

Indexicality and A Prioricity

James Pryor

Dept of Philosophy

Princeton University

<jimpryor@princeton.edu>

July 30, 2004

1 Introduction

The notion of “de re belief” has been used in a variety of ways. On the surface, it looks like all theorists are working with roughly the same idea, but on critical inspection I think that unity proves illusory.

I want to take a few minutes to sketch what I think are the separate dominant ways of understanding this notion, and which understanding I’ll be working with. We can structure our discussion around three questions.

Q1. What is the phenomenon to explain?

On the one hand, we have a certain style of **thought-attribution** that we call de re. This style of attribution is extensional and supports exportation.

On the other hand, some philosophers think there’s a **distinctive kind of mental state** that we can call de re.

Some philosophers acknowledge only the phenomenon of de re attributions. If they talk about “de re thoughts,” they’ll say these are only ordinary

thoughts that satisfy certain relational conditions. There's nothing intrinsically distinctive to thoughts of this kind. (Chisholm, Sosa, Lewis, Kaplan in "Quantifying In")

More commonly, philosophers acknowledge both phenomena: they agree that de re thoughts constitute a distinctive kind of mental state, and they say that *that's* the kind of mental state that de re attributions are in the business of attributing.

My own view is that both phenomena exist, but *it's unclear* how exactly they're connected. It's unclear whether de re attributions are always in the business of attributing de re thoughts. We'll talk more about this as we proceed.

Q2. What *is it* to have a de re thought?

One answer is that de re thoughts are just thoughts that have a certain type of content: e.g., object-dependent or singular propositions. Now even if it's *true* that de re thoughts have the proposed type of content, this won't be a satisfying answer to our question. We'll still want to know *why* a subject should count as having a thought with that type of content, as opposed to another type. What, in independent terms, will having a thought with that type of content *amount to*?

A more substantial answer to our question gives methodological priority to **de re attributions**. We aim first to give a semantics that's adequate to our existing attributional practices; and then we count you as having a de re thought just in case a de re attribution is true of you. On this approach, our notion of de re thoughts tracks whatever our willingness to make and accept de re attributions tracks. (This seems to be how most direct-reference theorists understand the notion of de re thought.)

Robin Jeshion answers our question differently. She says you count as thinking about an object de re when you think about that object in a "name-like" way. This involves having a certain kind of **underlying cognitive**

architecture. You need to have a “mental dossier” for the object; the thought in question has to be drawing on that dossier; and so on. This way of understanding what de re thought amounts to leaves it unsettled what the connection is between de re attributions and de re thoughts.

My answer to our question is different still. I distinguish de re thoughts from other thoughts by **what grounds your capacity to have** the thought. When you think that even the shortest spy is so-and-so, you may count in one sense (a thin sense) to be “thinking about” the shortest spy. Your capacity to be thinking about the shortest spy in *that* sense is grounded in his satisfying the descriptive condition you’re thinking. That descriptive condition determines which object you’re “thinking about.” When you demonstrate a man you see and think *he* is so-and-so, on the other hand, your capacity to think that thought is grounded in your acquiring perceptual information about the man. I understand de re thoughts to be thoughts that your capacity to have is grounded in informational relations of this sort.

Now, what is “information” and what are these “informational relations”? I find it most helpful to think of information in terms of its relation to justification. Information is what, in the right conditions, *constitutes* justification. For example, an experience as of something’s being F is one type of information that something is F. In creatures with the right kind of cognitive sophistication, and who meet certain enabling conditions, that experience *will constitute* prima facie justification to believe that something is F. It’s the sort of thing on which you could *base* a belief that something is F. It’s controversial exactly what kind of cognitive sophistication, and what enabling conditions, are needed. For example, do your experiences need to be reliable? Do you need to have background evidence that they are? But we don’t need to get into that. I understand information to be whatever *constitutes* your prima facie justification, *when you’re lucky enough to have* prima facie justification. Unlucky creatures will have the information but not the justification.

That's how I understand information *in general*. I understand *information about α* to be whatever constitutes your prima facie justification to believe *de re thoughts about α* . That's not supposed to be a reductive definition. (It can't be; I'm using the notion of information to explain what a de re thought is.) It's just helping us get a better handle on my notion of information. Information about α 's being F is just information about Fness, together with whatever would turn prima facie justification about Fness into prima facie justification about α 's being F. For example, we said that *a perceptual experience* as of something's being F, in lucky creatures, will constitute prima facie justification to believe that something is F. When that experience also meets *certain relational conditions*—e.g., if by having it you're seeing α to be F—then having it will also, in lucky creatures, constitute prima facie justification to believe, of α , that *it* is F. (Just as unlucky creatures could have the information about Fness without its constituting justification for them, so too can they have information about α 's being F without *that* constituting justification for them.)

So it's a constraint on my notion of information that in the right conditions it can constitute these different kinds of justification.

I approach the notion of information in this round-about way because I have a much better pre-theoretical understanding of what justification is, and solid intuitions about when we do and when we don't have certain kinds of justification. If I'm right to understand the relation between information and justification in this way, then I can leverage what I know about justification to help me better understand the notion of information.

Q3. When are you *able to have* a de re thought? When is it felicitous to make de re attributions to you?

Latitudinarians say it's relatively easy to have de re thoughts, or to have de re attributions be true of you. For example, if you think that the F is so-and-so, and know that there is a unique F, and α happens to be the unique

F, then for some philosophers that'll be enough for you to count as thinking, of the F, that it is so-and-so. Some Latitudinarians add extra constraints. For example, you may need to have introduced a directly referential term that you stipulated to refer to whomever is the F (a term like Kaplan's 'Newman-1').

Acquaintance Theorists say it's much harder to have de re thoughts about an object. You need to stand in some special epistemic connection to the object. Merely knowing that there's a unique F, and stipulating that a term will directly refer to that person, whomever he is, won't be enough.

"Acquaintance" here is a technical term for *whatever* special epistemic connection is required for you to think de re thoughts, if there is such a requirement. The term is also sometimes used in other contexts: e.g., as another name for what epistemologists call "direct apprehension"; and sometimes to describe a kind of introspective awareness that only you can have of your own pains. We shouldn't assume that these different relations coincide.

Explaining the present notion of acquaintance raises several difficulties. First, it's difficult to spell out what the special epistemic connection should be in a general way. We'd like to know what's in common to perceptual acquaintance, acquaintance through testimony, and so on. Can we be acquainted with abstract objects like numbers and fictions? Why are these various epistemic connections all instances of a single genus?

A second difficulty is how to understand cases of hallucination, miscommunication, and so on, where it *feels as though* you have a de re thought, but there is no appropriate real object for you to be thinking about. Do you have a de re thought of a different sort? (A thought with a gappy content? A thought about a creature of your imagination?) If so, is there still an acquaintance requirement in place? What do you need to be acquainted with?

These are difficulties of detail. It's good to be aware of them, but let's step back and ask the more general question:

Q3*. What motivates an acquaintance requirement in the first place? That is, why should we accept *any* form of Acquaintance Theory, over a more Latitudinarian theory?

This question interacts in complicated ways with Q1 and Q2. First of all, it matters what phenomenon we're trying to explain (Q1). If we're trying to explain the felicity of *de re reports*, then our linguistic behavior appears on the face of it too lax to uphold any Acquaintance requirement. In fact, it appears even laxer than any Latitudinarian theory would predict. A quick search on Google turns up thousands of examples like these:

- “Albert J. Pirrot, the husband of Jeanine F. Pirrot and a powerful man in his own right, wants you to know that he is no surburbosexual.”
- “Liberals think you and I are too stupid to make decisions for ourselves, so they design government to take decision-making out of the hands of citizens...”

This looks like a *de re* desire-attribution to Mr. Pirrot, and a *de re* belief-attribution to the Left. Yet we can quite consistently imagine that Mr. Pirrot and the Left are unacquainted with you and the speaker, and don't even know any descriptive condition that you uniquely satisfy. So I'm pessimistic about the prospects for a principled, well-motivated defense of an Acquaintance requirement for views that make *de re* attributions methodologically primary.

I'm also pessimistic about the prospects for an Acquaintance requirement on the cognitive-architecture-based understanding of what *de re* thought is.

If, on the other hand, we understand *de re* thoughts to be thoughts that your capacity to have is grounded in informational relations, then an Acquaintance requirement is built in at the ground floor. You're acquainted with α just in case you have information about α , and that's needed for you to have *de re* thoughts about α , because that's just what we *mean by* a *de re* thought about α . This is the understanding of *de re* thought that I'll be working with. (See also Bach...) It doesn't capture every intuition that

people have voiced under the name “de re,” but I don’t think that any sufficiently detailed view will. It’s enough for me if I’ve identified *one* coherent and interesting category of thought.

When I individuate thoughts of this category, I’m going to assign them a distinctive type of content. (It might, e.g., be object-dependent content. I want to be as ecumenical as possible about Millian/Fregean debates, and about how to account for cases of de re thought with no real object.) For now, let’s just call this distinctive type of content a “de re content,” leaving it unsettled what precise properties distinguish it. When I individuate thoughts that *aren’t* of this category, I’ll assign them *other* types of content. What having a de re content *amounts to*, then, is being such that a subject’s capacity to have you is grounded in certain informational relationships. The connections between those contents and the semantics of thought-*attributions* will be complicated.

For our discussion, though, we’ll need some way to talk about correlations between semantic properties of sentences and the contents of our thoughts. One way to do this would be to talk about the belief you *manifest or express* with a given sentence. But that has the difficulty that one generally manifests *many* beliefs by uttering a single sentence. Here’s how I’ll proceed. Sometimes we think in words. That is, we think thoughts *by* thinking or rehearsing sentences to ourselves. So we can talk about the semantic properties those sentences have (when understood as we understand them), and we can talk about the thoughts we have *by* rehearsing those sentences to ourselves, and what their contents are. This “by” relation is a tight one. If you rehearse the sentence “Mt Everest is 8850 m tall,” and simultaneously think metalinguistic thoughts *about* that sentence or about your act of rehearsing it, they won’t be thoughts you have *by* rehearsing the sentence. (They may be thoughts you have *in part* by rehearsing the sentence; but I’m not discussing those thoughts.) Similarly, if Everest reminds you of Hillary, and you have an image of Hillary cresting the summit, your thoughts about Hillary

won't be thoughts you think *by* rehearsing "Mt Everest is 8850 m tall." It'll just be a thought *prompted by* those words.

Two remarks. First, I don't assume that the semantic content of a sentence (understood as you understand it) and the content of the thought you have by rehearsing the sentence will always coincide. They may or they may not. This is a topic for us to investigate. Second, I happily acknowledge that we can think thoughts *without* rehearsing sentences to ourselves. I expect we can even think *one and the same thought*, both in words and without words. I'm just focusing on cases where *we do* think in words to make certain aspects of our discussion easier.

Well, that's my basic understanding of what de re thought amounts to, and a sketch of how I'll handle the connection between language and thought. There's a lot that needs defense and explication. But I'm not going to try to defend it here. Instead, I'm just going to take this understanding of de re thought on board, and use it to talk through some problem cases concerning which de re thoughts count as a priori. I'm going to try to motivate the answers it gives.

2 A Priori Justification

I understand **a priori justification** to be justification that doesn't derive from experience (and a priori knowledge to be knowledge where your justification is a priori). Under "experience" I include perceptual experience and I also include introspective awareness of your own sensations, intentions, and thought-processes. To be sure, there are important differences among these different types of awareness. But it seems natural to me to classify my knowledge that I have a toothache and my knowledge that I'm thinking of a number greater than 10 as a posteriori. My justification for believing those things is more like the justification I get from perception than it is like the justification I rely on when doing math and metaphysics.

Now, I have no *general* objection to the possibility of a priori justification for de re thoughts. For example, consider the de re thought, of Scott Soames, that if he exists then he exists. One needs to have had certain kinds of experiences, experiences that ultimately trace back to Soames, to be in a position to entertain this thought. But those experiences aren't the source of your justification for believing the thought. Your justification for believing it comes from your understanding of \supset . That justification seems to be a priori.

One sticky point concerning a prioricity and de re thoughts has to do with existential commitment. Consider the thought *Soames is self-identical*. Arguably that entails *Soames exists*. I'm reluctant to accept that you can have a priori justification to believe the latter. (This is disputed; but the natural view is that you don't have a priori justification to believe *Soames exists*.) So I'm reluctant to accept that you can have a priori justification to believe *Soames is self-identical*, either. There's a lot to sort out here. Some epistemologists deny that having a priori justification to believe one thing means you have a priori justification to believe what it entails. Some free logicians deny that *Soames is self-identical* entails *Soames exists*. My approach is to turn to "hedged" thoughts like *If Soames exists, he is self-identical*. I take these thoughts *not* to be existentially committing. That is, they *don't* entail that Soames exists. However, they are still de re thoughts about Soames. If Soames didn't exist, then arguably I wouldn't be able to think *any* of these de re thoughts about him. My situation looks like this:

- (1) Soames exists.
- (2) Soames is self-identical.
- (3) If Soames exists, he is self-identical.

If Soames doesn't exist, then I can't think de re thoughts about him by rehearsing *any* of these sentences. So from the premise *that* I'm thinking a de re thought about someone by rehearsing Sentence (3), I can infer that I'm

thinking a *true* de re thought about the same person by rehearsing Sentence (1). However, if it's *merely true* that I'm thinking a de re thought about Soames by rehearsing Sentence (3), then from the thought I thereby think, I'm not rationally entitled to infer the thought I'd be thinking by rehearsing Sentence (1). In general, it can take more to justify you in believing *you have* a de re thought by rehearsing some sentence to yourself, than it takes to justify the thought you thereby have.

I think this is the natural way to understand these sentences, and the thoughts I have by rehearsing them. It will take some care to articulate a free logic and semantics that validates the picture I've just sketched. But it can be done.

Examples we've considered show that experiences can be necessary for you to entertain a thought, without playing a role in your justification for believing the thought. In some cases, however, the experiences that enable you to entertain a thought *will be* a source of justification for the thought. Here are some examples. You look out the window and see a canary, and form the thought *That canary exists*. You hear a student's cellphone ringing, and think *That phone [auditorily demonstrating it] is ringing*. Stumbling about in a dark room, you place out your hand and feel a wall; you form the thought *This wall is perceived by me*. In each of these cases, you require certain experiences just to be able to entertain your demonstrative thought. And in each case, the experiences that enable you to entertain the thought arguably give you all the a posteriori justification you need, to be justified in believing that thought. Your visual experience of the canary both enables you to think the demonstrative thought *That canary exists* and justifies you in believing it. Your auditory experiences of the ringing phone both enable you to think the demonstrative thought *That phone is ringing* and justify you in believing it. And so on. We'll talk about these cases more later. But I'll announce up front that I don't regard your justification in these cases as a priori. Your justification to believe *That canary exists* seems to me

to come from the same source as your justification to believe *Some canary exists*: namely your perceptual experiences. Justification from that source is a posteriori. So the experiences that acquaint you with an object and enable you to have de re thoughts about it will in some cases be part of what justifies you in believing the thought, and in other cases won't.

You *may* have a priori justification to believe that thoughts *of the form* "That [auditorily demonstrated] phone is ringing" are true *whenever they're entertained*. But this won't be enough to give you a priori justification to believe a thought of that form. We will return to this issue later.

I said there's no general difficulty with having a priori justification for de re thoughts. We're going to discuss some problem cases, though, where I think you don't have a priori justification. These include the cases I just described. They also include cases where you rehearse to yourself sentences like these:

- (4) If that phone [auditorily demonstrating the ringing phone] exists, then it's ringing.
- (5) If I'm demonstrating any painting, then I'm demonstrating *this one* [pointing].
- (6) If I'm addressing anyone, then I'm addressing *you*.
- (7) I'm referring to myself.
- (8) I exist.

At first glance, these sentences all look to be what Kaplan called "indexically valid": their characters are such that in every context they express truths (if they express anything at all). But there are tricky details here, having to do with how we account for the demonstrations in the first two, and with whether a sentence needs to be uttered in every context that its

character is defined over. At the end of the day, I doubt that all of these sentences will still count as indexically valid. But they'll all be what we might call **utterance valid**: anyone who utters one of these sentences, understanding it as we do and with the same demonstrative and referential intentions that we have, will semantically express a true proposition, if he expresses any proposition at all. That's a nice property for a sentence to have.

Let me introduce some more shorthand. If a sentence is such that the thought you have by rehearsing it to yourself is justified a priori, I'll call the sentence **thinkable a priori**. Kaplan tells us that indexically valid sentences are always thinkable a priori. Many philosophers have agreed with him. They'd also agree that *utterance valid* sentences are always thinkable a priori.

Why should we believe them? There seem to be *many* sentences whose character is true in every context, but which wouldn't be thinkable a priori. For instance, there's:

(9) Cicero = Tully.

By Kaplan's own lights, names aren't indexicals and they have the same content with respect to every context. So if 'Cicero' and 'Tully' are coreferential, then (9) will be true in every context. But the thought I have by rehearsing that sentence to myself is not justified a priori.

But perhaps that last claim is too hasty. The thought you have by rehearsing to yourself:

(10) Cicero = Cicero.

is justified a priori, and on some accounts, *that's the same thought* you think by rehearsing (9) to yourself. You just think it under different "guises" in the two cases. This is all controversial. Fortunately, we can finesse the controversy by considering instead:

(11) Cicero \neq Tiberius.

Even on a direct reference account of what thought we're having by rehearsing this sentence, it isn't obviously the same as any thought that's justified a priori. Yet it too is true in every context.

At this point, some philosophers will argue that *we should* treat names as a kind of indexical, after all, and that it's only sentences whose characters *in that sense* are true in every context that count as thinkable a priori. I have many objections to that proposal, but let's not pursue them. For I think there's a general picture here that's driving those who accept a connection between being indexically (or utterance) valid and being thinkable a priori. That general picture emerges in two arguments I'll call the **Disquotation Argument** and the **Self-Verifying Argument**. It's a compelling picture, and I think the confidence it generates will resist being unseated by a few counter-examples or difficulties of detail about how exactly to formulate the connection between being indexically valid and being thinkable a priori. If we're going to resist such a connection, we'll have to do it in a more systematic and comprehensive way. That's what I'll try to do.

3 The Disquotation Argument

Here's how the Disquotation Argument will go for the sentence "If I'm demonstrating anything, then I'm demonstrating that."

- (12a) My understanding of 'that' justifies me a priori in believing that if I'm demonstrating anything, then 'that' refers to that thing.
- (12b) I understand my current use of 'that' and am thinking thoughts with it. That understanding justifies me a priori in believing that 'that' refers to that.
- (12c) From (a) and (b), I can infer that if I'm demonstrating anything, then I'm demonstrating that.

To diagnose this argument properly, we need to consider three kinds of case. (These cases deal only with sentences like (4)–(6). We’ll come back later to the Cogito sentences (7) and (8).)

I’ll say we have a **Perceptual Case** if you’re currently acquiring information about the object you’re referring to, and demonstrating it or referring to it *via* that information. The case where you auditorily demonstrate a ringing phone is of this sort. So too are ordinary cases of using ‘that’ and ‘you,’ when you intend to be referring to objects you’re then perceiving.

I argued earlier that your justification to believe thoughts like the one you have by rehearsing “That phone [auditorily demonstrating it] is ringing” comes in part from the experiences that acquaint you with the phone and enable you to think *de re* thoughts about it.

(Some sentences you rehearse to yourself in Perceptual Cases are very non-committal about what your external environment is like. They aim only to articulate your referential intentions. For example, consider the sentence “If I’m demonstrating any painting, I’m demonstrating this one.” In a case where your demonstrative intentions are *de re*, this is epistemically on a par with the sentence “My intent is to demonstrate that painting, if it’s there to be demonstrated.” Plausibly, awareness of what your referential intentions are will be enough to justify the thought you have by rehearsing that sentence. But that kind of introspective awareness is *a posteriori*, not *a priori*.)

[MORE DISCUSSION NEEDED HERE]

I’ll say we have a **Speaker Reference Case** if, regardless of whether you’re then acquiring information from anything, there’s some object you *independently* have in mind, and you intend and believe yourself to be referring to *that* object. It’s not important whether your belief that you’re referring to the object you have in mind is true or justified. For example, suppose you point blindly over your shoulder and say “I’ve had this picture for many years,” intending and believing yourself to be demonstrating your favorite picture of Carnap. In fact your picture of Carnap has been replaced, and

you're really demonstrating a new picture. Still, it's very plausible that by rehearsing "I've had this picture for many years," you'd be *thinking about* your picture of Carnap, not the new picture.

(Another example: I hear rumors about my friend Bill's latest hijinks. I shake my head and say, "Bill, you're a real handful." Here I've addressed Bill, and referred to him as 'you,' even though he's absent. Next day I see a stranger who looks a bit like Bill. I approach the stranger and ask, "Bill, is that you?" Plausibly this is another case where I'm addressing Bill, and referring to him as 'you.' I just have a false belief that Bill is the person I see. Contrast a case where I see the stranger pulling a cat's tail, and I say to him, "You'd better stop that, Bill." In this case I'm addressing the stranger rather than Bill. Or a case where I go up to the stranger and ask, 'Excuse me, are you Bill?'...)

How does the Disquotation Argument fare for these Speaker Reference Cases? Well, it's clear that the thoughts I have by rehearsing *I'm demonstrating this picture*, thinking thereby of my picture of Carnap, may not be true. I may be demonstrating some new picture. My understanding of *I'm demonstrating this picture* won't by itself give me any guarantee or promise that that's not so. So I think in Speaker Reference Cases, I'm not a priori justified in believing that the object I'm actually demonstrating, and semantically referring to with 'that,' is the same object that I'm thinking of by rehearsing 'that' to myself. In other words, I'm not a priori justified in believing the thought I have by rehearsing "That' refers to that."

I'll say we have a **Blind Case** if you're not acquiring information about the object you're trying to refer to, and you don't have an object independently in mind, either. For example, suppose you know on purely general grounds that there's a picture behind you, but you're not currently perceiving it, and there's no picture you're independently acquainted with that you're assuming it to be. You point over your shoulder blindly and ask, "I wonder if this picture is expensive?"

Or consider a case where you know your enemies have arranged for a thief to break into your house at midnight. You load your gun and wait in the dark for the thief to arrive. You doze off, and wake up at 12:15. You haven't heard the thief, but you believe that he'll be in the house by now. As you reach for your gun, you think: "You may have gotten into my house, but you're not getting out."

How does the Disquotatation Argument fare for Blind Cases? It's controversial what semantic properties your utterance has in a Blind Case; but to keep the discussion simple, let's suppose you manage to *semantically refer* to the painting and the thief you don't see. In a Blind Case, though, your epistemic position won't provide you with the means to *think de re thoughts* about those objects. Your demonstrative intention will have to be a descriptive one, e.g., to refer to whatever painting happens to be located behind you, or to refer to whatever thief happens to be in your house right now. Now, I said your epistemic position won't provide the means to think the relevant de re thoughts. I didn't say that it *prevents* you from thinking them, or that you'd necessarily be *unable* to think them. That depends on the details of the case. The object you're semantically referring to may happen to be an object you *are* acquainted with. For example, the picture behind you might happen to be your favorite painting, titled *Robin's Grandpa*. In that case, you *would* be able to think de re thoughts about the painting behind you. But you don't know that. Because your demonstrative intentions are descriptive, rather than de re intentions to refer to *Robin's Grandpa*, you won't be thinking de re thoughts about *Robin's Grandpa* by rehearsing sentences like "This painting is expensive" or "If I'm demonstrating any painting, then I'm demonstrating this one."

Does that mean you won't understand what you're saying? Well, by rehearsing those sentences, you don't manage to think the de re contents you're semantically expressing. That can fairly be regarded as *one* kind of failure of understanding. Note, though, that *this* failure has some precedents.

Whenever you rehearse a sentence containing terms that refer to Smith, while having Jones in mind as your speaker's referent, plausibly the thought you think by rehearsing the sentence (a thought about Jones) won't have the de re content that you're semantically expressing (about Smith). Cases of deferred reference can provide more examples. These are cases where you demonstrate one thing in order to demonstratively refer to another. For example, I might point to a house in order to demonstratively refer to its owner. Suppose the house looks undistinguished, and that perceiving it doesn't give me de re thoughts about its owner. Nonetheless, I can still demonstratively refer to the owner, by pointing to the house with the right intentions and saying 'he.' It's disputed what the semantic content of that utterance would be. But on some views it's a de re content. In this case, too, then, some would say I've semantically expressed a de re content that I'm not in a position to think. I lack that kind of understanding of what I've said.

Of course, understanding comes in different grades. The failure of understanding you're exhibiting here is compatible with various kinds of understanding *success*. You understand the standing meanings of the words you're using. You're acquainted with the context in which you're using them. You know what conditions have to be satisfied for something to be the referent. And so on. All of this plausibly amounts to *some* kind of understanding of what you've said. (Compare Salmon's distinction between strong/weak understanding.)

(If you don't think de re thoughts when you rehearse these sentences, then what do you think? (i) a descriptive thought? "The painting located there [mentally demonstrating an egocentric location you're already acquainted with]..." (ii) a non-referring de re thought? (iii) a de re thought about a painting you're imagining, and imagining to be behind you. This isn't a real painting you're independently acquainted with. It's *a product* of your imagination, in the way that Sherlock Holmes is a product of Conan Doyle's literary efforts.)

Let's return to the Disquotation Argument. In Blind Cases, your understanding of 'that' won't justify you in believing, of any object, that you're referring to *it* with 'that.' Nor will your understanding of 'you' justify you in believing of anyone that you're referring to him. You may happen to be *able* to entertain those de re hypotheses. For example, you're able to entertain the de re hypothesis, of *Robin's Grandpa*, that you're referring to it with 'that.' But there's nothing in your understanding of 'that' that gives you justification to believe that hypothesis, or raises its epistemic probability for you.

Sometimes it's claimed that you *will* have justification to believe you're referring to *Robin's Grandpa*—perhaps not when you think of it *as Robin's Grandpa*, but rather when you think of it under the demonstrative guise of "that." I don't see how this move can help. If I think of you *as* a hero, that doesn't by itself make me any more justified in believing that you're a hero. So why should thinking of *Robin's Grandpa as* the painting I'm demonstrating be epistemically better? (If *Robin's Grandpa* could somehow be *presented* to me as the painting I'm demonstrating, that may be more promising. But to say that it's *presented to me* in that way makes it sound like I'm getting de re evidence about *Robin's Grandpa*, to the effect that it's the painting I'm demonstrating. The Blind Cases were precisely constructed so that there'd be no source for that justification.)

4 Cogito Thoughts and the Self-Verifying Argument

[THIS SECTION WAS PUT TOGETHER HASTILY. THE ARGUMENTATIVE PIECES ARE ALL HERE, BUT I'M SURE IT'S HARD TO READ. APOLOGIES.]

There's a second argument I've so far been ignoring. This second argument goes as follows:

- (13a) If I understand a sentence, and it's such that anyone who thinks a thought by rehearsing it thereby thinks a true thought, then I can know a priori that it has that property.
- (13b) So long as there's no independent speaker's reference, many of the sentences we've been considering are of this sort. This is most obviously true for the Cogito sentences "I'm referring to myself" and "I exist,"
- (13c) I can think those sentences to myself, and know a priori that the thought I'm thereby having has to be true, no matter what my particular context turns out to be like.
- (13d) So I'm justified a priori in believing those thoughts.

There are two ways to think of this argument as working. On the first account, my justification for believing the thoughts derives in part from my justification to believe *the premise* that the thought I'm thinking has to be true.

The rebuttal to that is straightforward: so construed, the argument assumes that I'm justified in believing *I am* thinking a thought of the sort described. No doubt I will be entitled to that assumption, if it's true, but that's because I'll be introspectively aware of what sorts of thoughts I'm thinking. My justification will then in part be introspective and a posteriori, not a priori. All I can know a priori is that the thought-type is true whenever it's entertained. I won't know a priori that I'm entertaining it.

(Compare. Take a thought-type that's true whenever entertained by someone distinct from Brian, e.g., "If Brian exists, I am distinct from Brian." The mere fact that a thought of that type *is* entertained from someone distinct from Brian—say by Kevin—isn't enough to make Kevin justified in believing the thought. He may have no idea whether he's identical to Brian

or not. He only knows that when a thought of that type is entertained by a non-Brian, it's true. For that to help justify him in believing it, he'd need to know that *he's a non-Brian* entertaining it. Why shouldn't the same be true for thought-types that are true whenever entertained by *anyone*? For that to help justify you in believing the thought, wouldn't you need to know that *you are someone* entertaining it?)

The second account of the Self-Verifying Argument is harder to rebut. On this account, your justification for believing the self-verifying thoughts is *not* supposed to be inferential. Rather, *the mere fact* that a thought is self-verifying is supposed to give you justification to believe the thought. (Or, alternatively: *understanding* that a thought-type is self-verifying should give you justification to believe it.)

I'll resist this proposal with three analogies.

First, consider a practical analogy. The fact that, should you adopt X as an end, you'll achieve it, is *by itself* no reason for adopting end X. (Of course, achieving ends you've adopted may tend to have happy consequences, like lifting your mood; if so, then adopting end X may thereby acquire some instrumental merit.) If someone said to you, "The desire to have desires is an especially happy desire, because it's self-fulfilling. The desire to lack desires, on the other hand, is an especially unhappy desire, because it's self-thwarting. So you should have the desire to have desires." Why can't you respond, "Why? I neither desire to have desires, nor desire to lack them. How does the fact that the desire to have desires is self-fulfilling give me a reason to have it?" Similarly, the fact that, should you believe content C, your belief will be true, is by itself no reason to believe content C.

Some will protest: "But there's a difference! Having fulfilled desires isn't intrinsically valuable, true, but having true beliefs *is*."

It's not clear that having true beliefs is intrinsically valuable. It may be intrinsically better to believe the truth than to believe the false, but why should it be intrinsically better to have a belief at all, rather than to withhold

judgment? In any case, even if having true beliefs is intrinsically valuable, that would only provide a *practical* motive to believe self-verifying contents. It wouldn't be an epistemic reason to believe those contents.

Second analogy. Suppose Todd is especially stubborn. To date, he's never yet changed his mind once he's settled it. In fact, he believes he will never change his mind. You say to Todd, "Look Todd, consider what would happen if you came to believe that you will someday change your mind about something. That would itself constitute a change of mind; hence if you came to believe it, you'd thereby make it true." Todd says, "OK." You say, "So don't you see? That's an epistemic reason to believe it. That's a reason for you to think *you will* change your mind." Todd would look at you bemused. And rightly so.

Third analogy. Consider the following sentence:

(14) I am uttering a sentence.

(14) can only express a definite content if there is some agent to be the referent of "I." So it can only express a contentful thought when it's uttered. (Let's count rehearsing a sentence to yourself as a kind of utterance.) So given what the sentence means, it follows that whenever it's used to think a contentful thought, the thought it expresses is true. Anyone who understands the sentence is in a position to know that. However, suppose you do utter the sentence (either privately or aloud). What then justifies you in believing that you are uttering it, or any sentence? It can't be your understanding of the sentence. That would only justify you in having beliefs about what's true when the sentence is being uttered. It doesn't help you determine when that condition is fulfilled. The natural thing to say is that what justifies you in believing that you are uttering the sentence is your introspective or perceptual awareness of uttering it. Hence, your justification for believing the thought that (14) expresses is a posteriori—despite the fact that you know, just in virtue of understanding (14), that whenever it expresses a contentful thought, that thought is true.

I think the same holds for sentences like “I exist.” Your understanding of that sentence is enough to justify you in believing that whenever it expresses a contentful thought, that thought is true. We can count that justification as a priori. But your justification for believing the thought that “I exist” expresses is not itself a priori. What justifies you in believing that you exist would be something like your introspective awareness of your own thought-processes—including perhaps the act of entertaining the thought that you exist.