition"–which Lewis takes to be the content of that belief. So the guise of b will the property that x has just in case x will be mauled unless x curls up into a ball; and the guise of b' will be something like the property of having a hiking partner who will be mauled unless he or she curls up into a ball. But since we count as agreeing in virtue of my having b and your having b', these beliefs will have the same *content*. On Perry's approach, the content of b and b' would be the singular proposition that DN will be mauled unless he curls up into a ball. We might identify this with either a set of possible worlds or a Russellian proposition containing DN as a constituent.

4.3   A fourth and final theory

I want to consider one last approach, one which I think has an important advantage over the theories considered so far, but which also faces challenges of its own.

The preceding approaches all face a *prima facie* problem in characterizing agreement. This is perhaps most obvious when we consider Lewis and Frege. If I believe *de se* that I am in danger and you have the corresponding *de te* belief, then, intuitively, you and I should count as agreeing on something. But both Lewis and Frege deny that our beliefs have the same content, and so they are committed to denying Agreement. But what do they propose to put in its place?

It might be thought that Perry's approach enjoys an advantage here; I suggested as much in the previous subsection. But when we look a bit more deeply at Perry's account, it isn't so clear that it correctly characterizes the relevant

notion of agreement. To see what the problem is, imagine that neither you nor I know that I am DN. I believe *de se* that I am in great danger, but that DN is not. You, on the other hand, believe *de te* that I am in no danger, but that DN is. It seems that if we began to speak about these matters, we could quickly find ourselves in conflict.

But given AGREEMENT*, Perry's account would appear to say that you and I count as agreeing that I am in danger. (We would also presumably count as agreeing that I am not in danger.) For there is a guise under which I believe the singular proposition that DN is in danger, and there is a guise under which you believe that singular proposition. Thus, according to AGREEMENT*, we stand in the agreement-relation to that proposition. Now maybe there is a 'coarse-grained' sense of "agree" on which we count as agreeing that I am in danger. I am skeptical of this, but even if we grant it, it seems clear that there is a 'fine-grained' sense of "agree" on which we do not count as agreeing that I am in danger. It is this latter notion which Perry's account seems to have trouble capturing.

Thus, it appears that all three of these approaches to *de se* attitudes face a *prima facie* problem when it comes to characterizing agreement. Maybe the trouble is *merely prima facie*; maybe each of these theories has the resources to overcome this difficulty. Alternatively, maybe the difficulty is not that serious, since perhaps the only robust notion of inter-personal agreement that we have is a coarse-grained one. If so, then it seems to me all of these theories are in the clear, since both Lewis and Frege will likely have the resources to define a suitable coarse-grained notion of agreement.

But there is an alternative approach that appears to avoid these problems altogether. Versions of the approach I have in mind are defended in Stalnaker (2008, Ch.3) and Moss (2012), but I will describe the approach in my own way. Suppose we take the contents of attitudes to be Lewisian properties. Now Lewisian properties are relativistic in the sense that *some* of them vary in truth-value across world-mates. But of course *not all* Lewisian properties vary in truth-value across world-mates. For example, on this approach, the content of my belief that someone owns a typewriter is the property of being such that someone owns a typewriter. This does not vary in truth-value across individuals who inhabit the same world. If this property is true for me, it is true for you, and vice-versa. Properties that do not vary in truth-value across world-mates are called *boring properties*, and a property that is not boring is called *interesting* (cf. Egan 2006, 107). The property of being in danger, for example, is an interesting property, since it might be true for me but false for you.

One way of thinking of the approach I am presently outlining is that any of my *de se* beliefs can be associated with two properties, an interesting property and a boring property. We can, if we like, say that any given *de se* attitude has two *contents*, a boring content and an interesting content.[32] For example, when I believe *de se* that I am in danger, my belief has an interesting content which

---

[32]I think neither Stalnaker nor Moss says precisely this, though it is not too far off from things they do say.

is the property of being in danger. This interesting content corresponds to the functional role of the belief, in the sense that any belief that is functionally distinct from my belief will be associated with a different interesting content. But my belief is also associated with a *boring* content, one that doesn't vary in truth-value between me and you. This content can play a role in characterizing agreement. Suppose, for example, that you believe *de te* that I am in danger. Then this approach will say that the interesting content of your belief differs from the interesting content of mine, since our beliefs are functionally distinct. (This interesting content of your belief would be something like the property of being an x such that x's hiking partner is in danger.) But the boring content of your belief may well be the same as the boring content of mine. This is possible because, even though the boring content of my belief is a Lewisian property, it is a property that is true for me iff it is true for you. So on this approach, agreement is a matter of sameness of boring content.

The reason this is not simply a notational variant of Perry's theory is that, on this approach, the boring content of a *de se* attitude will not, in general, be a 'singular' property. For example, the boring content of my *de se* belief that I am in danger need not be the singular property of being such that DN is in danger (cf. Moss 2012, 230). Thus, the present approach avoids the trouble facing Perry's approach, for the boring content of my *de se* belief that I am in danger will be distinct from the boring content of of your belief that DN is in danger. Thus, we can combine this account with something like AGREEMENT or AGREEMENT* without having to say that you and I agree in virtue of having these beliefs.

This account strikes me as having the virtues of Perry's account without its vice. But this account faces a difficulty of its own. The trouble comes when we ask what the boring content of a *de se* attitude is. What, for example, is the boring content of my *de se* belief that I am in danger? If it is not the singular property of being such that DN is in danger, what is it? For reasons discussed in §2.1, it is doubtful that we could identify the boring content of my belief with any purely qualitative property, such as the property of being such that the F is in danger (for some qualitative property F). For if I think I might be in a reduplication universe, then there will be no qualitative property F that I believe is uniquely instantiated; in that case, there will be no qualitative property F such that I believe that the F is in danger. But if the boring content of my belief is not the singular property that DN is in danger, nor a qualitative property, what else is left?

This looks to me like a serious problem.[33] But even if this particular property-theoretic account is untenable, considering it is instructive in at least one respect. For it reveals what sort of *structure* a satisfactory theory ought to have. It ought to assign two content-like entities to each *de se* attitude, one that captures the functional aspect of the attitude, and another which captures the role the attitude plays in agreement. But in order to do the latter, 'agreement-contents' will need to be assigned to attitudes in a fine-grained manner (e.g. the

---

[33]Though see Stalnaker (2008, Ch.3) and Moss (2012, 230-231) for some discussion.

agreement-content of my *de se* belief that I am in danger ought to be distinct from the agreement-content of your belief that DN is in danger). So even if we don't yet know how to formulate a fully satisfactory theory of these matters, we at least now know what it is we are looking for.[34]

References

Anscombe, G. E. M. (1975). The first person. In S. Guttenplan, editor, *Mind and Language*. Oxford University Press, Oxford.

Boer, S. E. and Lycan, W. G. (1980). Who, me? *Philosophical Review*, **89**(3), 427–466.

Brown, C. (2011). Narrow mental content. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2011 edition.

Cappelen, H. and Dever, J. (2013). *The Inessential Indexical*. Oxford University Press, Oxford.

Cappelen, H. and Hawthorne, J. (2009). *Relativism and Monadic Truth*. Oxford University Press, Oxford.

Castañeda, H.-N. (1966). 'He': A study in the logic of self-consciousness. *Ratio*, **8**, 130–157.

Castañeda, H.-N. (1967). Indicators and quasi-indicators. *American Philosophical Quarterly*, **4**, 85–100.

Castañeda, H.-N. (1968). On the logic of attributions of self-knowledge to others. *The Journal of Philosophy*, pages 439–456.

Chalmers, D. J. (2002). The components of content. In D. Chalmers, editor, *Philosophy of Mind: Classical and Contemporary Readings*. Oxford University Press, New York.

Chisholm, R. (1976). *Person and Object*. Open Court, La Salle, IL.

Chisholm, R. (1981). *The First Person*. The Harvester Press, Brighton.

Churchland, P. (1979). *Scientific Realism and the Plasticity of Mind*. Cambridge University Press, Cambridge, Cambridge.

Devitt, M. (2013). The myth of the problematic *de se*. In A. Capone and N. Feit, editors, *Attitudes De Se: Linguistics, Epistemology, Metaphysics*. CSLI Publications, Stanford.

Douven, I. (2013). The epistemology of *de se* beliefs. In A. Capone and N. Feit, editors, *Attitudes De Se: Linguistics, Epistemology, Metaphysics*. CSLI Publications, Stanford.

Egan, A. (2006). Secondary qualities and self-location. *Philosophy and Phenomenological Research*, **72**(1), 97–119.

Evans, G. (1981). Understanding demonstratives. In H. Parret and J. Bou-

---

[34]My own hope is that *multi-centered worlds* will be of use in formulating an account with the right structure. See Ninan (2010) and Torre (2010) for relevant discussion.

veresse, editors, *Meaning and Understanding*. Walter de Gruyter, Berlin. Reprinted in Evans 1985, 291–321. Page references are to the 1985 reprint.

Evans, G. (1985). *Collected Papers*. Clarendon Press, Oxford.

Feit, N. (2008). *Belief baout the Self: A Defense of the Property Theory of Content*. Oxford University Press, New York.

Fodor, J. A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. The MIT Press, Cambridge, MA.

Fodor, J. A. (1990). *A Theory of Content and Other Essays*. MIT Press, Cambridge, MA.

Forbes, G. (1987). Indexicals and intensionality: A fregean perspective. *The Philosophical Review*, **96**(1), 3–31.

Frege, G. (1918/1956). The thought: A logical inquiry. *Mind*, **65**(259), 287–311.

Geach, P. T. (1957). On beliefs about oneself. *Analysis*, **18**(1), 23–24.

Hanks, P. W. (2013). First-person propositions. *Philosophy and Phenomenological Research*, **86**(1), 155–182.

Heck, R. (2002). Do demonstratives have senses? *Philosophers' Imprint*, **2**(2), 1–33.

Jackson, F. (1998). *From Metaphysics to Ethics*. Clarendon Press, Oxford.

Kaplan, D. (1989). Demonstratives. In J. Almog, J. Perry, and H. Wettstein, editors, *Themes from Kaplan*, pages 481–563. Oxford University Press, New York.

Kölbel, M. (2002). *Truth Without Objectivity*. Routledge, London.

Kölbel, M. (2004). Faultless disagreement. In *Proceedings of the Aristotelian Society*, volume 104, pages 53–73.

Kripke, S. (2008). Frege's theory of sense and reference: Some exegetical notes. *Theoria*, **74**, 181–218. Reprinted in Kripke 2011, 254–291. Page references are to the 2011 reprint.

Kripke, S. (2011). *Philosophical Troubles, Volume 1*. Oxford University Press, New York.

Lasersohn, P. (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy*, **28**(6), 643–686.

Lewis, D. K. (1979). Attitudes *de dicto* and *de se*. *Philosophical Review*, **88**(4), 513–543. Reprinted in Lewis 1983b, 133-159. Page references are to the 1983 reprint.

Lewis, D. K. (1983a). Individuation by acquaintance and by stipulation. *Philosophical Review*, **92**(1), 3–32. Reprinted in Lewis 1999, 373-402. Page references are to the 1999 reprint.

Lewis, D. K. (1983b). *Philosophical Papers, Volume 1*. Oxford University Press, New York.

Lewis, D. K. (1986). *On the Plurality of Worlds*. Blackwell, Oxford.

Lewis, D. K. (1999). *Papers in Epistemology and Metaphysics*. Cambridge

University Press, Cambridge.

Loar, B. (1976). The semantics of singular terms. *Philosophical Studies*, **30**(6), 353–377.

MacFarlane, J. (2014). *Assessment-Sensitivity: Relative Truth and its Applications*. Oxford University Press, Oxford.

Magidor, O. (Forthcoming). The myth of the *de se*. *Philosophical Perspectives*.

Markie, P. J. (1988). Multiple propositions and "*de se*" attitudes. *Philosophy and Phenomenological Research*, **48**(4), 573–600.

McDowell, J. (1984). *De Re* senses. *The Philosophical Quarterly*, **34**(136), 283–294.

McGrath, M. (2014). Propositions. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2014 edition.

Millikan, R. G. (1990). The myth of the essential indexical. *Noûs*, **24**(5), 723–734.

Moss, S. (2012). Updating as communication. *Philosophy and Phenomenological Research*, **85**(2), 225–248.

Ninan, D. (2010). *De se* attitudes: Ascription and communication. *Philosophy Compass*, **5**(7), 551–567.

Ninan, D. (2012). Counterfactual attitudes and multi-centered worlds. *Semantics and Pragmatics*, **5**(5), 1–57.

Ninan, D. (2013). Self-location and other-location. *Philosophy and Phenomenological Research*, **87**(2), 301–331.

Peacocke, C. (1981). Demonstrative thought and psychological explanation. *Synthese*, **49**(2), 187–217.

Perry, J. (1977). Frege on demonstratives. *Philosophical Review*, **86**(4), 474–497. Reprinted in Perry 1993, 3–32. Page references are to the 1993 reprint.

Perry, J. (1979). The problem of the essential indexical. *Noûs*, **13**(1), 3–21. Reprinted in Perry 1993, 33–52. Page references are to the 1993 reprint.

Perry, J. (1993). *The Problem of the Essential Indexical and Other Essays*. Oxford University Press, New York.

Pollock, J. (1990). Understanding the language of thought. *Philosophical Studies*, **58**(1-2), 95–120.

Putnam, H. (1973). Meaning and reference. *Journal of Philosophy*, **70**(19), 699–711.

Schiffer, S. (1978). The basis of reference. *Erkenntnis*, **13**, 171–206.

Spencer, C. (2007). Is there a problem of the essential indexical? In M. O'Rourke and C. Washington, editors, *Situating Semantics: Essays on the Philosophy of John Perry*. MIT Press, Cambridge.

Stalnaker, R. (1981). Indexical belief. *Synthese*, **49**(1), 129–149. Reprinted in Stalnaker 1999, 130-149. Page references are to the 1999 reprint.

Stalnaker, R. (1984). *Inquiry*. MIT Press, Cambridge, MA.

Stalnaker, R. (1999). *Context and Content*. Oxford University Press, New York.

Stalnaker, R. (2008). *Our Knowledge of the Internal World*. Oxford University Press, Oxford.

Stanley, J. (2011). *Know How*. Oxford University Press, Oxford.

Steward, H. (1997). *The Ontology of Mind: Events, Processes, and States*. Oxford University Press, Oxford.

Tiffany, E. (2000). What's essential about indexicals? *Philosophical Studies*, **100**(1), 35–50.

Torre, S. (2010). Centered assertion. *Philosophical Studies*, **150**(1), 97–114.

Williamson, T. (1994). *Vagueness*. Routledge, London.

Williamson, T. (2000). *Knowledge and Its Limits*. Oxford University Press, Oxford.