

Guide to Chapter Seven of Gareth Evans' *The Varieties of Reference*

Rick Grush
Department of Philosophy 0119
University of California, San Diego
La Jolla, CA 92093

7.0 Chapter Seven Overview

In this chapter Evans discusses 'I'-ideas. Like 'here'-Ideas, 'that'-Ideas and 'now'-Ideas they have i) a functional aspect consisting of unmediated connections to perception and action, and ii) a conceptual aspect that provides an objective understanding of the object. Because the first component is in play, 'I'-Ideas, like all the others, support IF judgments. But there are distinctions. First, unlike 'that'-Ideas, but like 'here'- and 'now'-Ideas the functional connections can be latent. And second, unlike 'now'- 'here'- and 'that'-Ideas, there is only one object tracked as 'I'. This makes the cognitive workload easier, but introduces issues concerning the identity of the self through time.

There are several main topics discussed in the chapter (my numbering of these topics does not correspond to section numbers). First, the fact that 'I'-judgments are often based upon special ways of gaining information — ways that only carry information about property instantiation and not about an object — has encouraged some authors to claim that 'I' does not refer.

Evans argues against this by maintaining that these connections are only one aspect of our 'I'-Ideas, the other 'objective' aspect does in fact specify a referent.

Second, issues about the connection between an objective understanding and special functional connections comes up in several ways — the fact that no indexical-free description can entail an indexically specified proposition (Castaneda, Perry); or that subjective differences have no objective consequences (Nagel). In such cases, the divide is one that is unbridgeable, so to speak, by inference. Evans agrees, but points out that the divide is bridged by functional (not rational or descriptive) connections. This issue is picked up again in the last three sections of the Appendix.

Third, Evans argues that our 'I'-Ideas are Ideas of a subject with both physical and psychological features. We have special ways of gaining both physical and psychological information about ourselves, and in both cases it supports IF judgments. But the case of our psychological judgments is complicated, which leads to the next topic.

Fourth, Evans discusses how we make psychological self-ascriptions, of both belief states and experiential states. The suggestion is that because the content of beliefs and experience concerns the world, we can get a handle on the representing states by assessing how they present the world to us. I assess whether *I believe* the apple is ripe by assessing whether the apple is ripe. The answer to the former question is *yes* if I answer the second question affirmatively. And I determine that *I am experiencing* red (when I look at an apple) by assessing whether *the apple* is red. In both cases, I use a conceptual understanding of the psychological predicate (belief, experience) to understand the result of such assessments. In the first Appendix, Evans continues this topic for the case of self-reference.

Fifth, Evans has an extended discussion of memory, and argues that memory supports IF 'I'-judgments, contrary to what some authors have claimed. The discussion here is parallel to the same topic in the context of 'here'-judgments.

7.1 Introductory

This initial section covers five issues. First, Evans discusses similarities and differences between self-identification and other forms of identification he has discussed earlier, and notes that 'I'-Ideas share features with 'here'-Ideas and 'that'-Ideas. Namely, they are constitutively tied to non-mediated ways of gaining information, and non-mediated ways of acting. Second, he points out that some people have argued that there is no specific content corresponding to 'I' in thought about oneself (that is, no such thing as an 'I'-Idea), on grounds that 'I am in pain' could be equally well expressed by 'there is pain', or 'I'm in an awful muddle' as 'This is an awful muddle!' Evans points out that it is a consequence of the fact that our 'I'-Ideas have functional connections to ways of gaining knowledge that are IF that information received from one of these channels to the effect that pain is instantiated (i.e., there is pain) just is for it to seem like I am in pain. But Evans will argue that while such functional links are part of what is involved in our 'I'-Ideas, there is more. Namely, that each of us understands that 'I' am identical to some objectively conceived person. (This is parallel to the fact that my 'here'-Idea has not only functional connections to my surroundings, but involves knowing what it is for it to be an objectively conceived location. One's 'I'-Idea has both subjective (ways of gaining information and connections to action) and objective aspects. Third, Evans addresses an argument to the effect that it is not possible for anyone to understand what it

would mean for there to be an identity between one's subjective self and an objectively conceived person. The argument hinges on the claim that the justification for such an identity would have to be found in the objective conception of the world. Evans agrees that there would be no justification to be found there, but insists that the justification is found in how the subjective situation is conceived. Fourth, Evans points out that while many authors have assumed that if there is anything like an 'I'-Idea, it concerns a purely mental subject. Evans points out that his doctrine is that this is not the case — our 'I'-Ideas are Ideas of things that are both mental and physical. Fifth, Evans issues some cautions about our understanding of the persistence of the subject's identity through time.

7.1.1 Preliminaries (beginning of 7.1 to bottom of p. 207)

First Evans sets up the discussion by first specifying what sort of 'I' thoughts are at issue, and then comparing such 'I'-thoughts with 'here'-thoughts and 'that'-thoughts. The sort of 'I' thoughts in question are self-conscious thoughts, that is, thoughts in which one is thinking of oneself *as oneself*. In Perry's example, he was thinking of himself when he saw the trail of sugar on the grocery store floor and thought 'That person should check their cart for the torn bag'. But though he was thinking of (the person who was in fact himself) because he didn't realize it was himself he was thinking of, he wasn't thinking an 'I'-thought in the sense we are concerned with.

Next, Evans points out that understood in this way, 'I'-thoughts have some similarities to 'here'- and 'that'-thoughts. One's 'I'-idea is constitutively tied to ways of gaining information and behavioral dispositions in a way that is identification free. There are information channels such that getting

information to the effect that '... is facing a tree' is instantiated (a visual scene that consists of a nearby tree) just is for me to have information to the effect that *I* am facing a tree — even though *I* don't explicitly figure in the perceived scene. Similarly, my knowledge that a bear is about to attack me will have certain kinds of direct consequences for *my* actions — consequences by no means guaranteed by my belief that *the tallest member of the UCSD philosophy department is about to be attacked by a bear*.

Evans summarizes a significant aspect of his account of the nature of 'here'-, 'that'- and 'I'-thoughts on p. 207:

... the Idea which one has of oneself involves the same kinds of elements as we discerned in the case of, say, 'here' (6.3): an element involving sensitivity of thoughts to certain information, and an element involving the way in which thoughts are manifested in action. Such an analysis would certainly explain the widely recognized irreducibility of self-conscious thoughts to thoughts involving definite descriptions: for no descriptive thoughts could guarantee the existence of these special dispositions. (p. 207)

So the phenomenon discussed by Anscombe, Casteneda, Perry and others to the effect that certain of these thoughts cannot be analyzed as descriptive thoughts is given a unifying treatment on Evans' account. It is constitutive of these Ideas that they have dispositional connections, unmediated by identity judgments, to ways of gaining information and acting. No description, not even one making reference to information links, can guarantee these connections.

7.1.2 Can there be 'I'-Ideas? (top of p. 208 to middle of p. 210)

From the top of page 208 to the middle of 210 Evans discusses the view that 'I' is a purely communicative device whose job is to get the *audience* to think of the correct object, but it doesn't correspond to any Idea one has of oneself that is used in thinking of oneself. This view was maintained (in slightly different forms), by Wittgenstein, Geach, Strawson, and Anscombe.

Geach and Strawson both point out that, in effect, in my judgment that *I am in an awful muddle* or that *I am in pain*, there does not appear to be anything corresponding to an *identification* of something that is in a muddle, or is in pain. (And here we are using 'identification of x' in the sense of 'discerning x from other things', not in the sense of "judging that x is identical to y.") Notice that the example involves only myself making a certain kind of *judgment*, not communicating anything to anyone. The judgments might as well be expressed as *What an awful muddle!*, or *There is pain!* Of course, if I am speaking to someone, I would need to use 'I', so that the person I am talking to would know who I am talking about, who it is that is in pain, for example. If I were to say "What an awful muddle!" or "There is pain!" that might mean I recognize that someone else is in a muddle, or in pain. And so 'I' is used to make sure that the audience knows who I am speaking of.

The premise in both cases is that there seems nothing in these first-person judgments that corresponds to getting myself to think of the right object. They then conclude that, at least in the case of one's own self-related judgments, 'I' does not have any substantive content, nothing corresponding to what Evans would analyze as an *Idea* of some sort (e.g., an 'I'-Idea). Evans has two lines of argument against this.

The first is to point out that at best the premise (there is no question of getting myself to think of the right object) just shows that 'I'-thoughts are identification-free, in the sense discussed in Chapter 6. The functional connections an 'I'-Idea would have to ways of gaining information will result

in IF judgments being made on the basis of *information* whose content doesn't include anything corresponding to 'I'. But such information channels are only part of the 'I'-Idea.

This leads to Evans second point, which is that my 'I'-thinking surely conforms to the GC. I can ascribe to myself predicates that are not decidable on first-person grounds (such as "... will be dragged unconscious **thought** the streets of Chicago"). And it is also the case that when I ascribe to myself a predicate like "...is in pain", the predicate I ascribe is one I understand to be ascribable to others. Both of these considerations point in the same direction: part of my self-knowledge is that I am a person like others. That is, I have a fundamental Idea of persons and in order for my thinking to be able to link up with this fundamental (objective) Idea, there must be an 'I'-Idea that I can conceive to be identical to the fundamental Idea. That is, I must be able to grasp $[I = \delta]$ where δ is a fundamental identification of a person, and this identity needs something on the left hand side. If something like this weren't the case, its not clear what could be made of my knowledge that I am judging that *the same person* is involved in my judgments that *I am in an awful muddle* and *I was dragged unconscious through the streets of Chicago*. (See my remarks on the fifth part of this section below.)

7.1.3. Subjective and objective

From the middle of p. 210 to the middle of p. 212 Evans contrasts his view on self-knowledge with that of Nagel, who claimed that such an identification – of ourselves with an element of the objective order – is not possible, because of an unbridgeable divide between the subjective and objective. Evans' recap of Nagel's argument is something like this: When we

conceive of an objective state of affairs, such as *T.N. standing before a burning house*, then the identification of myself with T.N. (or with any other person in the objectively construed world) makes no difference to how the world is conceived. All the objective facts are the same. Therefore, there isn't really any content to the supposition that I am identical with an objectively construed person. Since such content would presumably be the content I would need to grasp in order to judge that I am identical with an objectively construed person, there would seem to be no such thing as knowing what such an identity consists in.

Evans agrees that there must be some content to the supposed identity claim, but that Nagel's argument assumed that that content would have to be in terms of a difference made in how things are conceived objectively. Evans reply is that the difference can be in terms of how things are conceived subjectively. If I am identified with T.N., for example, then my environment will include a burning house; if I am identified with R.G., then it will not. To switch examples: suppose I am looking at a live video feed of a birds-eye view of a large mall. In the video feed I see one person (A) at one location, and another person (B) at another location. It is true that my supposition that 'I am A' will make absolutely no difference to anything represented in the video feed. So Nagel is right about that. But this supposition does make a difference to how things are conceived as being *with me*. If I am A, then there is a food court right behind me; but if I am B, then the food court is ahead and to the left.

Evans admits that such knowledge can only be expressed via indexical expressions, but says that this is no objection to treating it as knowledge.¹ (At

¹ Evans seems to have not been entirely satisfied with his remarks here, since the last few sections of the Appendix return to this topic and raise worries.

the very least, it puts the burden of proof back on Nagel's shoulders.) At the top of p. 212 Evans says:

... just as our thoughts about ourselves require the intelligibility of this link with the world thought of 'objectively', so our 'objective' thought about the world also requires the intelligibility of this link. For no one can be credited with an 'objective' model of the world if he does not grasp that he is modeling the world he is in — that he has a location somewhere in the model, as do the things that he can see. Nothing can be a cognitive map unless it can be used as a map — unless the world as perceived, and the world as mapped, can be identified. [p. 212]

The point may not be immediately clear. Think of it this way. If it were really the case that our conception of the world were purely 'objective', meaning that it had *no* connection to anything subjective, then it would be an odd sort of conception — something like an impersonal description that one might find in a very complete history book. But in addition to this oddness, if there are *no* connections to anything that is immediately subjective (none of the locations mentioned are places I could visit, or in any egocentric direction from my current location, or at any time in the future or past relative to *now*), then the difference between this impersonal description of 'the objective world' and an impersonal description of a fictional world vanishes. Part of what makes my objective conception of the world an objective conception of the *real* world is the fact that it is the world that *I* am in. In order for a world to be the objective world, it is not enough that I can entertain the supposition that I am identical to one of its inhabitants; rather, it is such that I have no choice but to take at least one such identity must be true. (To doubt this for any world just would be to doubt whether it was the real world.)

This also lets us see Nagel's argument in a new light. Nagel's argument finds the supposition that *I am R.G.* to be mysterious because it is setting up the required identity to be analogous to the supposition that *I am Sherlock Holmes*. How I could establish this latter identity would of course be mysterious. But it is of the essence of the case that the two suppositions are not parallel — one reason being that the only justification for taking the impersonally described world to be the actually existing world (as opposed to a fictional world) is that some such identity with the subjective is necessary. (And if there were any justification for me to identify myself with Sherlock Holmes, it would, as Evans correctly points out, not come from anything describable in objective terms, but from what I experience subjectively.)

7.1.4 'I' is not an exclusively mental self.

From the middle of page 212 to the middle of page 213, Evans discusses the view that our notion of ourselves is exclusively that of a mental subject. Evans here claims that it is true that it is essential that our self notion be a notion of a psychological subject — a subject of thought and judgment. But this does not impugn the notion that bodily and other physical self-ascriptions are also necessary. The discussion here is making clear what the doctrine is, there is no real argument in favor of the doctrine here. And in fact, there isn't a distinct developed argument for this anywhere in the chapter. But materials for such an argument are scattered throughout the chapter and the Appendix.

7.1.5 'I' through time.

Finally, from the middle of page 213 to the end of the section of the middle of page 215, the topic is personal identity through time. The issue is that in one's memories about oneself, and anticipations about oneself, it can seem as though there is some mysterious guarantee that one is always able to 'get at' the correct object. That in my current memory "I was cold yesterday" I can't be thinking of anyone other than myself; or in 'I shall experience pain tomorrow (during my dentist visit)' anyone else's pain but my own is at issue. This odd infallibility has made some suspicious of the proposal that there is any content to 'I' in such cases — a content that can't be wrong is no content.

Evans here merely points out that the assumed infallibility is just the result of the way the situation is described. If it is correct for us to describe the situation as one in which *a single subject* makes a number of self-identifications and identifies himself as the same subject in all cases, then of course the subject cannot be mistaken.

This topic is actually closely connected to the fact that 'I'-thoughts are IF, not just in terms of present-tense judgments, but also in terms of certain past- and future-tense judgments. This will be taken up in later sections in more detail.

7.2 Immunity to error through misidentification

This section has two parts. First, Evans explains why 'here' rather than 'this' is the closer parallel to 'I' as far as the kinds of mechanisms involved go. Essentially, it is because 'I' thoughts, like 'here' thoughts but unlike 'this' thoughts involve information links that can be latent, so long as the subject has the appropriate dispositions to act on any information that arises. Second, Evans discusses the phenomenon of immunity to

error through misidentification (which was discussed in the context of demonstrative identification in Section 6.6). His two concerns are first to argue that just because there is not a possibility of error, it does not follow that no object is identified (contra Wittgenstein, Anscombe, et al.); and second that the phenomenon is not tied exclusively to self-ascription of mental predicates. Rather is a function of ways of gaining information, and it can arise for self-ascription of both mental and physical predicates.

In the first part of this section, running from the beginning of the section through the first two-thirds of page 216, Evans compares 'I' thoughts with 'here' and 'this' thoughts. In all cases the Ideas are partially constituted by functional connections to information channels whose deliverances are directly deposited into the governing conception of the Idea. In the cases of 'I' and 'here', though, these channels can be latent, in that there need not be information actually coming in through the channel in order for the Idea to be in play. (The limits of the sufficiency of such latent dispositions is discussed in 7.6.)

The role of such information-links in 'I'-thoughts implies that such thoughts will exhibit the phenomenon of identification-freedom, in the sense Evans discussed in Chapter 6. Evans explores this in the second and larger part of this section, from the bottom third of page 216 to the end of the section at the middle of page 220. The issue is that the phenomenon of IEM (similar to, but different from, Evans' IF) has typically been thought to apply only to oneself qua *mental* subject. Evans traces this back to Wittgenstein, who chose examples such that mental predicates all turned out to be IEM ('use as

subject' in Wittgenstein's terminology), but physical predicates did not ('use as object').

But Evans claims that the distinction between identification-free and identification-dependent judgments is orthogonal to whether the judgments concern physical vs psychological attributions.

For a mental and physical case in which the ascription exhibits IEM/IF, consider 'I am seeing red' thought/said as one is looking at a red wall in good light; and 'My legs are crossed', thought/said on the basis of normal proprioceptive feelings (not looking and seeing crossed legs). In these cases, neither '*Someone* is seeing red, but is it *I* who am seeing red?' nor '*Someone's* legs are crossed, but is it *I* whose legs are crossed?' seem to make sense. In both cases, according to Evans, this is because the judgment is made on the basis of the deliverances of an information channel that is constitutively tied to the relevant Idea.

For a case of each that does not exhibit IEM: a neuroscientist is looking at a number of live fMRI images that are displaying the neural activity of several subjects, one of whom is himself (though he does not know which image corresponds to which subject) and looking at one of the scans that shows activity in the appropriate region 7677thinks '*Someone* is angry, but is it *I* who am angry?'; an anaesthetized subject in a room or mirrors who sees crossed legs in a mirror and says/thinks '*Someone's* legs are crossed, but is it me whose legs are crossed?' In both cases, the way the information was gained (that anger is instantiated, that legs are crossed) is via a route that is *not* constitutively tied to the relevant Idea, and hence requires the mediation of an identity claim in order to yield an 'I'-thought.

In all such cases, what makes or breaks the IF status of the judgment is whether or not the information to the effect that the predicate is instantiated is acquired through an information link that is constitutively tied to one's

'I'-Idea as part of that Idea's functional characterization. In the 'I see red' and 'My legs are crossed' (based on proprioception) it is; in the cases of 'someone is angry' and 'someone's legs are crossed (based on a visual image in a mirror) it is not.

So that's one mistake, the idea that somehow the IEM litmus test (whether it makes sense to wonder "Someone is F, but is it I am am F") singles out mental attributions as IEM. The second mistake is based on the first. It is to try to mount an argument against a Cartesian conception of the self on the basis of the alleged IEM status of mental self-ascriptions. The idea is this. Wittgenstein thought that the Cartesian conviction that through introspection we know ourselves as purely mental subjects was wrong. His anti-Cartesian argument was to try to show that this supposed idea (of ourselves as a purely mental subject) was empty. And the argument was basically that in all these alleged self-thoughts, the self drops out. 'I am in pain' could be re-expressed without loss as 'there is pain'. The reason, on this line of thinking, that it is not possible to misidentify the self in these cases is that one never was identifying anything at all.

Evans is no Cartesian, but he thinks this argument is flawed. One way to see Evans point here is as follows. A judgment based on information, including self-ascriptive judgments, identify an object (*a*) as the object that instantiates the property in question (*F*). There are two ways this can happen. *First*, the deliverances of the information channels (or set of cooperating information links) may provide information about the identity of the object *a* along with information about the property being instantiated. *Second*, the information link might be constitutively tied to a certain Idea in such a way that any information incoming along this channel to the effect that some property *F* is instantiated gets treated automatically by the subject as

information to the effect that the object corresponding to that Idea instantiates the property. But in the second sort of case there is still an Idea of an object in play. If one does not recognize the second sort of possibility, then cases in which the information channel does not carry any identification information (that is, all cases providing for identification-free ascriptions) will seem to be extremely mysterious, for it will seem to be the case that the judgment based on the information is ascribing a property to nothing.

Evans is no Cartesian. But denying that the self is essentially mental is not the only way to deny the Cartesian picture. That picture is equally denied by an insistence that the self is essentially physical (even if it is essentially mental as well).

The section closes with Evans countering a potential objection. The objection is that Wittgenstein's topic is not whether there are ways of gaining information which *are* IEM. Rather, it is whether there are ways of gaining knowledge which are *not* IEM. The point is this: even if it turns out that some bodily self-ascriptions are IEM, it remains true that some are *not* IEM. And if it also turns out that in the case of mental predicates there are not any ways of gaining knowledge that are *not* IEM, then there is still a principled asymmetry between the self-ascription of mental and physical predicates. Evans responds by pointing out that there is not much textual evidence to support the idea that this was Wittgenstein's topic. And moreover, there are situations in which mental self-ascriptions that are not IEM are made, and so there does not seem to be any asymmetry between mental and physical self-ascriptions. This topic is somewhat off the main positive account, so I won't explore it further.

7.3 Bodily self-ascription

This section has two parts, both aimed at self-ascription of physical predicates. The first part discusses identification-free bodily self-ascription proper — this includes things like “I am hot and sticky”, “my legs are bent” and so forth. The second has to do with our self-ascriptions of location based on perceptions, such as ‘I am in my bedroom’, or ‘I am facing a tree’. Evans points out how having IF knowledge of the latter sort is especially important for our objective conception of ourselves as spatial beings.

This section concerns self-ascription of physical predicates. Self-ascription of mental predicates will be covered in the next section. There are two kinds of physical predicates that are discussed here: predicates concerning the body's state, and predicates concerning the body's location/orientation in space. Both kinds of self-ascription support judgments that are identification-free.

The first part of this section, running from the beginning at the middle of page 220 to the top of page 222, discusses our self-ascription of predicates concerning the body's state. This would include things like ‘My legs are bent’, ‘My hair is blowing around’, ‘I am being pushed’. For each of these, there are various ways we could come to make the ascription. First, there are information channels such that when information to the effect that these predicates are instantiated is acquired through these channels, it is deposited directly (that is, in an identification-free way) into the controlling conception of oneself. I can feel my hair blowing, I can have proprioceptive information about my crossed legs, I can feel myself being pushed.

There are also channels that are not directly connected in this way, but might still provide information that could lead to me making the judgment, though in a way that employs an identification. I could see a pair of crossed legs in a game of twister, or see blowing hair or someone being pushed in a mirror. I might then, judge that my legs are crossed, or my hair is blowing, or I am being pushed. But in such cases the judgment would be based on my acceptance of an identity: those legs are my legs; that person who was just pushed is me, and so forth.

One might (following a similar objection that Evans discussed in Section 6.6) worry that the supposed IF judgments aren't really IF, because one can imagine situations in which one might question whether the predicate actually applies to oneself. One can imagine situations in which the information coming in through one of these channels is actually providing information that is not from one's own body (this would require some brain interfacing technology, etc.) such that the feeling of blowing hair is actually coming from someone else's body. In such a case there could be a misidentification.

Evans makes three counters. The first (in the first full paragraph on p. 221) is that this fact does not entail that in normal cases the identification is based on an identity claim. This argument parallels the argument to the same effect in section 6.6.

The second and third focus on the identity proposition that this position requires — *I am b*. The second counter concerns *b*. In the situation the objector is envisioning, the subject is getting information from a deviant causal chain: it is from some body that is not one's own. Evans says:

The supposed Idea *b* could be adequate only if it involved identification by description, on the lines of 'the body from which I hereby have information'. (VR p. 221)

Keep in mind that in such cases the information one is getting does not have any identifying (in the sense of determining which object is involved) elements. The information is just “... is hot” is instantiated. Given this, the only grip one has on *b* is descriptive, as *the body that is the causal source of the information*. Evans points out that nothing like this seems to be involved in “the normal case.”

The third counter concerns the ‘I’-Idea involved in the identity. The argument Evans is countering is trying to establish that bodily self-ascriptions are not IF, so it must argue that *all* of our self-ascriptions are based on such identity judgments. The argument is claiming that the examples of deviant causal chains makes these identity judgments apparent, but they are always in play. Evans claims that if one’s ‘I’-Idea were such that all ascriptions to it of such predicates involved the mediation of such identity judgments, then it is not clear how one could identify oneself with a physical thing at all.

While it might not be immediately obvious what Evans means by this, it is a reference to arguments, some of which Evans has outlined elsewhere — others of which he was familiar with due to, e.g. Strawson — to the effect that the kind of identification we have with our bodies is stronger than that provided via any sort of *inference*. (The last section of the Appendix follows up on this to some extent.)

The second part runs from the top of page 222 to the middle of page 224, and discusses another kind of bodily self-ascription — the self-ascription of predicates involving one's position, orientation and comportment with respect to the environment. For example, I ascribe to myself the location of being in my bedroom when I wake up and see the familiar objects; I locate myself as being next to a tree based entirely on a certain kind of visual

experience *of a tree*. Evans notes again that this appears to provide for identification-free knowledge – having visual experience of a tree and wondering “*Someone* is standing in front of a tree tree, but is it *me* who is next to a tree?” is anomalous.

Evans claims that such an ability for identification-free ascriptions of one's location are a necessary part of the ability to conceive of an objective world at all. The argument here is a bit thin, but clearly Evans has in mind the sort of thing that he has argued for in his ‘Things without the mind’, as well as doctrines argued for by his teacher Strawson in *Individuals* and *The Bounds of Sense*.

More concretely: on Evans account I must have an adequate Idea of myself in order to entertain genuine ‘I’-thoughts, and that an adequate Idea of myself is made available in part via a capacity to identify myself with an entity in the objective order (in the sense of having a capacity to understand an identity judgment to the effect that *I am δ*). It is via a mastery of a theory of the sort described at the top of page 222 that underwrites this ability. It goes hand in hand with the theory Evans described in Section 6.3, concerning how one can find one's way around, and coordinate egocentric space and objective space (pp. 162-4) Is it, then, what makes our Ideas of ourselves adequate.

It might be confusing to see Evans, in nearly adjacent paragraphs, arguing i) that self-ascriptions cannot be such that they (in the usual case) require identity judgments, and ii) that one's ‘I’-Idea, in order to be adequate, must be such that the subject is always able to understand an identity judgment between their ‘I’-Idea and that of an objectively conceived person: *I am δ*. This gets at the core of Evans' position. The views he is arguing against are views that recognize two things: an objective view of the world (Nagel's

objectively conceived world, describable by indexical-free descriptions), and something like subjective access to information (Anscombe's "unmediated agent-or-patient conceptions of actions, happenings and states"). And what has caused all the fuss, in one way or another — for Perry, Anscombe, Nagel, and others — has been the recognition that there's no *rational* way to bridge the two. E.g., no set of objective descriptive truths can entail that it is *my* bag that is leaking sugar.

Evans' program recognizes a third mediating entity: the 'I'-Idea. There are two key aspects to this Idea. First, even though it is not by itself a fundamental conception of a person, it is a conception of something such that it makes sense to identify it with an element of the objective order. Second, and more to the present point, it is an Idea that is partially constituted as the kind of Idea it is by *functional* (not *rational*) connections to ways of gaining information. An 'I'-Idea that could only support self-ascriptive judgments only via identity claims would be unsuitable as an 'I'-Idea because it would lack the non-rational functional connections that compose the substructure of one's self-conception. But an 'I'-Idea that has these connections must be such that it can be understood to be an element of the objective order.

7.4 Mental self-ascription

This section has five topics. The first is the self-ascription of belief states. The account is roughly that one ascribes belief states to oneself not via inner perception, but by reusing the same processes of belief formation that one employs when making

judgments about the world. In short, I answer the question “Do you believe that P?” by answering the question “P?” The second topic is the self-ascription of experiential states. Again, the account lacks anything corresponding to inner perception, but rather involves reframing the mechanisms of normal experience. For example, when looking at an apple, if someone asks “Are you having a red experience?” I answer this by looking at the apple, and assessing if the apple appears red. In both cases, my access to my own psychological state proceeds by way of probing the world. In both cases, the subject needs, in addition to this capacity for indirect access, the appropriate concepts: either the concept of belief, or of experience, as the case may be. Third, Evans discusses certain illusions that arise concerning self-knowledge and its putative infallibility, and how his account can help to avoid them. Fourth, it is pointed out that the sorts of ascriptions he discusses give rise to self-knowledge that is identification-free. Finally Evans makes a few remarks about why solipsism, though untenable, gets its appeal.

The topic of this section is the self-ascription of two kinds of psychological states: beliefs, and experiences. Though there are differences in the two accounts, they are analogous, and so it might help to outline the commonality. In both cases, what is being ascribed is a psychological state of the subject, in fact a representational state — they have a content that represents the world as being a certain way. The subject has no *direct* access to these states, but instead accesses them *indirectly*, though the way they represent the world. But in both cases, this indirect access must be conceptualized. That is, it must be understood as an instance of the concept

of either *belief* or *experience* — as a type of state that any believer or experiencer (not just the subject) could be in.

The first topic, from the top of page 225 to the middle of page 226, is the self-ascription of belief states. There are two parts to this account: i) the way that one accesses one's own belief states; and ii) the conceptual resources required to go from that access to a self-ascription of a belief.

The first part here is straightforward. I access my beliefs not directly, by any sort of inward glance, but indirectly through their content. For instance, I answer the question 'Do you believe p ?' not by accessing my inner states, but by answering the question 'Is p true?'. As Evans puts it:

... in making a self-ascription of belief, one's eyes are, so to speak, or occasionally literally, directed outward -- upon the world. If someone asks me 'Do you think there is going to be a third world war?', I must attend, in answering him, to precisely the same outward phenomena as I would attend to if I were answering the question 'Will there be a third world war?' I get myself in a position to answer the question whether I believe that p by putting into operation whatever procedure I have for answering the question whether p . (VR. p. 225)

Evans is careful to say that "I get myself in a position to answer the question whether I believe that p ..." because this process is not the whole story. It is plausible that many agents (some animals? young children?) have beliefs but do not self-ascribe belief states. (Indeed, the 'theory of mind' literature investigates precisely this.) The latter is an achievement not guaranteed by the former. What else is needed is the *concept of belief*. The subject must have something like an appearance-reality distinction at hand (such a distinction of course being central to any concept of belief), and with

this the subject will have the wherewithal to understand that how the world actually is is one thing, and how subjects believe the world to be is another.

The **second topic** is self-ascriptions of experiential states, and it runs from the middle of page 226 to the top of page 231. Though Evans says that such self-ascriptions of experiential states “follow a different model” from that of belief states, the account is analogous in structure, as described above. But there are differences.

The acquiring of perceptual *beliefs* is, according to Evans, a two-stage process; first the perceptual systems create, in the subject, an *informational* state that carries information about the perceived environment. These informational states are non-conceptual.² In forming a perceptual belief — a conceptual state — one is applying concepts *to the environment*, but this application is *guided by* the non-conceptual information state. It is crucial to note that what one is *not* doing in this case is applying concepts *to the information state* itself. The point is that in the normal case what one learns about is the environment, and the informational states are unnoticed tools that participate in this process.

Nevertheless, one can learn about these information states indirectly by doing exactly what one does when forming perceptual beliefs about the environment, and then exploiting the content of these beliefs to get a handle

² Evans also says that these states have content and permit of a “non-derivative classification as true or false”. This might seem to run counter to Evans’ description of informational content in Chapter 5, where it was claimed that the ‘content’ of such states did not concern specific objects (and hence not genuinely evaluable as true or false, see e.g. Chapter 6 Appendix Section 6). The suggestion seems to be that the Chapter 5 account is correct in terms of the bare informational content carried by such states, but when such states are connected to behavior this links these states to specific objects, and hence makes them truth evaluable. As he says: “For an internal state to be so regarded, it must have appropriate connections with behaviour -- it must have a certain motive force upon the actions of the subject.” (p. 226)

on the content of the associated information state. For example when I look at a stick partially submerged in water, I receive information, and am in an informational state. On the basis of this informational state, I might (or might not, as the case may be) apply concepts such as 'bent', and reach the perceptual in that I am inclined to form the perceptual judgment *the stick is bent*. However, I might also make a judgment *about my experiential state* such as 'it seems as though there is a bent stick'.

As in the case of self-ascriptions of belief, Evans points out that this process requires that the subject have a concept of experience Evans says:

The subject who genuinely has this capacity for self-knowledge must understand the content of his judgement 'It seems to me as though p', ... (p. 228)

The point is that 'I think' (or 'it seems to me') acquires structure ('thinks' or 'it seems to ', with 'I' is the argument-place) only when in is related to (at least possible) other exemplifications of the same predicate. (p. 228, n. 41)

At the bottom of p. 228, Evans turns to a certain kind of infallibility that appears to be present in cases of self-ascription of experiential states. While some kinds of self-ascription are fallible, such as 'I seem to see eleven points of light' when in fact the subject miscounted ten points of light as eleven, others appear not to be fallible, such as 'I seem to see red'. Some might try to leverage this observation to mount an objection to Evans' proposal, and Evans responds to this objection from the middle of page 229 to the top of page 230. What Evans fails to do is clearly spell out the objection to which he is responding. The objection is this. On Evans' account, perceptual judgments involve two components: a non-conceptual informational state, and a concept-involving judgment based upon (though

not *about*) that state. The competing account is that the information state is not distinct from the exercise of the concepts used in the judgment, but rather the information state *just is* the disposition to apply the concept. If this latter theory is correct, then Evans' account, which posits a non-conceptual-content-carrying information state independent of any judgments made on its basis, would be incorrect.

The fact that the self-ascription "I seem to see red" is infallible is relevant here because the proponent of the alternate view just described claims to be able to explain that infallibility. One can't mistakenly the concept F to an information state x whose content is inconsistent with F , because x just is nothing but the disposition to apply F .

And it would seem that on Evans' account, mistakes should be possible even in the case of simple things like 'I seem to see red', exactly because the judgment and the information state on which the judgment is based are not identical.

Evans' response is two-pronged. First, he points out that the opposition position is implausible. Two reasons are given. His first consideration is that perceptual judgments do not appear to be cases in which we simply have a disposition to make a certain kind of judgment. As Evans puts it '... it is not the case that we simply find ourselves with a yen to apply some concept — a conviction that it has application in the immediate vicinity.' To see the apple as red is more than just feeling like I really should apply the concept of redness to the apple. The second is that we appear to be able to have experiential states whose content far outstrips our conceptual repertoire, and on the opposition account (which simply identifies the content of the experiential state with the disposition to apply a concept) this should be impossible. To mention an example Evans provides elsewhere, we clearly can discern more shades of color than the number of color concepts we possess.

These are considerations aimed at making the opposition view less attractive by pointing out that experience seems to have more going on than just concept application.

The second prong is aimed at showing how his own proposal accounts for the infallibility. It is that a subject can only be credited with these sorts of observational concepts (*red, round*) to the extent that the subject reliably applies those concepts to perceptual information states of that sort. A subject who made a *prima facie* error — who sincerely said “I am seeing blue” when in front of a red wall in normal conditions — would be best describe not as making an error about their perceptual state (mistakenly applying the concept BLUE when they should have applied RED) but as not understanding what ‘red’ means. So the infallibility of some psychological reports is not due to the fact that there is a fool-proof inward gaze, or a link between the world and the applicable concept (without the intermediary of an informational state) but is rather the result of the fact that the only grounds for attributing to someone mastery of an observational concept of an appropriately basic sort is that they don’t make mistakes when applying it.

Evans briefly mentions experience that is not characterized by a content (attributing a property to an object), such as ‘I feel an itch in my foot’. In such a case it might be claimed to be a stretch to say that in answering a question about the experience my gaze is pointed outward. His suggestion is that in such cases the informational state in question might have additional features over and above its information-carrying features. In the case of itching, it might have functional connections to actions, such as scratching. (Though Evans does not expand on this, it seems that there are similarities between this account and his account of spatial perception, in which

experiences are said to be imbued with spatial content if they have the appropriate connections to behavior.)

Third, Evans turns to what he calls 'certain illusions about the self'. This discussion runs from the top of page 231 to the top of page 233. The *feature* Evans is alluding to is the fact that one's perceptual and judgmental attention is always directed outward, even when making self-ascriptions. And the self, one's body or one's mind (as a repository of mental states) is typically not among the things one is attending to in these cases, because the self is often not explicitly present in the content of the beliefs (*there will be a third world war*) or perceptual experience (*behold, a house!*). The *illusion* that this gives rise to is the belief that the self does not really figure at all in the content of the subject's judgments.

Evans reply is that the *content* of self-ascriptive judgments is not limited to the deliverances of the information that is the basis of the self-ascription. Any judgment, qua thought, involves Ideas of the entities and properties judged/thought about. In a self-ascriptive judgment, one's 'I'-Idea is in play. This Idea is constitutively linked to the appropriate information channels in such a way that one has a disposition to employ it in self-ascriptive judgments when information comes in on these channels. Thus the content of the judgment outstrips the informational deliverances of the information link. When I see a tree, the content of the informational state is limited to the tree's appearance. My judgment that "I am in front of a tree" goes beyond this because it brings in additional elements in the form of Ideas whose content goes beyond what is given in the informational state. (One can imagine a transcendental argument here, whose starting point is the fact that our *judgments* have a certain kind of content, but the content of the

informational input upon which they are based lacks some of those elements.)

Fourth, Evans points out (first full paragraph on page 233) that the kinds of self-knowledge he has discussed give rise to knowledge that is identification-free. This is closely related to the previous topic about illusions of the self. It is because the information gives rise to judgments that are identification-free that they *appear* to have no grip on any object — any such *apparent* grip would amount to an identification of that object, and would leave open the possibility that the object so identified is not the self.

Finally (from the middle of page 233 to the end of the section) Evans addresses a worry to the effect that the phenomenon of identification-free knowledge needs to be understood correctly for it can be a pressure toward solipsism. Here is one way to see what is happening. The solipsist has been challenged, e.g. by Strawson, to give an account of what he means by the 'I' that is the putative center of his world (Strawson, in *Individuals*, argues that a *true* solipsist has no self-concept). The relevant solipsist is one who thinks that his self-knowledge in no way depends on his employment of a conceptual scheme according to which he is one subject among others. Challenging this solipsistic claim — that one can have a self-conception without conceiving of oneself as one subject among others — is essentially Strawson's objection to solipsism in *Individuals*. And a similar challenge is to be found in Evans' own account of 'I'-Ideas, in that part of my conception of myself as objective is that I am one subject among others.

The solipsist might respond in the following way: I have a perfectly adequate conception of myself in absence of any conception of other subjects. I am the entity such that, when it is in pain, certain experiences are

to be felt; I am the entity such that when it sees red an experience of a certain sort is to be expected; and so forth. The solipsist would then take himself to have knowledge of himself as a subject that in no way depended on his recognition of other subjects.

Evans' point here is simply that this won't work, because the solipsist needs to be able to covertly invoke the notion of the self in order to correctly delimit those experiences that the definition makes use of. The full, non-tautological version of the solipsist's definition should be: I am the entity such that when it is in pain, certain experiences are to be expected *by me*. By illicitly phrasing it in impersonal terms, the solipsist makes it seem as though a workable, non-circular definition of the self is being offered.

The discussion here is almost certainly a response of an extended discussion of solipsism and the application of the predicate '... is in pain' that occurs in Chapter 3 of Strawson's *Individuals*. It is also very closely related to the discussion in Section 7.6.

7.5 Memory

This long section consists of four main parts. First, Evans discusses what he has called cognitive dynamics (he owes the expression to Kaplan, though Evans' treatment of the topic is quite different). The core idea is that for the relevant sorts of thoughts ('that', 'here', 'now', 'I', etc.) our understanding of ourselves as located agents in a spatiotemporal world underlies a process by which a belief can persist through changes

in context. And this process accounts for how different manifestations of such beliefs are IF. Second, this is applied to 'I'-thoughts specifically, where Evans points out that because of the way memory functions, it provides us with an ability to make IF judgments about our past states. If I make a judgment 'I am F', and retain this information in memory until a later time, I can judge "I was F" in a way that would make "Someone was F, but was it I who was F?" infelicitous. Third, Evans discusses another way that memory can operate, where what is retained is not a belief state ("I am in front of a burning house") but an informational state (a visual experience of a burning house). At a later time a belief state may be formed on the basis of the retained information, and one might make an IF judgment "I was in front of a burning house". Finally, Evans discusses at length arguments to the effect that the deliverances of memory of the second sort are not immune to error through misidentification. The main line of argument is parallel to the argument in 6.6 about 'here'. Some have thought that because i) one can imagine special circumstances in which the information is in fact not about oneself, and ii) in such cases one would have to arrive at the judgment that it was oneself that the information concerned via an identity judgment, it is illicitly concluded that iii) in all cases the first-person judgment is the result of an identity judgment.

The **first** issue, running from the beginning of the section to the bottom of p. 236, is about the cognitive dynamics of *beliefs* concerning 'that', 'here', 'that' and 'I'. It sets up a framework for understanding what is common in all these cases, as a preliminary to the second topic, which is how 'I'-thoughts differ from the others.

The topic concerns how to understand the relations between co-referential belief states (I will simplify by using 'belief states' as stand in

for cognitive states generally) as time, or more generally context, changes. This came up for the case of time in the first section of the Appendix to Chapter 6 where Evans discussed what he called Dynamic Fregean thoughts. There, recall, the point was to argue that one might give expression to the same thought by saying 'Today is fine' on d_1 , and 'yesterday was fine' on d_2 (the day after d_1). The dynamics involved concerned the tracking of a time — in this case, tracking a day (d_1) as time passed and one moved, so to speak, to a subsequent day. The same topic came up in the case of demonstrative thoughts in Section 6.4.

The issue then is how to individuate belief states. And the common theme in these discussions has been that a standard way of doing it (a way consonant with, if not determined by, the different words that would be used to express the belief states) is that, e.g. the belief state 'Today is fine' expressed on d_1 is different from the belief state 'yesterday was fine' expressed on d_2 . And similarly for analogous pairs involving 'here' and 'there' as one moves, or 'that is F' and 'that was F' as time passes. As Evans has pointed out in these various contexts, this raises issues about how to understand the required identity judgments. But that isn't the current issue.

The current issue is Evans' positive proposal, a reconception of what counts as a belief state. (Note that I am here augmenting what Evans strictly says, but in a way that I think captures the gist.) On Evans' view, a belief state is a combination of i) an informational state (one that is possibly fed by temporally extended information channels) together with ii) what one might call spatiotemporal world knowledge — knowledge concerning the interplay between oneself as a located subject moving in space and time among other objects, etc. This combination yields a situation in which the judgments the subject would make in one location or time will be *systematically related* to the

judgments the subject would make at a different location or time. As Evans puts it:

If we take a belief state, as I think we should, to be a disposition to have certain thoughts or to make certain judgements, then we can say that any rational being must have a cognitive system which brings it about that the dispositions to make judgements he has at one time should be systematically dependent on the dispositions to make judgements he had at earlier times. (There being, presumably, a single persisting structural feature of the nervous system underlying both sets of dispositions.)

Persistence of belief does not always, or indeed even usually, involve persistence of a disposition to make the same judgements (if judgements are individuated simply in terms of the forms of words which would express them). For instance, the persistence of a belief that I would have manifested at some time by the judgement 'John is now angry' involves the disposition to judge later not that John is now angry but that John was then angry. (p. 236)

On this reconception, any 'different but systematically related' judgments that concern the same object, because they will be products of the same underlying complex state (as described above), will be IF.

The second topic, running from the top of p. 237 to the top of p. 239 is the application of this conception to 'I'-thoughts. The initial point (actually expressed earlier on) is that in the case of 'I', there is no *tracking* of an object over time, because there are no competing objects. In the cases of 'here'-thoughts, 'now'-thoughts, and 'that'-thoughts there are different objects/locations that can be the objects of those Ideas. And so the complex dispositional state that underlies judgments about such objects is one that must track those objects. As I change location, my 'here'-thoughts will

concern different place, 'now'-thoughts will concern different times. But my 'I'-idea is not variable in this way. This is what Evans means by this:

We might say: if a subject remembers, at time t' , being in a position at time t to assert 'I am F ', then he is in a position, without further information, to assert 'I was F (at t)'. (Hence if he is in a position at t to assert 'I am now G ', he is in a position to assert 'Something was F (at t) and is now G '.) There is no such simple rule relating memory to 'here', 'this', or 'now'. [p. 235]

It doesn't work for 'here' because remembering thinking 'it is cold here' at t_1 , together with my thinking 'it is sunny here' at t_2 , doesn't license the conclusion that some one place is both cold and sunny, because my 'here' on t_1 may have been a different location than my 'here' on t_2 . Similarly for 'this' and 'now'. But with 'I', because the Idea is never used for more than one object, the 'simple rule' works. The phenomenon Evans is pointing to is just that the 'I'-Idea is never used for different objects, and so there is no need for 'tracking' an object if one wants to maintain thought about it.

At the middle of p. 237, Evans misleadingly suggests that it is because of this special feature of 'I'-thoughts — they do not involve the tracking of an object — that they provide for judgments that are IF. But this does not seem to be correct. Evans himself (see Section 6.4) made a point of arguing that demonstrative judgments, which clearly involve tracking an object, support IF judgments provided one has kept track of the object (e.g. the infelicity of "*something* just rolled off the table, but was it *that pencil* that just rolled off the table"). And the same point clearly holds for locations and times — again, provided everything has been tracked appropriately.

Rather, the point should be that in the case of 'I'-thoughts the IF status of the relevant judgments is particularly pervasive and robust, because the

complex state underlying these various judgments is, so to speak, simpler to maintain than the counterpart states of 'here'-, 'now'- and 'that'-thoughts.

Evans' point from the bottom of p. 237 to the top of p. 238 is that this sort of IF does not depend on whether the original judgment was IF. Suppose the original judgment was identification-dependent. Suppose on Monday I saw an image in a mirror and judged "I am sunburnt". This judgment is not IF because the image in the mirror is not one of the dedicated channels for my 'I'-Idea. I could in this case wonder "Someone is sunburnt, but is it me?" perhaps because I wonder if the mirror was angled and I saw someone else. But what about my past-tense judgment on Tuesday "I *was* sunburnt (yesterday)". Here, there are two places where identity judgments are *potentially* in play. The first is the identity judgment between "I" and the person I see/saw in the mirror. The second is between "I" (on Tuesday), and "I" (on Monday). The latter is IF, even if the first is not. Or to put it another way: if on Tuesday I wonder "Someone was sunburnt, but was it I who was sunburnt?", this can only be because I am questioning, on d_2 , the identification made on d_1 , of myself on d_1 and the person I saw in the mirror on d_1 . It can't be because I am questioning, on d_2 , the identity of myself on d_2 and the person who judged, on d_1 , 'I am sunburnt.'

Evans states that this cognitive dynamics can be regarded as another part of the functional characterization of 'I'-Ideas: they not only have certain kinds of direct connections to various information channels, but they are also characterized by a kind of cognitive dynamics that, e.g., exploits information retained in memory in an identification-free way.

The **third topic** runs from the top of page 239 to the middle of page 240, where Evans introduces another way in which memory can operate, where what is retained is *information*, rather than a persisting *belief*. In the type of memory currently under discussion, the distinction concerns two

possible ways that a memory can be achieved. Either a), at t_1 the perceptual systems produce an informational state that is, also at t_1 , conceptualized to yield a belief state, and this belief state is retained in memory from t_1 to t_2 ; or b) at t_1 the perceptual systems produce an informational state, and this informational state is retained in memory from t_1 to t_2 , and then at t_2 this informational state is conceptualized so as to yield a memory-based belief state. The (b) type of case is what is currently under discussion. Evans points out that retained information is subject to a similar sort of cognitive dynamics as retained belief states, in that retained information is information that makes it seem to the subject as though something *was* experienced by the subject. And as with self-ascription of experiential states, the retained information is not information *about the psychological experience*, but information about the world.

The final topic, from the bottom third of page 240 to the end of the section concerns memory of the (b) sort, and whether it gives rise to knowledge that is identification-free. This discussion has a number of components. First, at the bottom of p. 240 and very top of p. 241, Evans claims that such cases exhibit IF. Next, he points out that Shoemaker has argued against this conclusion from considerations based on the possibility of quasi-memories (henceforth q-memories).

Shoemaker's argument is described from the top of p. 241 to the middle of p. 242: it is possible to have memories that *seem* to be memories of one's own but are not (perhaps they are someone else's memories, or perhaps they are simply false memories, the events putatively remembered happened to nobody). We can, following Parfit, call these *q-memories* — a q-memory is a psychological state that presents itself to the subject as a memory, but may or may not be an actual memory. It is of course possible that one might either suspect or know that one's own apparent memories are merely q-memories.

In such a case, it can make sense for one to wonder, on the basis of a remembered informational state 'Someone watched a tree burning, but was it I?' Thus, Shoemaker concludes, this sort of memory does not provide for judgments that are IEM.

Evans has a number of responses to this argument. First, he points out that the situation here is parallel to the case of 'here' discussed in Chapter 6. From the fact that in some situations a question can be raised such that it can make sense to think 'somewhere it is noisy, but is it *here* that is noisy?' in such a way as to require an inference in order to establish that in fact it *is* here that is noisy, it is illicitly assumed that such an inferential articulation is always present. In the current case, from the fact that a subject can suspect that she is merely q-remembering, and therefore make an explicit inference to establish herself as the real subject of the memories, it is illicitly assumed that such an inferential articulation — one that would render memory identification-dependent — is in play in the normal case as well. This line of reasoning, recall from Chapter 6, is precisely why Evans' notion of IF differs from IEM.

And just as in the case of the discussion of 'here'-thoughts, Evans not only points out that his notion of identification freedom accommodates such cases (as described above), but he also provides independent reasons for thinking that memory cannot, in general, be identification-dependent in the way suggested. Such an identity would be one of the form [I = that person], where *that person* is the person who presumably experienced the q-remembered event. Evans points out (from the bottom of p. 242 to the bottom of p. 243) that for a proponent of the Shoemaker position, there are problems with both parts of this identity statement. First, Evans claims that if a person's 'I'-Idea were such that her memory actually worked that way, then that person wouldn't in fact have a workable 'I'-Idea. The suggestion is that

having genuinely IF memory is required in order to mobilize the sort of cognitive mechanisms Evans discussed previously that allow the subject to conceive of herself objectively, to have an adequate Idea of herself. The full paragraph on p. 245 picks up on this point, by showing that while in some circumstances we can suppose a subject, appraised of the situation, might be able to use her apparent memories to get information about another subject, that this cannot be the default way for memory to work.

He goes on to argue (from the bottom of p. 243 to the top of p. 245) that the 'that person' side of the identity is also dicey. Not only would we not have an adequate Idea of any such person (Evans here appeals to his dictum that information-links are not sufficient), the judgment in question — 'someone was *F*' or 'that person was *F*' — would not constitute knowledge, for if one's memories are subject to such interference, the inference from their deliverance of *F* being instantiated to *someone* actually having been *F* would be a shot in the dark if true; and false if false; and in neither case sufficient to support a knowledge claim.

From the middle of p. 246 to the end of the section, Evans discusses why many resist his conclusions. The text is structured in such a way as to give the impression that there are two different points here — one about Williams' remarks on imagination, and another about a fallacy concerning the nature of q-memory — but I think it is more revealing to treat them together. The fallacy is to move from the conceptual possibility of q-memory, and even of a subject's coming to believe that her apparent memories are q-memories, to the conceptual possibility of a kind of memory that presents contents to the subject in such a way that leaves the identity of the subject open. Call this q*-memory (the expression is mine, not Evans'). We now have three kinds of 'memory':

memory: The process as normally conceived. Memories are of events that one in fact experienced oneself, and they are presented to the subject as such.

q-memory: Are of events that may or may not have been experienced by the subject. But are presented to the subject as though the subject did experience them herself.

q*-memory: Are of events that may or may not have been experienced by the subject. And are presented to the subject in an impersonal way, a way that does not suggest that the subject herself had the experience.

The slide is easy enough to describe: a subject who realizes that she is merely q-remembering and has only *apparent* memories of things that (possibly) happened to someone else will realize that she cannot judge that she was F on the basis of her q-memories. Any self-judgment she might make will need to be the result of an inference such as

1. Someone (*a*) was F
 2. I am *a*
- Therefore, 3. I was F.

From this, a philosopher illicitly assumes that the sort of content described in premise 1 captures the content of q-memory. That is, q-memory is really q*-memory.

It is here that Williams' observation is relevant. That point is that to imagine *being in the West Indies* — or maybe more perspicuously for present purposes, to imagine 'what it is like for someone to be in the West Indies' — is not the same as to imagine *oneself* being in the West Indies. That is, there is a difference between an episode of imagining that is framed in first-person terms, and one that is framed impersonally. (Similarly, there seems to be a difference between i) imagining what it is for *a location* (in general) to be loud, and ii) imagining that it is loud *here*.

The suggestion is that the difference between these two is that the first-personal imaginings are a composite of the impersonal imaginings plus some surplus, expressible as a premise or inference to the effect that the *one* that is imagined to be in the West Indies is *oneself*.

Evans' reply is to the effect that Williams' observation to the effect that to imagine being F is not the same as to imagine oneself being F is correct, but that this does not support the conclusion about memory that Evans' opponent wishes to draw. Evans is surely correct to point out that this gets things exactly backwards — the normal operation of memory provides states that present to the subject experiences in a first-personal way, and it is only in unusual circumstances that one will engage in an inference that overrides that element of the content. That is, it is not that memory is a composite of q*-memory plus an inference involving an identity claim. Rather, i) q*-memory isn't a thing; and ii) q-memory is a composite of normal memory plus an inference to the effect that the apparent first-personal nature of its content might be inaccurate.

7.6 The possibility of reference failure

In this section, Evans discusses a possibility that his views on reference and 'I'-thoughts provide: that a subject's 'I' thoughts may lack a referent, not because there is no subject, but because the subject may be unable to think about herself if she lacks discriminating knowledge. Two kinds of case are discussed. The first are cases where the controlling conception of oneself is being fed information from an object that is not oneself. In this case there is not a single object that the information is coming from, and hence the attempted self-thoughts are ill-grounded. The second sort of case is the brain in a vat. In this situation, the problem is that the subject has no resources to identify herself as an entity in the objective order.

Even in a book full of novel, fascinating and counter-traditional positions, this section stands out. Having rejected the “no ownership/'I' does not refer” view earlier in the chapter, Evans now rejects its main rival, the Cartesian view to the effect that 'I'-thoughts have a guaranteed of successful reference, to the subject's own mind. Note that the sort of “reference failure” possibility being discussed here is not a situation where there is *no referent, no subject*. Rather, it is assumed that there is a subject. The question is whether the subject is able to refer to herself, and the claim is that in some cases a subject might lack the ability to refer to herself.

This discussion is intimately related to the discussion of solipsism in Section 7.4. There, recall, Evans argued against the attempt of the solipsist to identify oneself along the lines of “I am that object such that when it is in pain something frightful is to be expected” (p. 234). Evans diagnosed the problem there as a sort of circularity — it only works if one is illicitly

supplementing the definition with a self-specification that builds in the capacity for self-discrimination it is trying to explain. Something along the lines of “I am that object such that when it is in pain something frightful is to be expected *by me*.”

Here Evans discusses the requirements for a subject's having an adequate 'I'-Idea. The upshot will be that a capacity for successful self-reference is not automatic. The first sort of case he discusses in which a subject's 'I'-thought might lack a referent involves elements such as q-memories and information links that are configured in deviant ways. This discussion runs from the beginning of the section at the top of page 249 to the top third of page 250. The point here is that one's 'I'-Idea is one that has as its target a single object from which the various information channels and memories derive. This is similar to 'here' thoughts and 'this' thoughts. In these cases — the unknown substitution of one object for another in what is *taken* to be a temporally extended demonstrative thought, or unknown movement during an attempted series of 'here' thoughts — there is the possibility that there is no single object and hence no thought. Similarly Evans here argues that a multiplicity of objects from which the information in one's controlling conception of oneself derives (or via which one's intentions to act are manifested — the other side of the functional characterization of 'I'-Ideas) can render one's 'I'-thoughts ill-grounded, and without an object.

Evans points out that in such cases it is nevertheless true that if one is appraised of the situation, one might (depending on specifics) be able retreat to a more limited self-conception in 'I'-thinking, one in which one abandoned the attempt to identify oneself prior to the onset of the deviant processes, and restrict one's self-directed thought to times when the thoughts would be well-grounded. This might work if one learned which information

channels were suspect (one could then purge the information from those channels), or if there was some time before or after which memories became suspect (in which case one could purge the suspect memories). It would be possible, in this way, to have a sort of modified but still legitimate 'I'-Idea.

The second kind of case, whose discussion runs from the middle of page 250 to the end of the section at the middle of page 255, involves the brain in a vat. Initially we consider a case where the brain was never actually embodied, but was always in a vat. Here Evans says that the subject will lack resources to identify herself, and hence lack any adequate 'I'-Idea, and hence be unable to think of herself. And in such a case, there isn't even the possibility of the sort of limited fall-back option there would be in the situation described in the prior paragraph.

This might initially seem to be obviously false, for in the standard brain-in-a-vat case the brain is connected to inputs and outputs in such a way as to exactly mimic the inputs and outputs a normal brain would have, and also such as to yield a subject of thought that is fooled into thinking that it has a body, is moving around in a vast world, etc., when in fact none of this is true. And surely it might be thought that in such a case the subject is capable of thinking 'I'-thoughts, and had in fact thought quite a number of them, assuming that the brain had been around for a while (as was the case for most humans in *The Matrix*, for example).

Evans argues that in fact the subject cannot successfully think any 'I'-thoughts, for the subject has no adequate Idea of herself such as would be required for such a thought. When appraised of her situation, the subject would have to realize that the entire universe of her experience has been a sham, that her body, including what she took to be her brain, is illusory. The one thing on which her 'I'-thoughts clung — her body/brain qua

spatiotemporally extended objective entity in the world — has been revealed to be non-existent, along with what she took to be the objective world.

Evans discusses the possibility that the subject will attempt to think of herself as some brain somewhere. He says:

The subject is to think 'Somewhere in the world there is a small parcel of grey matter, wrinkled, moist, and soft, about three inches high, and that is me.'

...

But obviously the physical side of the conception of himself which our unfortunate subject possesses does not encourage any identification of himself with a brain; and anyway, we are now considering a case in which such elements are extruded from the subject's conception of himself, being, as he now realizes, without any foundation in fact. So the identification is, so to speak, wholly theoretical, and it remains quite obscure what mode of thinking about himself renders it even thinkable. [p. 251]

The idea here is that a subject in such a position, if appraised of her situation and if she is thinking it through clearly, will realize that the very idea that there are people with biological brains in their bodies that in some sense support their thoughts is all part of the charade. And so the supposition that this is how things are 'in reality' is a guess. It might be the case that such general elements of the charade (people's thoughts are sustained by brain-like things) are modeled on facts of the reality in which the brain and its computer interfaces are maintained. But it might not.³

³ A decent chunk of people's intuitions about what is going on in the Matrix is subtly infected by the fact that as the film depicts things, there are significant similarities between actual reality and the virtual reality of the matrix itself. The bodies people have are similar, the fact that there are brains in heads that appear to be the seat of thought is similar. Even the rough geography is similar (the year is different, as are a variety of specifics). This no doubt makes things easier from a narrative standpoint, but an actual subject in such a situation could not presume that such similarities obtain.

So the thinker lacks the capacity to identify herself as a being identical to any objective thing, like a body or brain. The next possibility is that the subject might try think of herself via description, as 'the thinker of this thought' (where the 'this thought' might either reflexively refer to that thought itself, or to some distinct 'demonstratively identified' thought). Whether this will work depends on whether one has an adequate Idea of the supposed anchoring thought itself. Does the subject here have the ability to distinguish that thought from all others?

Notice first that the necessary sort of distinguishing is, so to speak, distinguishing a *token*, not a *type*. I can of course tell that a given thought is a *dogs are mammals* thought, as opposed to a *triangles have three sides* thought. But that sort of distinguishing won't do the job, since many people might be thinking that dogs are mammals, and so trying to identify myself as *the person who is thinking that dogs are mammals* won't single me out. Rather, what 'this thought' has to mean is not this *type* of thought (or thought with *this content*) but rather *this (token) thought*.

The issue is that typically we distinguish tokens of mental states by reference to the person having the state. If I know that there are two mental states of the same type (headaches, say), occurring in an auditorium, I identify the tokens by reference to the people: one is Mary's headache, and the other is Rudy's headache. If you asked me to identify just one of two qualitatively identical headaches, but without referring to the people who have the headaches, it is not clear what I could do.

Given this, the attempted descriptive identification is circular. It is trying to anchor an identification of a subject via an identification of a

And so guesses about there being a brain somewhere would be just that, mere guesses.

thought, but fails to notice that the only anchor available for identifying (token) thoughts is the subjects who have them.⁴ It is analogous to this:

A: Do you know who you are?

B: Yes, I am the person at location *L*.

A: And where is location *L*?

B: *L* is where I am.

Evans goes on to claim that in other kinds of case, such as paralyzed subjects, or subjects that were normally embodied up to a point and then had their brains removed, residual dispositions might be sufficient to support 'I'-thoughts.

7.7 Conclusions

In this section, Evans summarizes some key points. The section is short enough that a separate overview is unnecessary.

This concluding section is broken down into four numbered points. The first is that our 'I'-Ideas constitutively involve functional connections to sensory input and action output in such a way as to render any attempted reduction of 'I' to a description unworkable, because no description can

⁴ Evans' discussion here owes much to the early chapters of Strawson's *Individuals*, and Evans' discussion here is a reflection of the extent to which a good deal of Strawson's thinking infuses the DNA of Evans' program. My point is that anyone who wants to get a better idea of what is driving Evans' intuitions here would be well-advised to read *Individuals*.

guarantee the correct functional connections. In addition to these functional considerations, a subject's 'I'-Idea, in order to be adequate, must be such that the subject has the capacity to identify herself as an entity in an objectively conceived world. This line of thought will be explored further in the Appendix.

The second point is a recap of the discussion in Section 7.1 concerning the view, which Evans opposes, to the effect that our self-identification involves a subjectivity that is impossible to reconcile with the objective order. He points out that our self-conception is a conception of an objectively conceived subject.

The third point concerns the fact that, on Evans' account, our 'I'-Ideas span the physical and mental. They are connected to ways of gaining knowledge that are identification-free.

The fourth point concerns the dependence that our 'I'-Ideas have on presuppositions — such as a continuing bodily referent. Because of this, the notion of the mind as a independent entity, cognitively self-sufficient in absence of a body or knowable worldly location, is mistaken.

7.A1 Appendix Section 1

In the main body of Chapter 7, Evans discussed the ability to self-ascribe belief states and experiential states. This was by way of trying to explain our self-concept and self-knowledge — we conceive of ourselves as believers and experiencers. This section of the appendix is a start on explaining another aspect of our self-conception, our reflective self-consciousness or self-awareness.

This section of the Appendix is one of the least understood parts of the book (which is saying a lot). Partly this is due to the fact that it is not as clearly written as it could have been. Accordingly, I will spend a little more space on this than I otherwise might.

Evans' main goal in this chapter is to explicate our 'I'-Ideas, and part of this has been to explicate our self-understanding. A passage from the first paragraph of the chapter is crucial for framing this:

... there can be no complete understanding of self-identification without an understanding of the self-ascription of mental predicates; and no adequate understanding of the self-ascription of mental predicates without an account of the significance of those predicates... (p. 205)

Two applications of this have been our self-ascription of beliefs, and self-ascription of experiential states, as part of an account of our understanding of ourselves as believers and experiencers. In both of those cases, there have been two components to Evans' account, what might be called a *source* component and an *understanding/conceptualization* component. The source in both cases was an ability to re-purpose capacities whose primary home is learning about the world in order to get at, indirectly, the psychological states. For example, I learn that *I believe that P* by asking whether *P*. The understanding/conceptualization component in both cases had to do with our concept of the ascribed state. A self-ascriber needs to have a concept of belief, and a concept of experience. This is what Evans means by "an account of the significance of those predicates."

This section of the Appendix is in the same spirit. Though the self-ascription in question is not belief or experience, but *self-reference*. Evans introduces this topic with some remarks on Anscombe's paper, but this is just to introduce the topic. Evans' goal is not to engage with any of Anscombe's substantive claims specifically. (And it should be pointed out that Evans' program in this chapter owes more to Anscombe's paper than he acknowledges.)

As Evans reconstructs Anscombe, her challenge is to understand what it means to use "I" to refer to oneself. (Following Castaneda and Anscombe, an asterisk on a reflexive pronoun indicates the relevant self-awareness — Oedipus was thinking of *himself* when he intended to kill the slayer of Laius, but he was not thinking of *himself**.) The worry is that either in using "I" one refers to oneself in the sense of "X refers to X" (*Alice refers to Alice*) or in the sense of "X refers to herself" (*Alice refers to herself**). The first sense won't work, since *that* sense applies to the case where Oedipus referred to himself in thinking that *the slayer of Laius is F*. And the second formulation, while correct, makes use of the very notion, referring to oneself*, that we are trying to understand.

Another suggestion discussed by Anscombe is that the first person pronouns is what one uses to *knowingly and intentionally* refer to oneself. The question now is in understanding the content of that intention. Again, it seems there are two ways to fill this out. First, one could fill it out third-personally: *X refers knowingly and intentionally to X*. This doesn't work, since Oedipus did knowingly and intentionally refer to himself. So this addition doesn't help.

The goal then is to say something substantive about self*-reference. Evans proposal is one that, essentially, splits this work up into two parts,

neither of which presupposes self*-reference, but which jointly get us at least some way in the direction of understanding self*-reference.

First, let's take a quick detour though the locution *S intends to A*. This is a common way to express propositional attitudes. In the following examples, the subordinate clause that expresses the proposition that is the object of the attitude is underlined:

1. S [attitude] CLAUSE

For instance

2. I believe the cat is on the mat.
3. John wants everyone to be happy.
4. Oedipus intends Laius to be killed.

In English the subject of the subordinate clause is often suppressed, especially when context or grammatical clues make it clear what the referent of the suppressed subject argument is. This suppressed element is often called PRO — essentially a non-overt pronoun (the doctrine in generative linguistics is that it is represented grammatically even though it is not expressed phonologically). The antecedent of this pronoun can be indicated by an index. So in cases where there seems to be no subject for the subordinate clause as in (5) and (6), one might recognize a non-overt pronoun, as in (7) and (8):

5. Alice wants to be happy.
6. The pilot intends to be killed.
7. Alice_i wants PRO_i to be happy. (Compare this to (3))

8. The pilot_j intends PRO_j to be killed. (Compare to (4))

The first thing to note is that in cases where PRO is coreferential with the matrix subject (as in (7)), the sense of PRO is essentially self*-directed. The sentence “Alice_i wants PRO_i to be happy” cannot be interpreted to mean that *Alice wants someone (who she doesn't realize is herself) to be happy*. So if you and your friend Alice see a news report about a forest fire, and she says “I want the person who started the fire to get caught,” and you know that she unwittingly started the fire, you cannot accurately (non-mischievously) report this by saying “Alice wants to get caught.”

One might worry about the appearance of self*-directedness here, but the idea is that *this* sort of self-directedness — between the person indicated as the subject of the matrix clause and the referent of PRO — is not one that is appealing to the sort of self*-reference we are trying to understand. This sort of self*-involvement is often used precisely in cases where, as Evans' analysis would describe it, the functional *action* component of one's “I”-Idea is in play. For example, when I see a bear charging, one could say “Rick wants to flee”, and in such a case we would have:

9. Rick_n wants PRO_n* to flee.

But note that we could equally correctly say the same thing of a rabbit (“The rabbit_m wants PRO_m to flee”), without attributing to the rabbit any conceptual understanding of self*-reference. It is just an expression of the special action-oriented component of ‘I’-Ideas, and it does not depend on a capacity for self*reference.⁵

⁵ On this analysis *of the grammar*, the object of my attitude is a proposition that has a subject. But this shouldn't be taken to indicate that there is an intermediate identity involved, of the sort that would render such ascriptions

Let's apply this to the original suggestion. The intention we wanted to understand (with minor grammatical adjustments) was:

10. X_a intends PRO^*_a to refer to X_a .

The problem was how the *object* X of the subordinate clause is specified: (10) either fails to capture the intention (if we don't put an asterisk on the last 'X'), or it is circular (if we do). The challenge, accordingly, is to spell out the subordinate clause in a way that doesn't face this dilemma. But note the problem is *not* with the PRO^* subject. That is just expressing X 's ability to engage in IF actions.

Evans' proposal is that we can explicate the intention in this way: S *intends to satisfy* " ξ refers to ξ ", or as it could be spelled out:

11. S_a intends PRO^*_a satisfy ' ξ refers* to ξ '

And to reiterate the self*-directedness in PRO^*_a is a manifestation of the 'action' component of 'I'-Ideas. It does not presuppose self*-reference in anything like the sense we are trying to elucidate, and hence not introducing an objectionable circularity.

Now let's turn to the one-place concept expression ' ξ refers* to ξ '. The first thing to note is that the sort of reference required here is self*-reference, which is why I am including an asterisk. What S must intend, of ξ , is *not* that ξ satisfy ' ξ refers to ξ ' in such a way that Oedipus referring to Oedipus (in thinking *the slayer of Laius is F*) counts. The concept expression

identification-dependent. The rabbit is not wanting *something* to flee, and then inferring that the thing it wants to flee is *itself* on the basis of an identity judgment. The connection is functional and non-mediated.

is one that involves self*-reference. S is intending that ξ refer to itself as itself.

This realization might make one worry that the account is just as circular as the simpler version Evans derided earlier, but that he just tried to hide the circularity in a lower level of argument structure. But this would be to miss the essential point. The sort of self*-reference specified in ' ξ refers* to ξ ' is an *objective* understanding. What is required is that the subject *S* understand what it means for a *thinker* to refer to *itself**. A subject could conceivably understand this without necessarily having to understand what it means for *her** to refer to *herself**. Imagine here a being that thinks thoughts about the world, and about other subjects in the world. This being understands what it means for subjects in the world to refer to themselves self*-consciously, but does not grasp the idea that it is, *itself*, a thinker capable of such self*-reference. Call such a being the *oblivious observer* — oblivious in the sense that it does not realize that it itself is an agent/subject like the ones it is observing. The oblivious observer, if it met Oedipus, might very well intend that Oedipus satisfy " ξ refers* to ξ ", and might try to bring this state of affairs about by saying, to Oedipus, "Look, *you* killed Laius, *you're* the one who is *F!*" (There is nothing especially outlandish about the notion of an oblivious observer. A primatologist studying chimps might conceptualize many chimp activities — physical or psychological — using concepts that apply to herself even though she is unaware that they do.)

In any case, that the understanding of self*-reference involved here is 'objective' in this sense is what Evans is getting at when he says:

... it does not follow that in order to elucidate the intention of satisfying ' ξ refers to ' ξ , we need a grasp of the self-conscious Idea-type that we have of ourselves. (*VR* p. 258)

It is now clear how this discussion is parallel to the discussion of self-attribution of belief and experiential states. In both of those cases, part of what was required was an objective understanding of the relevant predicate, a *concept*. One had to understand what beliefs are, that is, to have an objective understanding of what it means for “X believes that P” to be true.

Now we can put the pieces together. The challenge was to explicate the content of the intention “knowingly and intentionally refer to oneself” in a way that was substantive but not circular. The claim is that we’ve now at least made some headway. In (11) there are two places that self*-ish-ness comes into play. First, it is present in the action-component of the intention. This is the part captured by the asterisk on the PRO_a^* . But as we’ve seen, this does not sneak in self*-reference in the sense currently being examined.

The second place it occurs is in the sort of reference specified by the concept expression. This also does not introduce circularity, since it brings in an objective, so to speak, understanding of self*-reference.

So imagine again the oblivious observer, capable of understanding thinking subjects and grasping the thought that subjects can self*-refer. This being might, then, at some point get an inkling that it is itself a thinker, like the thinkers it has been contemplating. Part of this realization might involve an intention to satisfy the very concept expression it previously considered only in relation to other subjects. This is what Evans means when he continues the passage quoted above by saying:

Indeed, it seems plausible that the explanatory direction goes the other way: the fully self-conscious use of ‘I’ can be partly explained, precisely, as a use in which the subject knowingly and intentionally refers to himself (satisfies $\lambda x(x \text{ refers to } x)$). (VR pp. 258-9)

The reason that I described this as 'making headway' on the challenge is that Evans doesn't think that this is quite enough. It is an advance, but not a full solution, since it is little more than a gesture at what that objective conception of self*-reference is. Evans does some perfunctory exploration of what the content of that could potentially be, starting around the middle of p. 259 to the middle of p. 260.

Evans then turns to some consequences of the 'self-reference principle'. The first is that it explains why there has been a tendency to equate the *self* with the *mind*:

It seems plausible that ... the propensity to think of the self primarily as a thinker ... come[s] about from the application of the self-reference principle. We enquire about ourselves as the objects of 'I'-thoughts. (Our question is 'What is the object of an "I"-thought?') And the self-reference principle tells us that the object of an 'I'-thought is its subject. So the identity of the self and the thinker becomes established as indubitable, simply because we enquired into the nature of the self by asking about the object of certain thoughts. (The identity of the self and the speaker might have assumed an equally -- and equally misleading -- central position in an enquiry which focused on the reference of the pronoun 'I'.) (VR p. 260-1)

The self-reference principle is *consistent with* the suggestion that the self is separate from the body (the body is the self's physical organ); but it is not consistent with the contrary suggestion that the self is separate from the mind, that the mind is the self's thinking organ, because then the self-reference principle would not apply: the object of the self*-thought would not be identical to the thinker of that thought. Evans does not think that the self is just a mind, but he admits that that fact cannot be established by the self-reference principle.

7.A2 Appendix Section 2

Here Evans is invoking Perry ('The essential indexical') and Lewis ('Attitudes de dicto and de se') as well as others. And it is plausible that Evans is here thinking of considerations similar to those raised in this Chapter in relation to Nagel's argument about the gulf between the objective and the subjective. Rather, learning starts with the functional component, and then learning to approximate a more objective understanding of the relation between our thought and the (objective) world.

7.A2 Appendix Section 2

In this section, Evans explores aspects of various kinds of Ideas he has discussed. In particular, Evans discusses the relation between the 'functional' aspect of such Ideas and the 'background' component (which is the component that accounts for the subject's capacity to know which object is in question). The upshot is that a full account of these Ideas requires attention to both aspects, and the two aspects are not fully separable, despite the fact that his discussion of them has treated them as separable. In particular, Evans suggests that the functional component must be seen as providing at least some of the materials for discriminating knowledge.

Evans in this section summarizes the main features of his account of Ideas. His account of 'here'-, 'that'- and 'I'-Ideas have two aspects. The first is

what he has called the *functional component* of these Idea types he has discussed, which has to do with the information links between these Idea types and relevant avenues of informational intake and behavioral expression. This component would, by itself, be of limited use. As Evans puts it:

... the evidential component of the functional characterization of a demonstrative Idea of a material object based on a visual information-link with it would deal only with judgements involving concepts whose application can be determined on a visual basis. (VR p. 262)

The second component is given a number of descriptions by Evans, and also a number of jobs. It is linked to our fundamental Ideas of the relevant objects; it accounts for our understanding of the referents as objective; it provides for the satisfaction of Russell's Principle; and ultimately it accounts for the fact that our thoughts satisfy the Generality Constraint.

Evans has argued at many points, directly and indirectly, that both components are necessary. And he has also argued at a few points that they are not only both necessary, but that neither can be reduced to the other, and he reiterates this point at the bottom of p. 262.

Here Evans is invoking Perry ('The essential indexical') and Lewis ('Attitudes de dicto and de se') as well as others. And it is plausible that Evans is here thinking of considerations similar to those raised in this Chapter in relation to Nagel's argument about the gulf between the objective and the subjective. Rather, learning starts with the functional component, and then learning to approximate a more objective understanding of the relation between our thought and the (objective) world.

But now he takes on another issue directly (though he has hinted at it before, e.g. the end of Section 4.3) which is the relations between the two aspects. The first point is that the “general conception” — the second aspect, responsible for our *understanding* of the Ideas and concepts involved and the satisfaction of Russell’s Principle — is arrived at in part from training based on the first component.

In addition, he is here suggesting that at least some of the resources that had been located in the second level have their source in the functional aspects. As Evans puts it:

It is essential to see that the material of the functional characterization -- which connects the subject, via a complex battery of dispositions, to a specific object in the world which is disposed to affect his thinking and to be affected by his actions -- can enter into the account of how the subject satisfies the requirement that he know which object is at issue. Otherwise we would be liable to an unnecessary worry over whether our singular thoughts are adequately pinned down to a particular set of objects, when we contemplate the possibility of Twin Earth cases. (P. 263)

The connection between the levels solves what would otherwise be a problem. Here Evans puts the problem in terms of Twinearth and massive reduplication (pinning our thoughts down to a particular set of objects). My twin and I will have, let us suppose, identical *objective* conceptions of the universe, exactly because this conception is by definition bereft of the one sort of element that could distinguish the two: the kinds of functional connections made available by the first component of our Ideas. contextual elements. (One could also express it in terms of the right functional connections not being deducible from the general conception.)

So in one way or another, a contextual element is required as part of the apparatus that satisfies RP. And materials such as this are infused in these Ideas. The suggestion is vague, though. It is not clear what form this 'infusion' is supposed to be taking. Though subsequent sections of the appendix are on roughly the same topic.

7.A3 Appendix Section 3

In this section, Evans raises a question for his account, but does not answer it. The question concerns the role of the fundamental level of thought. Evans claims its point was to explain objectivity, and so even if someone thinks it is unnecessarily complicated, something doing that job will have to be present. But he raises a worry that, regardless of whether one appeals to a fundamental level of thought, or something less elaborate, it is not obvious how we are to account for this objectivity. On his own account, the cognitive map is still built up from landmarks that are encountered by the individual, and hence are saturated with egocentricity.

Evans begins by pointing out that while his apparatus of the fundamental level of thought might strike some readers as unnecessarily baroque, the job that it is intended to do is necessary. Evans says the job it is intended to do is to explain what it is in virtue of which a subject's thought concerns an objective world. And since our thought clearly is objective in this way, any adequate account of thought will have to have something playing this role, even if it is not exactly of the form of Evans' proposal concerning the fundamental level of thought. (The photograph model lacks any such

apparatus, and throughout the book this failure has been what Evans has harped on, in one way or another.)

...even if the idea of fundamental identifications had been dropped, the overall picture of 'this'-thinking, 'I'-thinking, and 'here'-thinking would have been similar in an important respect: the role played, in the book as it stands, by the notion of the fundamental level of thought would have been played instead by the notion of the objective or impersonal conception of the world, with mastery of such thinking dependent on an understanding of how it relates to the world as objectively conceived. (P. 264)

Evans then points out that how this is supposed to work, even on his own account, is not obvious. The problem is, roughly, that it is not at all clear how a subject can bootstrap herself up to fully objective conception of anything. The example Evans gives here is his account of objective space in 6.3, according to which one's cognitive map is given by a representation of the world anchored to recognizable landmarks that holistically constitute the frame of reference. But these landmarks are necessarily landmarks that the subject must be able to encounter in egocentric space. Given this, the worry is that "the seemingly objective mode of thinking about space is, after all, contaminated by egocentricity."

This is a question that Evans does not have an answer to, but it is one that he has asked before. In 4.3 he says of the generality constraint that it is perhaps an ideal to which our actual mode of thought only approximates; in Part 2 of this Appendix, he suggests that the 'objective' aspect of an Idea has some elements of the ground level functional part built into it in one way or another.

Note that when the apparatus of the fundamental level of thought was introduced, it was not introduced in terms of underwriting *objectivity*, but

rather in terms of underwriting *discriminating knowledge*. The thought, at least for physical objects and locations, seems to have been that in order to discriminate an object in a way that allowed for satisfaction of Russell's Principle, would require an objective conception of space and time, and an ability to discern which object, of all objects represented in this objective conception, was the target of the thought. Here Evans is worrying about where this objective conception can possibly come from. If it is built up from egocentrically presented objects and locations, then it is not fully objective.

I will return to this discussion when we get to the Appendix of Chapter 8, since a similar worry is discussed there.

7.A4 Appendix Section 4

This section is brief enough that a separate overview isn't needed.

This very brief section, in which McDowell puts Evans in the third person, makes some remarks on some central ways in which Evans would likely have expanded his discussion of 'I'-ideas had he been able. It continues the theme of the last few sections of this appendix, namely that there is a puzzle regarding the identification of the self as manifest in the 'I'-Idea proper and the self as an element of the objective order. This discussion — in this section of the Appendix and the previous few sections — is grappling with worries of the sort raised in Section 7.1 with respect to the Nagel argument. This section consists mainly of quotes from philosophers that are getting at the issue that Evans is worried about. The vague suggestion seems

to be that there is something about the exact nature of our embodiment, and the functional component of our Ideas, that anchors them in an objective order more directly than merely supporting a possible identity judgment between the entity with the right functional connections and an objective entity. But the section is little more than a glance in the direction of the worry.⁶

⁶ McDowell himself has an extremely helpful and fascinating discussion of the import of the action element of our self-conception, in a way that might very well have been at least roughly along the lines Evans would have pursued, in his 'Referring to oneself' (In Hahn, Ed. *The Philosophy of P.F. Strawson*).