

Self-Location and Other-Location

Dilip Ninan
dilip.ninan@tufts.edu

1 Introduction

This paper concerns the proper treatment of self-locating and other-locating attitudes. I take as my starting point the *modal theory* of attitudes, the idea that the content of a belief, desire, etc. is best characterized by the set of possibilities at which the belief, desire, etc. is true. The two best known modal theories are the possible worlds account and the centered worlds account; the former has been developed most fully by Robert Stalnaker, the latter by David Lewis.¹ In more recent work, Stalnaker has come to accept a modified version of the centered worlds theory (see Stalnaker 2008, Ch. 3), and so I will focus my discussion on centered worlds, since the pure possible worlds theory now appears to have few advocates.

The motivation for the centered worlds theory has primarily to do with self-locating – or *de se* – attitudes. The focus of this paper is on the less-discussed question of how other-locating – or *de re* – attitudes ought to be treated within this framework. Most advocates of modal theories, including Stalnaker and Lewis, adopt a *descriptivist* treatment of other-locating attitudes. There are intramural differences between Stalnaker, Lewis, and other modal theorists (e.g. ‘two-dimensionalists’) on a number of issues: on the precise nature of descriptivism (global vs. local), how attitude content relates to the asserted content of the sentences we utter, and on the proper semantic treatment of attitude reports. I shall pass over these differences to focus on a problem common to these various approaches: all face a problem when it comes to characterizing the contents of *counterfactual attitudes* like imagining, dreaming, and wishing.

I then show how the problem can be solved if we adopt a different account of *attitudinal alternatives*, the possibilities used to characterize the content of an attitude. Rather than using centered worlds, I propose to use *multi-centered worlds*, or triples consisting of a possible world w , a time t , and a group of distinguished individuals in w , each of whom represents one of the individuals with whom the relevant agent is acquainted. I discuss two versions of this proposal: on the first, *de se* contents are what Perry (1979) calls “propositions of limited accessibility,” i.e. propositions that only a single agent can entertain; on the second, *de se* contents are ‘relativistic’ as they are on the centered worlds approach. I close with a tentative discussion of how we might try to decide between these two versions of the proposal.

¹See Stalnaker (1984) and the papers in Stalnaker (1999); and Lewis (1979, 1983a, 1986, 1995).

2 Self-location

Although my primary interest concerns other-locating attitudes, I begin with self-locating attitudes, since it is these that have been taken to motivate the centered worlds theory. There are a number of different arguments that Lewis and others have offered in favor of the centered worlds theory, but I will focus on one that concerns the role of *de se* attitudes in the explanation of action.²

Building on earlier work by John Perry, Lewis observes that two agents who share an appropriate set of *de se* beliefs and desires will, *ceteris paribus*, be disposed to behave in the same way. If, for example, you and I are both in a belief state that we could express by saying, “I am being attacked by a bear”, we will, other things being equal, be disposed to: curl up in ball or climb a tree or engage in other bear-averting actions. On Lewis’s view, the point of assigning content to attitudes is to characterize the causal role of attitudes vis-à-vis stimuli, behavior, and other attitudes. Since your belief that you will be attacked by a bear and my belief that I will be so attacked have the same causal role in our respective cognitive systems, they ought to be assigned the same content.

There is much more to say about this premise of Lewis’s argument – that since our respective beliefs ought to be assigned the same content – and I will return to this issue later in the paper (see §8).³ But if we *do* we accept Lewis’s premise for the moment, we will be forced to admit that contents have *relative truth values*: one and the same content can be true for one person, false for another. To see this, let p be the content that you and I both believe. Now suppose that I am in fact being attacked by a bear, but you are not (it’s merely a man in a bear suit). Then my belief is true and yours is false. But a belief is true (false) if and only the content of the belief is true (false). Since my belief is true, and the content of my belief is p , p must be true. Since your belief is false, and the content of your belief is p , p must be false. But p can’t be *both* true *and* false, so something has gone wrong.

Lewis’s solution is to give up the idea that p is true or false *simpliciter*; rather it is something that can be true relative to one agent, false relative to another. Since my belief is true, p is true for me; since yours is false, p is false for you. Problem avoided. But what kind of thing is p ? Lewis accepts the modal theory of attitudes, so p should be a set of possibilities. But it cannot be a set of *possible worlds*, for the following reason. A set of possible worlds (a possible worlds proposition) q is true at a world w just in case w is an element of q . You and I inhabit the same world, α . So p is either true at α or false at

²One can find something similar to this argument in Lewis’s discussion of the ‘Hume-Heimson’ case (Lewis 1979, 142-143). See also the introduction to Stalnaker (1999), Egan (2010), and Moss (2012). For other arguments for the centered worlds theory, see Lewis (1979), Lewis (1983a), and Ninan (2012).

³Perry (1979) denies that you and I believe the same content. He captures the psychological similarity between us by saying that we are in the same *belief state*; he thinks belief states in this sense can be characterized by centered (‘relativized’) propositions. It is not entirely clear whether this difference between Lewis and Perry is substantive or verbal; see Lewis (1979, 151-152). See Perry (2001) for Perry’s more recent view.

α . If it is true, then it cannot be the content of your belief, since your belief is false; if false, it cannot be the content of my belief, since my belief is true.

Lewis's alternative is to take p to be a set of *centered worlds*, where a centered world is a triple consisting of a world, a time, and an individual who exists at the time and world in question.⁴ This proposal requires some explanation.

The key notion of the centered worlds proposal is what it is for a centered world to be compatible with what an agent believes. In his papers on this topic, Lewis essentially treats this as a primitive notion, not one that is defined in independent terms. But although it is a primitive notion, we can say various things to convey its intuitive content. For example, we might say the following, for any agent x , time t and world w :

- (1) A centered world (w', t', x') is compatible with what x believes at t in w iff x' has, at t' in w' , all of the properties that x believes *de se* (at t in w) that she herself has.

Notions of compatibility for other attitudes (desire, imagination, etc.) can be given parallel characterizations.⁵

A *centered proposition* is a set of centered worlds, and an agent believes a centered proposition q just in case all of her centered belief worlds – all of her doxastic alternatives – are contained in q . If an agent x believes centered proposition q at time t in world w , her belief is true just in case (w, t, x) is in q .

Now let us apply all this to the bear attack case. I believe, at τ in α , that I am being chased by a bear. So every centered world (w, t, x) compatible with what I believe at τ in α is such that x is being chased by a bear at t in w . So I believe p , where:

$$p = \{(w, t, x) : x \text{ is being chased by a bear at } t \text{ in } w\}$$

Since I am being chased by a bear, $(\alpha, \tau, \text{me})$ is in p , and so my belief is true.

You also believe, at τ in α , that you are being chased by a bear. So every centered world (w, t, x) compatible with what you believe at τ in α is such that x is being chased by a bear at t in w . So you believe p as well. But since you are not being chased by a bear, $(\alpha, \tau, \text{you})$ is not in p , and so your belief is false. Thus, we believe the same thing – p – and that thing is true for me, false for you. So my belief is true, yours is false, as desired.

⁴Actually, since Lewis takes individuals to be world-bound, and believes in temporal parts, Lewis identifies p with a set of world-bound time-slices. Using centered worlds allows us to consider the same proposal within alternative metaphysical frameworks.

⁵Note that our gloss on the notion of compatibility employs the notion of a *de se* belief. As I understand the centered worlds theory, it can be seen as providing not a reduction of facts about intentional mental states, but a regimentation of those facts that reveals certain aspects of their structure. Reduction may come at a later stage, when we ask in virtue of what an agent should count as believing or desiring a centered proposition.

3 Other-location

How do we treat *other-locating* – or *de re* – attitudes on the centered worlds approach? We can bring out a potential problem for the approach by considering a well-known case due to Quine (1956):

There is a certain man in a brown hat whom Ralph has seen under questionable circumstances; suffice it to say that Ralph suspects he is a spy. There is also a gray-haired man who Ralph sees at the beach one day; Ralph recognizes this man to be Bernard J. Ortcutt, the town mayor, and Ralph believes that this man is no spy. Now Ralph does not know it, but the men are one and the same.

What does the centered worlds theorist say about this case and others like it: Pierre who fails to realize that London is Londres, the astronomer who has yet to discover that Hesperus is Phosphorus, etc.?

Let's say we are considering what Ralph believes at time t in his world w . Since Ralph believes (at t in w) that the man in the brown hat is a spy, we might be tempted to say that each of the centered worlds (w', t', x') compatible with what Ralph believes is such that this man is a spy at t' in w' . And since Ralph believes (at t in w) that Bernard J. Ortcutt is no spy, we might want to say that each of the centered worlds (w', t', x') compatible with what Ralph believes is such that Ortcutt is not a spy at t' in w' . But since the man in the brown hat just is Ortcutt, if we say both these things, then we must admit that no centered worlds are compatible with what Ralph believes at t in w , since no centered worlds (w', t', x') are such that that man/Ortcutt both is and is not a spy at t' in w' .

There are few different problems with being committed to this last point. One is that if no centered worlds are compatible with what Ralph believes, he counts as believing every centered proposition whatsoever. Another is that Ralph seems to have a coherent picture of what the world is like. Indeed, it seems that we have no trouble conceiving of what things are like from Ralph's perspective: there is one man he saw in a brown hat, and this man is a spy; there is another man, different from the first, who has gray hair and who Ralph saw on the beach, and this man is not a spy. The present approach to other-location doesn't seem to capture Ralph's intuitive picture of the world.

Note that this problem is not a problem for the *centered* worlds proposal *per se*: exactly the same problem would arise if we were using (uncentered) possible worlds to represent Ralph's doxastic alternatives instead. The general problem being raised here is for the idea that the content of a *de re* attitude about individual y is a *singular* centered or possible worlds proposition about y , i.e. the set of possible/centered worlds in which y is a spy.

Most advocates of the modal approach to attitudes accept some kind of descriptivist resolution to this problem. For example, the approach taken by Stalnaker is of essentially this kind – this is quite clear in Stalnaker (2008, Ch.

3).⁶ The same goes for *two-dimensionalists* like David Chalmers and Frank Jackson.⁷ But instead of trying to describe this descriptivist approach in very general terms, I will, for the sake of concreteness, focus on Lewis’s particular version of the descriptivist solution. But the problem I will later raise for Lewis’s theory should be easy to extend to the other versions of this general approach.

Lewis adopts a variation on the descriptivist treatment of the *de re* pioneered by Quine (1956) and Kaplan (1968). Lewis starts with the notion of a *relation of acquaintance*, a relation “of a sort apt for the reliable transmission of information” (Lewis 1979, 155). Note that Ralph bears two different relations of acquaintance to Ortcutt: on one occasion, he sees Ortcutt wearing a brown hat; on another occasion, he sees gray-haired Ortcutt on a beach. Let Q be the relation that x bears to y just in case y is the unique individual that x has seen wearing a brown hat. Let S be the relation that x bears to y just in case y is the unique gray-haired individual that x has seen on the beach. So Ralph bears both Q and S to Ortcutt.

On Lewis’s account, the basic notion of other-locating or *de re* belief is the notion of believing something about someone *relative to an acquaintance relation*.⁸ I take this notion to have reasonably clear intuitive content. Ralph is acquainted with Ortcutt in two different ways, but fails to realize this. He bears relation Q to Ortcutt, and relative to this relation, Ralph believes that Ortcutt is a spy. (We might suppose that, upon seeing the man in the brown hat, Ralph points and whispers, “That man is a spy.”) Ralph also bears relation S to Ortcutt, and relative to this relation, Ralph believes that Ortcutt is not a spy. (We might imagine Ralph nodding as we point to the gray-haired man on the beach and say, “That Bernard Ortcutt is no spy.”)

It will often be useful to put the “relative to” qualification next to the referring expression at issue; for example, I will often say things like: “Ralph believes that Ortcutt (relative to Q) is a spy, and believes that Ortcutt (relative to S) is not a spy”. This is somewhat inelegant, but serves the purposes of clarity.

Lewis then offers the following *analysis* of the intuitive notion of believing something about someone relative to an acquaintance relation:

- (2) An agent x believes, at t in w , that y is F , relative to acquaintance relation

⁶Stalnaker’s method of ‘diagonalization’ is not an alternative to the descriptivist approach; rather, it is a way of repairing a mismatch between the (‘locally’) descriptivist belief contents intuitively expressed or reported by certain utterances and the normal semantic content of the sentences uttered.

⁷See, for example, Jackson (1998) and Chalmers (2002). (Chalmers might resist being described as a ‘descriptivist’, but this subtlety is irrelevant here.)

⁸Lewis (1979, §XIII) talks about “ascribing a property to an individual under a description.” But he takes a “description” to be a relation, and in the case of *de re* belief, he requires the relation to be a relation of acquaintance. So this comes to the same thing as my ‘believing something about someone relative to an acquaintance relation’.

But this is, in any case, something of a simplification; something more general is needed for the ‘multiply *de re*’ case, as when Ralph believes that he is standing in between Jones and Ortcutt. The general notion is something like ‘believing something about a plurality of individuals-relative-to-acquaintance relations’.

R iff:

- (i) x bears R uniquely to y at t in w , and
- (ii) x believes *de se* (at t in w) that the thing to which he bears R is F .⁹

Given Lewis's account of *de se* belief, (ii) says that all the of centered worlds (w', t', x') compatible with what x believes at t in w are such that the thing to which x' bears R at t' in w' is F at t' in w' .

Note that, on this approach, there *will* be centered worlds compatible with Ralph believes in Quine's case. At time t in world w , Ralph believes that Orcutt (relative to Q) is a spy. On Lewis's account, this means that Ralph believes (at t in w) that the man he saw in the brown hat is a spy. So every centered world (w', t', x') compatible with what Ralph believes (at t in w) is such that the man x' saw in a brown hat at t' in w' is a spy at t' in w' .¹⁰ Ralph also believes (at t in w) that Orcutt (relative to S) is not a spy. So he believes (at t in w) that the gray-haired man he saw on the beach is not a spy. So every centered world (w', t', x') compatible with what Ralph believes (at t in w) is such that the gray-haired man that x' saw on the beach at t' in w' is not a spy.

But there is no conflict between these two claims about Ralph's doxastic alternatives. Each of the centered worlds (w', t', x') compatible with what Ralph believes at t in w contains two relevant individuals, both of whom correspond, in some sense, to Orcutt. One of them will be the person whom x' saw in a brown hat under questionable circumstances in w' , and this individual will be a spy in w' ; the other will be a gray-haired person that x' sees on the beach in w' , and this person will not be a spy in w' . Since the two men are distinct, there is no contradiction in saying that the one is a spy and the other is not.

Given this account of *de re* belief, an agent x believes *de se* at t in w that she is F just in case she believes *de re* that she is F relative to the relation I , identity, as Lewis (1979, 156-157) observes. This is because x believes, at t in w , that x is F relative to I iff:

- (i) x bears I to x at t in w , and
- (ii) all the centered worlds (w', t', x') compatible with what she believes (at t in w) are such that the thing to which x' bears I at t' in w' is F at t' in w' .

But since I is identity, (i) is trivially satisfied at any world and time at which x exists, and so is entailed by (ii); and (ii) is equivalent to the claim that all the centered worlds (w', t', x') compatible with what x believes (at t in w) are such that x' is F at t' in w' . Since the latter is just what it means, on the centered worlds account, to say that x believes *de se* at t in w that she is F , belief *de se* is just belief *de re* where the *res* is the agent and the relation of acquaintance is I , identity.

⁹This is Lewis's account of the singly *de re*. One would need something more general to handle the multiply *de re*.

¹⁰I am being a bit sloppy about the temporal aspects of Ralph's beliefs, but this will not affect any of the issues discussed in the paper.

4 Counterfactual attitudes

Up until now, our discussion has mostly focussed on the attitude of belief. But of course we want an account that applies to all the other content-bearing mental states: hopes, desires, episodes of imagining, dreams, and so on. But a problem arises when we try to extend Lewis's account of other-location to *counterfactual attitudes* like imagining, wishing, and dreaming.

Consider imagining. As with belief, we have an intuitive notion of imagining something about someone relative to an acquaintance relation. Suppose Ralph is watching the brown-hatted Ortcutt behaving suspiciously. Ralph can imagine something about the man before him, and intuitively this is different from his imagining something about the gray-haired man he saw on the beach. Imagining is fine-grained in the familiar way that virtually all psychological attitudes are. In the language we have adopted for speaking about these things, we can say that Ralph's imagining something about Ortcutt relative to Q (the 'see in a brown hat' relation) is different from his imagining something about Ortcutt relative to S (the 'see on the beach' relation). If, for example, Ralph imagines the brown-hatted man meeting the grey-haired man, we can say that he is imagining that Ortcutt (relative to Q) is meeting Ortcutt (relative to S).

The natural extension of Lewis's analysis of belief *de re* to imagination *de re* is this:

- (3) An agent x imagines, at t in w , that y is F , relative to acquaintance relation R iff:
 - (i) x bears R uniquely to y at t in w , and
 - (ii) x imagines *de se* (at t in w) that the thing to which he bears R is F .

Again, (ii) breaks down to this: every centered world (w', t', x') compatible with what x imagines at t in w is such the individual to whom x' bears R at t' in w' is F at t' in w' . Note that this means that the definite description in (ii) is understood to have a narrow-scope reading.

But there are two problems with (3), one for each direction of the biconditional. To see the problem with the left-to-right direction, let's suppose that Ralph is watching the brown-hatted Ortcutt behaving suspiciously. Notice that, of the various things Ralph might imagine about Ortcutt (relative to Q , the 'seeing in a brown hat' relation), some of them are compatible with Ralph's beliefs about Ortcutt (relative to Q), while some are not. For an example of the latter, note that although Ralph believes that Ortcutt (relative to Q) is a spy, Ralph might imagine a scenario in which Ortcutt (relative to Q) never took up espionage. That is, Ralph might point to the brown-hatted man he is watching and say, "I am imagining a scenario in which he is not a spy." Lewis's account has no trouble characterizing the content of this imagining; it is simply the set of centered worlds in which the individual to whom the center bears Q – the individual who the centers sees wearing a brown hat – is not a spy.

But the centered worlds account does run into a problem in cases in which what Ralph imagines about Ortcutt (relative to Q) conflicts with his bearing

relation Q to Ortcutt. For example, suppose Ralph is again looking at the brown-hatted man before him, and has an imagining he could report by saying, “I’m imagining a situation in which I never laid eyes on that man [Ralph points at the brown-hatted man before him]. In my imaginary scenario, that man is a recluse who few people have ever seen.” Now it is intuitively true that Ralph is imagining that Ortcutt (relative to Q) is a recluse whom he has never seen. But according to Lewis’s account, this is true if and only if two things hold (assume t and w are the time and world of the imagining):

- (i) Ralph bears Q uniquely to Ortcutt at t in w , and
- (ii) all the centered worlds (w', t', x') compatible with what Ralph imagines at t in w are such that there is an individual y' to whom x' bears Q at t' in w' , and y' is a recluse whom x' has never seen at t' in w' .

Claim (i) is true, since Ortcutt is the man Ralph has seen in a brown hat at t in w , and Q is the ‘see in a brown hat’ relation. But (ii) is false. To see this, first note that since Ralph’s imagining is intuitively coherent, there should be at least one centered world compatible with what he imagines, which means that (ii) is not vacuously true. Now recall that Q is the relation that x bears to y just in case y is the unique individual that x sees wearing a brown hat. And note that there is no centered world (w', t', x') in which there is an individual y' such that: (a) y' is the unique individual who x' sees wearing a brown hat at t' in w' , and (b) y' is a recluse whom x' has never seen at t' in w' . There is no such centered world because y' cannot both have been seen in a brown hat by x' and never have been seen by x' . Since there are no centered worlds that meet this condition, it cannot be that all the centered worlds compatible with what Ralph imagines at t in w meet this condition. Hence (ii) is false, and the relevant biconditional (the relevant instance of (3)) fails in the left-to-right direction.

I’ll return to consider some responses to this objection in a moment, but first I want to explain why I think (3) also fails in the right-to-left direction. Suppose Ralph is again watching Ortcutt behaving suspiciously in his brown hat. But now Ralph’s thoughts wander away from Ortcutt, and he begins to imagine a scenario in which he sees exactly one man in a brown hat, and in which this man is flying a kite. As I am thinking of this situation, Ralph’s imagining is not about Ortcutt or any other individual of his acquaintance (other than himself): he is simply imagining that he is watching *someone or other* wearing a brown hat and flying a kite. Other than the fact that his imaginary scenario contains himself, his imagining is purely general or *de dicto*. This seems clearly possible, and it seems that, if t and w are the time and world of his imagining, that it should *not* be true that Ralph is imagining at t in w that Ortcutt (relative to Q) is flying a kite. This should not be true, since, by stipulation, Ralph is not imagining anything about Ortcutt or any other individual of his acquaintance. But consider this biconditional, which I take to be an instance of (3):

Ralph imagines, at t in w , that he is watching Ortcutt (relative to Q) flying a kite if and only if:

- (i) Ralph bears Q uniquely to Ortcutt at t in w , and
- (ii) all the centered worlds (w', t', x') compatible with what Ralph imagines at t in w are such that the individual to whom x' bears Q at t' in w' is flying a kite at t' in w' .

As I said, it seems to me that the left-hand side of this biconditional is false. But the right-hand side is true: Ralph *does* bear Q to Ortcutt uniquely, and each of Ralph's centered imagination worlds *is* such that the individual to whom the center bears Q is flying a kite. Each of Ralph's centered imagination worlds contains exactly one individual who the center sees wearing a brown hat, and that individual is flying a kite. So Lewis's account predicts that Ralph should count as imagining that he is watching Ortcutt (relative to Q) flying a kite. But Ralph is plainly not doing this: his imagining is purely general and is not 'about' the brown-hatted man in front of him.¹¹

Two points about the scope of these objections. First, it is not just the mental state of imagining that is at issue; other *counterfactual attitudes* like wishing and dreaming give rise to a similar problem. For example: Ralph might wish that he had never seen Ortcutt (relative to Q). Second, although I have been focussing on the details of Lewis's theory, the objection would seem to apply quite generally to any modal theory of attitudes that employs some sort of descriptive treatment of the *de re*. I will have a little bit more to say about this below, when I discuss the possibility of a 'two-dimensionalist' solution to this problem.

5 Replies

5.1 First reply

One might reply that both of these apparent problems arise only because we are taking a too-simplistic view of the the acquaintance relation relative to which Ralph is thinking of Ortcutt. After all, it is not really plausible that Ralph has only ever seen one man wearing a brown hat. So the relevant acquaintance relation is not really Q , but something richer, something based on Ralph's perceptual experience of Ortcutt. Perhaps it is relation T , the relation x bears to y just in case y is the unique individual that looks ϕ , where ϕ is the way Ortcutt looks to Ralph on the occasion in question.

¹¹Some people have wondered whether some version of counterpart theory could be brought in to resolve these problems. This is a tricky question, and depends in part on what one means by 'counterpart theory'. The account I shall be proposing in the next section could be construed as a version of counterpart theory. But it is difficult to see how any account on which counterpart relations are *qualitative similarity relations* could work, given the problem just mentioned. For consider two imaginings Ralph might have: he might imagine that there is *some man or other* possessing certain qualitative characteristics; he might instead imagine that *Ortcutt* possesses those very same characteristics. It is hard to see how a qualitative-counterpart-theoretic account could distinguish these two cases. (A modal theory of attitudes that employs non-qualitative counterpart relations faces a different sort of problem; see Ninan (2012) for discussion.)

But this move won't solve the general problem. Take the left-to-right problem again. Ralph could imagine a scenario in which Ortcutt (relative to T) does not look ϕ , perhaps because of an unfortunate accident that he had as a child (he says, "I'm imagining that that man [he points at the brown-hatted man before him] looks very different from the way he in fact looks"). Or Ralph could imagine a scenario in which Ortcutt (relative to T) dies as a child; presumably the doomed child of Ralph's imagining looks little like the strapping man before him, and so again is unlikely to look ϕ .

In fact, I think there is a very general reason why this sort of strategy for avoiding the problem is not going to succeed. Whatever the 'right' acquaintance relation relative to which Ralph is thinking about Ortcutt is, the fact that Ralph stands in that relation to Ortcutt is likely to be a contingent fact. After all, what typically puts us in a position to think about an object is our standing in some contingent relation to that object. This makes it plausible that, no matter exactly what that relation is in the Ralph-Ortcutt case, Ralph will be able to imagine a scenario in which Ortcutt exists but in which he (Ralph) does not bear that relation to Ortcutt. So the problem cannot be avoided simply by making sure that we've selected the 'right' acquaintance relation.

5.2 Second reply

Let's put the account of *de re* imagining we are discussing before ourselves again:

- (3) An agent x imagines, at t in w , that y is F , relative to acquaintance relation R iff:
- (i) x bears R uniquely to y at t in w , and
 - (ii) x imagines *de se* (at t in w) that the thing to which he bears R is F .

I said that the definite description – "the thing to which he [the agent] bears R is F " – is read with narrow scope. That means that the content of Ralph's imagining that he never saw Ortcutt (relative to Q) is the set of centered worlds (w', t', x') in which x' never sees at t' in w' the individual he saw in a brown hat at t' in w' . But what if we read the same description with *wide scope* or (what amounts to the same thing) with a tacit "actually" between "he" and "saw"? I take it the proposal here is to replace (3) with (4), or to read (3) as if it expressed (4):

- (4) An agent x imagines, at t in w , that y is F , relative to acquaintance relation R iff:
- (i) x bears R uniquely to y at t in w , and
 - (ii) every centered world (w', t', x') compatible with what x imagines (at t in w) is such that the thing to which x bears R at t in w is F at t' in w'

What this means for the problem case is that Ralph imagines, at t in w , that he never saw Ortcutt (relative to Q) only if every centered world (w', t', x') compatible with what Ralph imagines at t in w is such that x' never sees, at t' in w' , the individual to whom Ralph bears Q at t in w . Now the individual to whom Ralph bears Q at t in w is just Ortcutt, so, on this proposal, Ralph imagines, at t in w , that he never saw Ortcutt (relative to Q) only if every centered world (w', t', x') compatible with what Ralph imagines at t in w is such that x' never sees Ortcutt at t' in w' . Since there are centered worlds in which the center fails to see Ortcutt, this move would appear to avoid our worry.

But this proposal really just takes us back to the problematic view that we discussed at the beginning of §3, the singular proposition approach. So the problem for that view is equally well going to be a problem for the view just proposed. Suppose, for example, that Ralph imagines a scenario in which Ortcutt (relative to Q) is distinct from Ortcutt (relative to S). For example, suppose that Ralph is watching the gray-haired man on the beach, while at the same time holding a photograph of the man in the brown hat that he earlier saw behaving suspiciously. Then Ralph might have an imagining he could report by saying, “I’m imagining a scenario in which I never saw that man [he points to the gray-haired man on the beach] nor this man [he points to the man in the photograph], and in which the two of them are having a fist-fight.” The present proposal cannot make sense of this imagining, an imagining which is evidently coherent.

Note that the ‘fist-fight’ case just mentioned is one that is problematic for *both* Lewis’s descriptive account of *de re* imagining *and* the singular proposition account discussed at the beginning of §3. This means that even if we adopt some kind of two-dimensionalism according to which attitudes have two kinds of content (descriptive and singular), we cannot evade the present problem.¹² For the above case – a case which combines the problem of counterfactual attitudes with the main problem for the singular proposition approach – will be one in which both of the contents posited by the two-dimensionalist will be empty.¹³

¹²For the sort of view I have in mind, see, for example, Chalmers (2002).

¹³An anonymous referee suggests that the centered worlds theorist might resist this last argument by maintaining that Ralph’s imagining in the ‘fist-fight’ case is not actually coherent, in spite of initial appearances. The referee suggests that the centered worlds theorist might describe the case as one in which Ralph is imagining a ‘disguisedly contradictory content’.

This is a possible response, but it is not one that Lewis and his followers are likely to find congenial. As I understand the proposal, centered propositions are not believed or imagined *simpliciter*, but only believed or imagined *under a disguise*. But if Lewis regarded this sort of move as legitimate, then it seems mysterious that he should have gone down the descriptive route in the first place: why not say, of Ralph in Quine’s original case, that he believes the singular centered proposition that Ortcutt is a spy under one disguise, and that he believes the singular centered proposition that Ortcutt is not a spy under a different disguise? Clearly, the referee’s suggested proposal abandons one of Lewis’s ambitions, which is to represent intuitive distinctions between contents using no more than sets of possibilities. Maybe that ambition ought ultimately to be abandoned, but that issue obviously exceeds the scope of the present paper.

5.3 Third reply

An anonymous referee makes the following observation:

As Lewis says in *On the Plurality of Worlds*, the connection between belief contents and belief reports is “complicated and multifarious” (Lewis 1986, 34). Lewis gives a number of rules linking beliefs and belief reports, but only one has been considered here. But Lewis is clear that not every report of the form “ x believes that y is F ” (let alone every report of the form “ x ϕ ’s that y is F ”) should be understood along these lines. From this perspective, the mere observation that this particular rule yields implausible results for some attributions involving imagination or desire is neither surprising nor especially problematic.

The first point worth noting about this is that I have not explicitly been discussing attitude *reports*; rather my topic has been self- and other-locating *attitudes*, a topic which, as Lewis (1979, 154) emphasizes, it is important to distinguish from the proper semantic treatment of attitude reports. Lewis’s own proposal is presented as an analysis of *de re* attitudes, not as analysis of *de re* attitude reports, and it is his analysis that I am challenging.

But in any case, it is worth examining the question of whether Lewis’s remarks on the connection between attitude reports and attitude contents provides him with a way out of the present difficulty. To that end, let us suppose that Ralph is looking at Ortcutt in the brown hat and says, “That man [pointing at Ortcutt] is a spy.” He then imagines that the man in question is a recluse who he (Ralph) has never seen. So the following sentence appears to be true in the relevant context:

- (5) Ralph imagined that he never saw that man [we point at the man Ralph is looking at].

Lewis (1986, 32-34) discusses five ways we might associate truth-conditions with an attitude report. The first of these does not obviously apply to (5), since it concerns *de dicto* reports, reports that do not contain any singular terms in the scope of the attitude verb. The second concerns *de se* reports whose only referring expression is “he himself”; it is difficult to draw any conclusions about sentences like (5) from what Lewis says about such sentences. The third possibility Lewis discusses is the acquaintance-based approach on which this essay has been focussing.

The fourth involves the ‘acceptance of sentences’:

Each of Peter’s doxastic alternatives is in a position to say truly, ‘Santa brings presents’; what is more, Peter and his alternatives more or less understand what this sentence means; and that is how Peter believes that Santa brings presents. (Lewis 1986, 33)

Adapting this to the case in which we are interested, we can suppose that if every centered world (w', t', x') compatible with what Ralph imagines at t in w is such that x' is in a position to say truly, ‘I never saw that man,’ then it is true to say of Ralph that he imagines that he never saw that man.

It may be that (5) is true if Ralph has imagining of the sort just described. But it is enough for our purposes to point out that (5) may also be true if Ralph imagines a scenario in which he is simply not in a position to refer demonstratively to anyone at all. To see this, note that the sentence is true if Ralph imagines a scenario in which the man in question is a recluse who he (Ralph) has never seen; and note that Ralph can imagine this without imagining that he is in a position to refer demonstratively to anyone.

Lewis’s fifth and final proposal also involves the acceptance of sentences: it explains how “Pierre believes that Father Christmas brings presents” can be true even though Pierre has never heard the name “Father Christmas”. The explanation involves the fact that each of Pierre’s doxastic alternatives accepts the sentence, “Père Noel brings presents”, and the fact that the two denotationless names emerge from a common tradition. This proposal appears to have little relevance to the present problem.

I conclude that Lewis’s alternative ‘recipes’ for associating truth-conditions with attitude reports do not help to resolve the problem of counterfactual attitudes.

6 Cheap haecceitism

On the centered worlds account of other-location, if Ralph has (at t in w) a *de re* attitude about Ortcutt (relative to Q), then each of Ralph’s attitudinal alternatives contains an individual y who represents Ortcutt (relative to Q) there. The problem of counterfactual attitudes seems to arise because of a certain constraint the theory puts on what this individual y can be like. The constraint is this:

a centered world (w', t', x') contains a representative of Ortcutt (relative to Q) only if there is an individual y such to whom x' bears Q uniquely at t' in w' , and that individual y is the one who represents Ortcutt (relative to Q) in that centered world.

There is, in effect, no other way for an individual in one of Ralph’s attitudinal alternatives to represent Ortcutt (relative to Q) as existing there. This leads straight to the problem of counterfactual attitudes, for if Ralph is to imagine (at t in w) anything about *that man* [we point at the brown-hatted man that Ralph is watching], he must imagine (at t in w) that he bears Q to that man.

The constraint requires that the individual who represents Ortcutt (relative to Q) in one of Ralph’s attitudinal alternatives be qualitatively similar to Ortcutt himself in a particular way. More precisely, it requires that *the pair* (x', y') that represents *the pair* (Ralph, Ortcutt) in (w', t', x') be qualitatively similar to (Ralph, Ortcutt) in a certain way: x' must bear Q to y' at t' in w' just as Ralph bears Q to Ortcutt at t in w .

It seems that this requirement of qualitative similarity is the source of our trouble. So we should be able to avoid the trouble if we can solve the problem of ‘*de re* representation’ without requiring the relevant representatives to be qualitative similar to what they represent in this way. To this end, it will be instructive to turn our attention briefly to a parallel discussion concerning the analysis of *metaphysical* modal claims.

On Lewis’s well-known account of possible worlds, the domains of any two worlds are disjoint. Given that doctrine, *de re* modal claims like “Romney could have won the election” cannot be understood as saying that there is a possible world in which Romney himself won the election. On Lewis’s original version of counterpart theory (Lewis 1968), the above sentence would instead be understood as saying that there is a possible world w which contains a *counterpart* of Romney who won the election in w . Romney’s counterparts in another world w are those things in w that more similar to Romney than are any other things in w , and which are sufficiently similar to Romney. Multiply *de re* claims are handled in a similar fashion: “It could have been that Romney won and Obama lost” is true just in case there is a possible world w which contains a counterpart of Romney who won in w and a counterpart of Obama who lost in w .

In an early response to Lewis, Fred Feldman argued that this theory has a number of questionable consequences. For example, Feldman (1971) suggests that the following sentence might be true, but would come out false on the counterpart-theoretic analysis:

I could have been more similar to the way you in fact are than to the way I in fact am; and at the same time, you could have been more similar to the way I in fact am than to the the way you in fact are.

Now there are probably a number of different ways for a defender of standard counterpart theory to respond to this objection, but I don’t want to consider them here. Instead, I want to look at a solution to this problem due to Allen Hazen (1979), a variant of which was adopted by Lewis himself (cf. Lewis 1983a, 1986, §4.4). What’s interesting for us is that Hazen’s proposal suggests a way of solving our problem of *de re* representation without entangling ourselves with the notion of qualitative similarity.¹⁴

On Hazen’s approach, possibilities – the things modal operators quantify over – are no longer identified with possible worlds. Instead, Hazen takes a possibility to be a pair of a possible world and what he calls a *representative function*, which is simply a function from the domain of the actual world into the domain of w . The idea is that f maps an individual x to his or her representative

¹⁴The more general interest of Hazen’s proposal concerns the manner in which it deals with certain logical difficulties that arise in connection with the ‘actuality’ operator. Fara and Williamson (2005) discuss these logical issues in more detail, and number of authors have recently responded to their paper by invoking theories similar to Hazen’s; see Dorr (2010), Bacon (Forthcoming), and Russell (Forthcoming). (Fara (2009), Kment (2012), and Stalnaker (2012) are also relevant to these issues.) The Hazen-style view is sometimes called *cheap haecceitism*.

in the possibility given by (w, f) . So on Hazen’s approach, “Romney could have won the election” is true just in case there is a possibility (w, f) such that $f(\text{Romney})$ wins the election in w . Similarly, “Romney could have won and Obama lost” is true just in case there is a possibility (w, f) such that $f(\text{Romney})$ wins in w and $f(\text{Obama})$ loses in w .

On Hazen’s version of this view, a pair (w, f) of this sort is a genuine possibility only if, for any individual x , f maps x to something in w that is *of the same kind* as x . Lewis (1983a, 1986, §4.4), on the other hand, thinks that, at least in some contexts, there are no such constraints, and so regards *every* pair (w, f) of this sort as a genuine possibility.¹⁵ But on either view, Feldman’s sentence should come out as true, since there are possibilities (w, f) such that the way $f(\text{me})$ is is more similar to the way you in fact are than to the way I in fact am, and such that, the way $f(\text{you})$ is is more similar to the way I in fact am than to the way you in fact are. The reason for this is that the individual who represents me in (w, f) is no longer the individual in w who most closely resembles me – it is simply $f(\text{me})$, whoever that may be. The individual who represents me in a possibility is no longer determined by relations of qualitative similarity; it is, so to speak, simply *stipulated* that $f(\text{me})$ represents me in (w, f) , irrespective of how qualitatively similar (or dissimilar) the two of us are. For this reason, Hazen calls these (w, f) pairs *stipulational worlds*.

7 Multi-centered worlds

Note a similarity here between stipulational worlds and centered worlds. When we are assessing a centered world (w', t', x') for compatibility with what an agent x believes or imagines (at time t in world w), the center x' represents the agent x at t' in w' . But x' represents x not in virtue of being identical to x nor in virtue of being qualitatively similar to x , but simply because this is how the theory works – the theorist simply stipulates that the center of a centered world is the agent’s representative in that centered world. So here again we see the idea that the individual that represents someone in another possible situation is determined via theoretical fiat.

But the similarity between stipulational worlds and centered worlds stops there: the centered worlds theorist adopts this approach only for the agent’s *de se* representative; all other representatives must be qualitatively similar to the individual they represent in the manner described earlier. Now if the problem of counterfactual attitudes arises from this requirement of qualitative similarity, then perhaps we can avoid the problem by treating *de re* representation for attitudinal possibilities the way Hazen treats *de re* representation for metaphysical possibilities.

¹⁵This is a bit misleading, since Lewis adopts a variation on Hazen’s proposal according to which modal claims are analyzed by quantifying over sequences of compossible individuals that are accessible from a given sequence of actual individuals. But he thinks that, at least in some contexts, there are no constraints on this accessibility relation, which is basically equivalent to the proposal I attribute to Lewis in the text. (On the difference between Hazen-style proposals and Lewis’s variant, see Dorr (2010, 16-17).)

But how do we do this? Well, perhaps we can simply identify attitudinal possibilities with Hazen’s stipulational worlds rather than with centered worlds. But there is at least one obstacle to doing this: any individual x in the actual world has at most one representative $f(x)$ in any given stipulational world (w, f) – this follows from f ’s being a *function*. But because of cases of identity confusion – cases like the Ralph-Ortcutt case – we need to allow for the possibility that an attitudinal possibility contains *more than one* representative of a given individual in the agent’s world. In any of Ralph’s belief worlds, there should be two distinct individuals, both of whom correspond to Ortcutt.

To appreciate this last point, think for a moment about how the centered worlds account handles the Ralph-Ortcutt case. In each of Ralph’s centered belief worlds (w', t', x') , there is one individual y' who x' has seen in a brown hat at t' in w' , and there is a distinct individual z' who x' has seen on the beach at t' in w' . Both of these individuals represent or correspond to Ortcutt, but they do so relative to two different ways in which Ralph is acquainted with Ortcutt. The first individual y' represents Ortcutt relative to Q (the ‘see in a brown hat’ relation), while the second z' represents Ortcutt relative to S (the ‘see on the beach’ relation). So on the centered worlds theory, an individual in an attitudinal possibility does not represent an individual in the agent’s world *full-stop*; rather an individual in an attitudinal possibility represents an individual in the agent’s world only *relative to an acquaintance relation*.

This observation suggests a way we might adapt Hazen’s proposal to the case of attitudinal modalities. Instead of using functions from *individuals* in the relevant world to individuals in some world (as Hazen does), we should instead use functions from *individual-acquaintance relation pairs* to individuals in some world. We could then identify attitudinal possibilities with pairs of a possible world w and a function from individual-acquaintance relation pairs to individuals in w . To see how this would work, suppose we gathered together all of the individual-acquaintance relation pairs (y, R) such that Ralph bears R to y at t in w . Let σ be the set of all of these pairs. The idea would be to represent Ralph’s attitudinal alternatives using pairs (w', f') consisting of a possible world w' and a function f' from σ into the domain of w' . The function f' would map each (y, R) in σ onto its representative in w' . So $f'(y, R)$ will be the individual in w' who represents y relative to R there. Since Ralph is acquainted with Ortcutt via Q and via S at t in w , both $(\text{Ortcutt}, Q)$ and $(\text{Ortcutt}, S)$ will be elements of our set σ . So in a possibility (w', f') – f' a function from σ into the domain of w' – $f'(\text{Ortcutt}, Q)$ will represent Ortcutt relative to Q , while $f'(\text{Ortcutt}, S)$ will represent Ortcutt relative to S . Since f' can map each of those pairs – $(\text{Ortcutt}, Q)$, $(\text{Ortcutt}, S)$ – to different individuals in w' , we can, like the centered worlds theorist, adequately represent Ralph’s confused belief state.¹⁶

Let us develop these ideas more precisely, and then illustrate how they help to solve the problem of counterfactual attitudes. We begin by defining the

¹⁶Note that nothing in what follows depends on Lewis’s idea that the domains of any two possible worlds are disjoint.

notion of an *acquaintance set*:

An agent x 's *acquaintance set* at time t in world w is the set of all pairs (y, R) such that x bears relation R to y uniquely at t in w .

An *acquaintance set simpliciter* is the acquaintance set of some agent at some time in some world.

In the example discussed above, σ is Ralph's acquaintance set at t in w .

To distinguish the functions employed in our theory from Hazen's representative functions, let us call the former *tagging functions*.

A *tagging function* is a function from an acquaintance set into the domain of some possible world.

Where Hazen took metaphysical possibilities to be pairs of a possible world and a representative function, we will instead take the attitudinal alternatives of an agent x at at time t in world w to be triples of a possible world w' , a time t' and a tagging function f' , where f' is a function from x 's acquaintance set at t in w into the domain of w' .¹⁷ Let us call this general type of object a *multi-centered world*:

A *multi-centered world* is a triple (w, t, f) consisting of a possible world w , a time t , and a tagging function f whose range is included in the domain of w .

'Multi-centered' is appropriate because this approach extends the centered worlds treatment of *de se* representation to the general *de re* case. Each of the outputs of the tagging function f is a distinguished individual – a 'center' – of the multi-centered world (w, t, f) .

A multi-centered world (w', t', f') will be compatible with what an agent x imagines at time t in world w only if the domain of f' is x 's acquaintance set at t in w , and the individuals who are the outputs of f' have all the properties that the agent imagines that the corresponding inputs have. More precisely:

(6) A multi-centered (w', t', f') is compatible with what an agent x imagines at at time t in a world w iff:

- the domain of f is $\{(y_1, R_1), \dots, (y_n, R_n)\}$, x 's acquaintance set at t in w ,
- each of the $f'(y_i, R_i)$ has, at t' in w' , all of the properties that x imagines, at t in w , that y_i has (relative to R_i),
- $f'(y_1, R_1), \dots, f'(y_n, R_n)$ stand, at t' in w' , in all of the relations that x imagines, at t in w , that y_1 (relative to R_1), ..., and y_n (relative to R_n) stand in.

¹⁷The time coordinate will not play a crucial role in what follows, but I add it for the sake of continuity with the centered worlds proposal.

Since any one-place property can be represented as an n -place relation, we can drop the middle clause on the right-hand side of the biconditional.¹⁸

As an alternative to the centered worlds account of *de re* imagining (cf. (4)), we offer the following:

- (7) Agent x imagines, at t in w , that y (relative to R) is F iff
 every multi-centered world (w', t', f') compatible with what x imagines
 (at t in w) is such that $f'(y, R)$ is F at t' in w' .¹⁹

Before illustrating how the account works – and how it resolves the problem of counterfactual attitudes – I should point out how *de se* attitudes are treated on this account. Here we follow Lewis (1979) in taking *de se* attitudes to be *de re* attitudes where the *res* is the agent and the acquaintance relation is I , identity. So if (w', t', f') is compatible with what agent x believes or imagines at t in w , $f'(x, I)$ is x 's '*de se* representative' in (w', t', f') .

We can now see how this account avoids the problem of counterfactual attitudes. Recall the original problem case:

Suppose Ralph is again looking at the brown-hatted man before him, and has an imagining (at t in w) he could report by saying, “I’m imagining a situation in which I never laid eyes on that man [Ralph points at the brown-hatted man before him]. In my imaginary scenario, that man is a recluse who few people have ever seen.”

Ralph is imagining (at t in w) that Ortcutt (relative to Q) is a recluse whom he has never seen. For Ralph to imagine this (at t in w) on our account is for all of the multi-centered worlds (w', t', f') compatible with what he imagines (at t in w) to be such that $f'(\text{Ortcutt}, Q)$ is a recluse at t' in w' who $f'(\text{Ralph}, I)$ has never seen at t' in w' . And there certainly are multi-centered worlds meeting this condition, since there are worlds w' and times t' such that some individual x' in w' is a recluse at t' in w' and some individual y' in w' has never seen x' at t' in w' . If (w', t', f') is such that $f'(\text{Ortcutt}, Q) = x'$ and $f'(\text{Ralph}, I) = y'$, then (w', t', f') is a multi-centered world meeting this condition.

That dispenses with the first problem involving counterfactual attitudes discussed earlier, the left-to-right problem. What about the second problem, the right-to-left problem? Recall the case: Ralph bears Q to Ortcutt at t in w ,

¹⁸Compare this account of compatibility to the one we gave for centered worlds in §2: for any agent x , time t and world w :

A centered world (w', t', x') is compatible with what x believes at t in w iff x' has, at t' in w' , all of the properties that x believes *de se* (at t in w) that she herself has.

¹⁹This is our account of the singly *de re*. The general account is this:

Agent x imagines, at t in w , that y_1 (relative to R_1), ..., y_n (relative to R_n) stand in relation T iff

every multi-centered world (w', t', f') compatible with what x imagines (at t in w) is such that $f'(y_1, R_1), \dots, f'(y_n, R_n)$ stand in relation T at t' in w' .

but engages in a purely general imagining – he simply imagines that there is a unique man in a brown hat who he sees, and that this man is flying a kite. Here his imagining is intuitively not about Ortcutt or anyone else. This case is again no problem for the multi-centered account. For Ralph to imagine this (at t in w) is for all the multi-centered worlds (w', t', f') compatible with he imagines at t in w to be such that there is exactly one man y' in a brown hat who f' (Ralph, I) is watching at t' in w' , and this man y' is flying a kite at t' in w' . In contrast, if Ralph were to imagine that he was watching *Ortcutt* (relative to Q) flying a kite, all of Ralph's imagination alternatives (w', t', f') would be such that f' (Ralph, I) was watching f' (Ortcutt, Q) flying a kite at t' in w' .²⁰

On the centered worlds account, if (w', t', x') is a centered world compatible with what Ralph imagines at t in w , then x' is the individual who represents Ralph (relative to I). If y' is the individual who represents Ortcutt (relative to Q) in (w', t', x') , the account demands that x' bears Q to y' , just as Ralph bears Q to Ortcutt. I have suggested that this requirement of qualitative similarity between (x', y') and (Ralph, Ortcutt) is what leads to the problem of counterfactual attitudes. Note how the multi-centered worlds account differs here. Suppose (w', t', f') is a multi-centered world compatible with what Ralph imagines at t in w , and that f' (Ralph, I) = x' and f' (Ortcutt, Q) = y' . Then x' represents Ralph (relative to I) and y' represents Ortcutt (relative to Q). But on this account, there is simply no constraint as to how (x', y') and (Ralph, Ortcutt) are related – there is simply no demand that those two pairs be qualitatively similar in any particular respect. It is this feature of the present account that enables it to avoid the problem of counterfactual attitudes.

8 Recovering relativism

Although I have been talking only about imagination in the last few sections, I assume that we want a uniform account of attitudes and their content. So the foregoing all applies, *mutatis mutandis*, to attitudes other than imagining, such as belief. The account of doxastic compatibility for multi-centered worlds is easily read off of our account of ‘imagination compatibility’: just take (6) and substitute the word *believes* in for every occurrence of the word *imagines*.

And so far we have talked about the characterization of attitudinal alternatives on this approach, but we have said nothing about what the *content* of an attitude is. Usually in a modal theory, the content of an attitude is a set of possibilities, e.g. the centered worlds theorist takes contents to be sets of centered worlds. The natural thing to do on the multi-centered approach, then, is to take contents to be sets of multi-centered worlds, aka *multi-centered propositions*. To believe or imagine a multi-centered proposition p is for all of the multi-centered worlds compatible with what one believes or imagines to be contained in p .

But one consequence of moving from centered worlds to multi-centered worlds is that we lose the relativistic individuation of *de se* content that we discussed in

²⁰The reader can verify that the present account has no trouble handling the ‘fist-fight’ case discussed in §5.

§2. Recall the argument we discussed in favor of relative contents: Two agents, a and b , each of whom is in a belief state (at time t in world w) that she could express by saying, “I am being attacked by a bear” will, other things being equal, be disposed to act in similar ways. So the content of their beliefs should be the same (according to Lewis). But since a ’s belief might be false, while b ’s is true, the thing they believe needs to be something that can vary truth value across a and b . Lewis’s centered propositions have just this feature, and this was taken to be an argument for Lewis’s approach.

But consider what the multi-centered content of a ’s belief would be, if a believed (at t in w) that she was being chased by a bear. Individual a has a *de se* belief, which we construe as a *de re* belief about herself relative to I , the identity relation. So in each multi-centered world (w', t', f') compatible with what she believes (at t in w), her representative – $f'(a, I)$ – is being chased by a bear at t' in w' . So the content of her belief is the following set of multi-centered worlds:

$$(8) \{(w', t', f') : f'(a, I) \text{ is being chased by a bear at } t' \text{ in } w'\}$$

But the content of b ’s belief, on the other hand, is going to be something different. She has a *de re* belief about *herself*, relative to I . So in each multi-centered world (w', t', f') compatible with what she believes (at t in w), *her* representative – $f'(b, I)$ – is being chased by a bear at t' in w' . So the content of her belief is:

$$(9) \{(w', t', f') : f'(b, I) \text{ is being chased by a bear at } t' \text{ in } w'\}$$

But since a is distinct from b , (8) and (9) are different sets. So a and b do not believe the same thing, despite the doxastic similarity between them.

In fact, it is not hard to see that, on the multi-centered approach as developed thus far, *de se* contents are what John Perry has called “propositions of limited accessibility” (Perry 1979, 45ff.). For (8) is a multi-centered proposition that *only a can believe*, and (9) is a multi-centered proposition that *only b can believe*. To see this, let us suppose that an agent x believes (8) at a time t in a world w . We want to show that x must be identical to a . Since x believes (8) at t in w , it follows that every multi-centered world (w', t', f') compatible with what x believes at t in w is in (8). Note that (w', t', f') is in (8) only if f' is defined for the pair (a, I) . Since (w', t', f') is compatible with what x believes at t in w , the domain of f' must be x ’s acquaintance set at t in w . From these two facts it follows that (a, I) is in x ’s acquaintance set at t in w , and from this it follows that x bears I to a at t in w . Since I is identity, this means that x is identical to a . So if an agent believes (8), that agent is a ; so only a can believe (8). Parallel reasoning shows that only b can believe (9).

Is this a bad result? Lewis would say so, but the issue is controversial. Frege (1918/1956) famously held the view that the content of an agent’s *de se* attitude was something that could only be believed by the agent herself, a view defended by some contemporary philosophers, such as Evans (1981). This is an interesting result: generalizing on the centered worlds theory in a certain way

yields a theory of content that is broadly similar to the neo-Fregean approach advocated by Evans.²¹

This feature of the multi-centered theory also brings out two different aspects of the original *centered* worlds theory. On the one hand, there is the relativistic aspect of that theory, which we have just now been discussing, and which we discussed in §2. On the other hand, there is the way the centered worlds theory handles the issue of ‘*de se* representation,’ the question of who, in an attitudinal alternative, is the agent’s *de se* representative there. The centered worlds theorist responds to this question by stipulating the answer: in a centered world (w', t', x') that we are assessing for compatibility with what the agent x believes (or imagines), x' represents x . This aspect of the centered worlds theory was emphasized in §7. What we see here is that a theory can have this second feature without having the first. So these two features of the centered worlds theory are not as tightly connected as they might have seemed at first.

In any case, Lewis and his followers will not be happy with these ‘interesting results’, since we are no longer able to say that a and b believe the same thing, in spite of the doxastic similarity between them. I will return to the question of whether this really is what we ought to say about the bear attack case (and cases like it). But before I do that, I want to point out that there is an alternative way of formulating the multi-centered approach which Lewis and his followers should find more congenial. For on this alternative formulation, *de se* contents are relativistic in just the way they are on the centered worlds approach. After demonstrating this, I will discuss which version of the approach (if either) ought to be preferred.

In our development of the multi-centered theory, we have been taking the ‘*de re* representation’ relation to be a relation between individuals in a possibility and individual-acquaintance relation pairs (y, R) , where the agent bears R *uniquely* to y . Let me emphasize that last point: all of the acquaintance relations we’ve been talking are ones that the agent bears *uniquely* to some individual (at the time and world in question). For example, we have been assuming that an individual-acquaintance relation pair (y, R) is in an agent x ’s acquaintance set at time t in world w only if x bears R uniquely to y at t in w .

Now suppose we have a multi-centered world (w', t', f') , where the domain of f' is some agent x ’s acquaintance set at t in w . And suppose we do the following: we take each element (y, R) of x ’s acquaintance set and ‘throw out’ the individual component y of the pair. What we’re left with then is a *set of acquaintance relations*: that set of acquaintance relations R such that x bears R uniquely to someone at t in w . We can make a parallel adjustment to f' : let the domain of f' be the set of acquaintance relations $\{R_1, \dots, R_n\}$ obtained in this manner. I submit to you that, despite having been transformed in this way, the triple (w', t', f') can still be regarded as representing a potential attitudinal

²¹There are a number of interesting similarities between (this version of) the multi-centered theory and Evans’s account. For example, the contents of *de re* attitudes are ‘object-dependent’ propositions on both accounts, and both accounts make important use of individual-acquaintance relation pairs in order to resolve cases of identity confusion.

alternative for x at t in w . To see this, let “ $R^{w,t,x}$ ” abbreviate “the individual to whom x bears R uniquely at t in w ”. Then we can regard each of the $f'(R_i)$ as representing $R_i^{w,t,x}$ relative to R_i .

To put it loosely: since we can always ‘recover’ an individual from an acquaintance relation (together with the agent, time, and world), we can throw the individuals out of the agent’s acquaintance set, and everything will work just as before. The main difference between a theory formulated along these lines and the theory developed in the previous section of the paper concerns the individuation of content. In particular, the version of the theory we are developing yields a relativistic individuation of *de se* content, as we shall soon see.

Instead of taking an acquaintance set to be a set of individual-relation pairs, we will now take an acquaintance set to be a set of acquaintance relations:

An agent x ’s *acquaintance set* at t in w is the set of all acquaintance relations R such that there is someone to whom x bears R at t in w .

We leave the other definitions as they were: a tagging function is still a function from acquaintance sets to individuals in the domain of some world, and a multi-centered world is a triple (w, t, f) , where f is a tagging function. But since acquaintance sets are now sets of acquaintance relations, tagging functions are now functions from acquaintance relations to individuals.

Compatibility is then characterized as follows (I use the state of belief here, rather than the state of imagining):

(10) A multi-centered (w', t', f') is compatible with what an agent x believes at at time t in a world w iff:

- the domain of f is $\{R_1, \dots, R_n\}$, x ’s acquaintance set at t in w ,
- $f'(R_1), \dots, f'(R_n)$ stand, at t' in w' , in all of the relations that x believes, at t in w , that $R_1^{w,t,x}$ (relative to R_1), ..., $R_n^{w,t,x}$ (relative to R_n) stand in.

The account of *de re* belief on this version of the proposal goes like this:

(11) Agent x believes, at t in w , that y (relative to R) is F iff

- (i) $R^{w,t,x} = y$,
- (ii) every multi-centered world (w', t', f') compatible with what x believes (at t in w) is such that $f'(R)$ is F at t' in w' .

It is straightforward to verify that this version of the multi-centered approach resolves the problem of counterfactual attitudes in much the same manner as the first version did

The content of an attitude is still a set of multi-centered worlds, a multi-centered proposition. But now with our new understanding of what acquaintance sets, tagging functions, and multi-centered worlds are, the content of a *de se* attitude will be relativistic in the way centered propositions are. To see

this, suppose again that, at t in w , a believes *de se* that she is being chased by a bear, and b believes *de se* that she is being chased by a bear. Then every multi-centered world (w', t', f') compatible with what a believes (at t in w) will be such that $f'(I)$ is being chased by a bear at t' in w' . So, at t in w , a believes the following multi-centered proposition:

$$(12) \{(w', t', f') : f'(I) \text{ is being chased by a bear at } t' \text{ in } w'\}$$

But everything is exactly the same for b : every multi-centered world (w', t', f') compatible with what *she* believes (at t in w) will be such that $f'(I)$ is being chased by a bear at t' in w' , and so b too will count as believing (12).

But in spite of believing the same thing, a 's belief might be true, while b 's is false. To see how the account allows for this possibility, we need to say something about what it is for a belief to be true on this account:

$$\text{Let } x\text{'s acquaintance set at } t \text{ in } w \text{ be } \{R_1, \dots, R_n\}, \text{ and let } f^{w,t,x} = \{(R_1, R_1^{w,t,x}), \dots, (R_n, R_n^{w,t,x})\}$$

If x believes a multi-centered proposition p at t in w , then x 's belief is true iff $(w, t, f^{w,t,x}) \in p$.

Suppose a is being chased by a bear at t in w , but that b is not. And suppose that a 's acquaintance set at t in w is $\{I, R_2, \dots, R_n\}$. Then, by the above account, her belief is true iff:

$$(w, t, f^{w,t,a}) \in (12)$$

where $f^{w,t,a} = \{(I, I^{w,t,a}), (R_2, R_2^{w,t,a}), \dots, (R_n, R_n^{w,t,a})\}$. So a 's belief is true iff $f^{w,t,a}(I)$ is being chased by a bear at t in w . Note that $f^{w,t,a}(I) = I^{w,t,a}$, where $I^{w,t,a}$ is the individual to whom a bears I at t in w . Since I is identity, $I^{w,t,a} = a$. So a 's belief is true iff $f^{w,t,a}(I) = a$ is being chased by a bear at t in w . Since a *is* being chased by a bear at t in w , her belief is true. Parallel reasoning shows that, when b believes (12), her belief is true iff b is being chased by a bear at t in w ; since we are supposing that b is *not* being chased by a bear, her belief is false. So a and b believe the same thing – viz. (12) – even though a 's belief is true, while b 's is false. This was the result desired by Lewis and his followers.

So we now have two versions of the multi-centered account: one which, like Fregean accounts, takes *de se* contents to be propositions of limited accessibility; and one which, like the centered worlds account, takes them to be relative propositions, things that vary in truth value between individuals. Which account should we adopt? I will not try to settle this issue conclusively here, but I will try to give the reader a sense of the relevant considerations.

The proponent of relativistic content can, of course, point to the ‘bear attack’ argument in defense of her position. The first premise in that argument is that two agents who are in the same *de se* belief state – for example, two agents who are both in a belief state that they could express by saying, “I am being attacked by a bear” – are disposed to behave in similar ways (other things being

equal). One could object to this claim, but I don't want to get into that here. The second premise is that, in virtue of the fact reported by the first premise, we ought to say that the two agents believe the same thing. It is this second premise that I wish to examine in more detail.

Let us agree that the role *de se* attitudes play in the explanation of action requires us to accept that agents in similar *de se* states will, other things being equal, be disposed to act in the same way. *One* way to capture this similarity would be to say that they believe the same thing. But another way to capture it would be to say that, although such agents don't believe the same thing, they share a common property, the property of being in the same *de se* belief state. It is this property that is relevant for explaining why they are disposed to behave similarly. For example, someone who wanted to defend the 'limited accessibility' version of the multi-centered approach could point out that, in our bear attack example, both *a* and *b* have the property of being an *x* such that *x* believes:

$$\{(w', t', f') : f'(x, I) \text{ is being chased by a bear at } t' \text{ in } w'\}$$

The fact that they share this property might then be relevant for explaining why they are disposed to behave similarly, perhaps in conjunction with other properties of this sort that they share (e.g. the property of believing that one can avert a bear attack by playing dead, the property of wanting to avert a bear attack, etc.).

That seems like a good reply, but there remains a question about the 'limited accessibility' approach. According to that approach there is a doxastic difference between *a* and *b*: there must be, since they believe different things. But, the relativist asks, why should we believe that there really is any doxastic difference here? She might continue as follows:

“On my account, the central constraint on the individuation of attitude content concerns the role that attitudes play in the explanation of behavior. So if two agents, in virtue of being in the same *de se* state are thereby disposed to behave in similar ways, we ought to say they believe the same thing. If you say they believe different things, what difference between the agents are you tracking? Not one relevant to the explanation of action. So the difference you posit in what these agents believe is a difference that makes no difference.”

Of course, the defender of the 'limited accessibility' approach does have a possible answer to this charge: she can say that she is seeking to capture the fact that the two agents' respective beliefs might differ in truth value, as is the case in our bear attack example. That is a potential difference in their belief states, one we might well be interested in, and the sort of thing that might be taken to constrain a theory of attitude content.

Let me close by pointing out one further difference between these two versions of the multi-centered approach, a difference which may again be relevant to the choice between them. The first version of the multi-centered approach is

externalist in a way in which the second view is not. To illustrate this, consider two individuals, Jones and Smith, on opposite ends of the globe. Jones bears some acquaintance relation R uniquely to a particular individual – Jane, say. Smith finds himself in an epistemic situation which, ‘from the inside’, is exactly like the situation Jones finds himself in. And he too bears R uniquely to some individual – Susan (where Susan is distinct from Jane). Now suppose that Jones comes to believe that Jane (relative to R) has brown hair; Smith similarly comes to believe that Susan (relative to R) has brown hair.

According to the first version of the multi-centered approach, the content of Jones’s belief is (13), while the content of Smith’s is (14):

(13) $\{(w', t', f' : f'(Jane, R) \text{ has brown hair at } t' \text{ in } w')\}$

(14) $\{(w', t', f' : f'(Susan, R) \text{ has brown hair at } t' \text{ in } w')\}$

Since Susan is distinct from Jane, these are distinct contents, so what Jones believes differs from what Smith believes. But on the second version of the multi-centered approach, the two men believe the same thing:

(15) $\{(w', t', f' : f'(R) \text{ has brown hair at } t' \text{ in } w')\}$

In this respect, the second version of the account is again more similar to the centered worlds approach, which would also say that Jones and Smith believe the same thing.

References

- Bacon, Andrew. Forthcoming. “Representing Counterparts.” *Australasian Journal of Logic* .
- Chalmers, David. 2002. “The Components of Content.” In David Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press.
- Dorr, Cian. 2010. “How to be a Modal Realist.” Unpublished manuscript, Oxford University.
- Egan, Andy. 2010. “Disputing about Taste.” In Ted Warfield and Richard Feldman (eds.), *Disagreement*. Oxford: Oxford University Press.
- Evans, Gareth. 1981. “Understanding Demonstratives.” In H. Parret and Jacques Bouveresse (eds.), *Meaning and Understanding*. Berlin: Walter de Gruyter. Reprinted in Evans 1985, 291-321. Page references are to the 1985 reprint.
- . 1985. *Collected Papers*. Oxford: Clarendon Press.
- Fara, Delia Graff. 2009. “Dear Haecceitism.” *Erkenntnis* 70:285–297. doi:10.1007/s10670-008-9149-3.
- Fara, Michael and Williamson, Timothy. 2005. “Counterparts and Actuality.” *Mind* 114:1–30. doi:10.1093/mind/fzi001.

- Feldman, Fred. 1971. "Counterparts." *Journal of Philosophy* 68:406–409. doi:10.2307/2025039.
- Frege, Gottlob. 1918/1956. "The Thought: A Logical Inquiry." *Mind* 65:287–311. doi:10.1093/mind/65.1.289.
- Hazen, Allen. 1979. "Counterpart-Theoretic Semantics for Modal Logic." *The Journal of Philosophy* 76:319–338. doi:10.2307/2025472.
- Jackson, Frank. 1998. *From Metaphysics to Ethics*. Oxford: Clarendon Press. doi:10.1093/0198250614.001.0001.
- Kaplan, David. 1968. "Quantifying In." *Synthese* 19:178–214. doi:10.1007/BF00568057.
- Kment, Boris. 2012. "Individuals, Supervenience, and Fundamentality." *Philosophical Review* 121:573–609. doi:10.1215/00318108-1630930.
- Lewis, David K. 1968. "Counterpart Theory and Quantified Modal Logic." In Lewis (1983b), 113 – 126. doi:10.1093/0195032047.001.0001.
- . 1979. "Attitudes *De Dicto* and *De Se*." *Philosophical Review* 88:513–543. doi:10.1093/0195032047.003.0010. Reprinted in Lewis 1983b, 133-159. Page references are to the 1983 reprint.
- . 1983a. "Individuation by Acquaintance and by Stipulation." *Philosophical Review* 92:3–32. doi:10.2307/2184519. Reprinted in Lewis 1999, 373-402. Page references are to the 1999 reprint.
- . 1983b. *Philosophical Papers, Volume 1*. New York: Oxford University Press. doi:10.1093/0195032047.001.0001.
- . 1986. *On the Plurality of Worlds*. Oxford: Blackwell.
- . 1995. "Reduction of Mind." In Samuel Guttenplan (ed.), *A Companion to Philosophy of Mind*, 291–324. Oxford: Basil Blackwell.
- . 1999. *Papers in Epistemology and Metaphysics*. Cambridge: Cambridge University Press.
- Moss, Sarah. 2012. "Updating as Communication." *Philosophy and Phenomenological Research* 85:225–248. doi:0.1111/j.1933-1592.2011.00572.x.
- Ninan, Dilip. 2012. "Counterfactual Attitudes and Multi-Centered Worlds." *Semantics and Pragmatics* 5:1–57. doi:10.3765/sp.5.5.
- Perry, John. 1979. "The Problem of the Essential Indexical." *Noûs* 13:3–21. doi:10.2307/2214792. Reprinted in Perry 1993, 33-52. Page references are to the 1993 reprint.
- . 1993. *The Problem of the Essential Indexical and Other Essays*. New York: Oxford University Press.
- . 2001. *Reference and Reflexivity*. Stanford: CSLI Publications.
- Quine, W.V.O. 1956. "Quantifiers and Propositional Attitudes." *Journal of Philosophy* 53:177–187. doi:10.2307/2022451. Reprinted in Quine 1976, 185-196.
- . 1976. *The Ways of Paradox and Other Essays*. Cambridge, MA: Harvard University Press.

Russell, Jeff Sanford. Forthcoming. "Actuality for Counterpart Theorists." *Mind*

Stalnaker, Robert. 1984. *Inquiry*. Cambridge, MA: MIT Press.

—. 1999. *Context and Content*. New York: Oxford University Press. doi: 10.1093/0198237073.001.0001.

—. 2008. *Our Knowledge of the Internal World*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199545995.001.0001.

—. 2012. *Mere Possibilities*. Princeton, NJ: Princeton University Press.