

# 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 What Happens in Chapter Seven

*In this chapter Evans discusses 'I'-ideas. They are similar to 'here'-Ideas and 'that'-Ideas in that they pivot on functional connections to perception and action. But they differ in key respects. First, unlike 'that'-Ideas, the connections can be latent. And second, unlike 'here'- or 'that'-Ideas, there is only one object tracked as 'I', which makes the cognitive workload easier, but introduces issues concerning identity through time. Also like these other modes of thinking, 'I'-thinking is IF. There are several primary topics. First, 'I' does refer. Some arguments to the contrary have concluded that it does not because no object is identified, but Evans points out that this is a consequence of their IF status. The information doesn't identify the subject, but it is identified by the Idea that the information is linked to. Second, memory supports IF 'I'-thinking, despite some arguments to the contrary. These contrary arguments move from the fact that in some cases (where q-memories are involved) the identity can be questioned "Someone was captured by the KGB, but was it me who was captured?" to the conclusion that in the normal case they are based on an identity. Third, Evans points out that while some have felt that IEM applies only to psychological predicates, Evans shows that in fact there is no asymmetry between physical and psychological predicates with respect to their potential identification-freedom. Bodily self-ascription includes not only predicates about one's physical body, but also about one's location. For example, I come to know that I am next to a tree based entirely on IF visual experience of a tree ("Someone is standing before a tree, but is it me who is standing before a tree?"). And this is crucial for one's ability to locate oneself in space, and hence the adequacy of one's I-Idea. The ascription of psychological predicates, such as beliefs and experiences, is also IF.*

## 7.1 Introductory

*This initial section covers four 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. Second, he argues that 'I' is a referring expression, contrary to a number of prominent views. The opposition view is that since misidentification is not possible, this suggests that nothing is identified to begin with. Evans points out that this is not the case, it just means that the relevant information link is constitutively and functionally tied to the relevant Idea. Third, Evans argues that self-identification is not a matter of some elusive pure subjectivity nor is it even exclusively a matter of the mental. Fourth, Evans issues some cautions about our understanding of the persistence of the subject's identity through time.*

From the top of page 205 to the bottom of page 207, Evans discusses some initial similarities between 'I'-thoughts and other kinds of thoughts discussed already, such as 'here'-thoughts, in order to motivate the treatment of self-identification (he also notes some differences). The main kind of similarity is in the form of some apparent functional isomorphisms between 'I'-thoughts and the thoughts discussed in Chapter Six. For instance in the case of 'here' the subject's 'here'-thinking involves an Idea that is constitutively tied to certain ways of gaining information, and certain ways of acting on the subject's part — getting information along an appropriate channel to the effect that '... is hot' is instantiated serves for the subject to have information to the effect that 'here it is hot'. And there is a parallel action component: being *too* hot here must, *ceteris paribus*, dispose me to leave the vicinity.

Similarly, the kind of self-thinking we are interested in is thinking that has analogous functional connections to perception and action. My thought about 'the tallest member of the UCSD philosophy department' need not have any special connections to my perceptual input or action; but my 'I' Idea must have such connections. There are information channels such that getting information to the effect that '... is facing a tree' is instantiated just is for me to have information to the effect that *I* am facing a tree. 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 knowledge that the tallest member of the UCSD philosophy department is about to be attacked by a bear.

In short, I-Ideas are constituted in part by functional connections to specific ways of gaining information and dispositions for action. In this way they are similar to here-Ideas. And also because of this, we should expect that like demonstrative Ideas and here-Ideas they will exhibit identification-freedom.

Though there do appear to be similarities of these functional sorts between 'I'-Ideas and other kinds of Ideas, Evans points out that 'I'-Ideas have features that are novel. The first element concerns the role that our memories play in our 'I'-Ideas, and this will be the topic of 7.5. The second involves the fact that our 'I'-Ideas essentially involve a kind of self-reference. (This will be discussed in Section 1 of the Appendix.)

From the top of page 208 to the middle of 210 Evans discusses and rejects the view that 'I', while grammatically a noun, is empty. The view is that the verbal expression 'I' has a role in the language, and this is mistakenly taken to imply that it has a referent, or that it is associated with some kind of content. This view was maintained (in slightly different forms), by Geach, Strawson, Wittgenstein, and Anscombe. Evans argues that this is mistaken, that there is substantive content of, and indeed a referent corresponding to, our 'I'-Ideas. 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. The judgments might as well be expressed as *What an awful muddle*, or *There is pain*.

The premise in both cases is that there seems nothing in such first-person judgments that corresponds to getting myself to think of the right object (as there is in communicative situations with others that might require the use of 'I'). But the conclusion, that nothing corresponding to a substantive 'I'-Idea is involved in such judgments, does not follow. Evans has two connected lines of thought here, an objective line and a subjective line.

The subjective line is that at best the premise (there is no question of getting myself to think of the right object) shows that I-thoughts are identification-free, in the sense discussed in Chapter 6. The objective side 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 through 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 lines 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  $\delta$  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, it's not clear what could be made of my knowledge that the same person is involved in my judgments that *I am in an awful muddle* and *I was dragged through the streets of Chicago*.

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 unbridgable 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 that content a grasp of which would constitute knowledge to the effect 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' reply is that while it is true that no difference will be manifest in how things are objectively conceived, there will be great differences in how 'the subject's immediate environment' is conceived. 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. Evans admits that such knowledge can only be expressed via indexical expressions, but says that this is no objection to treating it as knowledge.<sup>1</sup> (At the very least, it puts the burden of proof back on Nagel's shoulders.)

At the top of p. 212 Evans says:

---

<sup>1</sup> 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.

... 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 if there are *no* connections to anything that is immediately subjective, then the difference between an impersonal description of the objective world and an impersonal description of a fictional world vanishes. If Nagel is right then the supposition that I am R.G. is much on a par with the supposition that I am Sherlock Holmes. But of course no two suppositions could be more distinct. 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 — not that it is merely a world such that I can entertain the supposition that I am identical to one of its inhabitants, but is such that I have no choice but to presuppose at least one such an identity.

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. There is no real argument 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.

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 certain kinds of illusions about the identity of a subject over time exist because ways of describing situations can covertly build substantive assumptions into our conception of some states of affairs in such a way that surprising conclusions follow. The example Evans gives is Anscombe's conclusion that self-identification has the mysterious property that misidentifications of myself over time are not possible: in various of my 'I'-thoughts over time, there is no possibility that there is an inadvertent switch of referent. Anscombe takes this to be reason to be suspicious of the idea that in self-identification one is identifying anything at all (a conclusion she shares with Geach and Strawson as discussed above). But Evans remarks that the remarkable guarantee in this case is a consequence of our description of the case, and not anything genuinely mysterious. As Evans points out, 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. A subject might misidentify himself, but in those cases the description that was used to describe the situation will not apply. Cases of this sort of thing will come up in the course of the chapter.

## 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. 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.). The second concern with respect to identification-freedom is to argue that the phenomenon is not tied exclusively to self-ascription of mental predicates, but rather is a function of ways of gaining information.*

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. The reason is that although the Ideas in all cases constitutively involve functional connections to information links, in the case of 'I'- and 'here'-thoughts, these links can be latent, provided the subject has the appropriate dispositions to act on any such information in the appropriate way should any arise. (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. Evans discusses other treatments of the phenomenon of immunity to error through misidentification (henceforth IEM) and points out that while it is often taken to be a phenomenon concerned with self-ascription of mental predicates, it is actually tied to *ways of gaining information*, and hence applies to some cases of bodily self-ascription, and does not apply to some cases of mental self-ascription.

For a mental and physical case in which the ascription exhibits IEM, 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 information. 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. For a case of each that does not exhibit IEM: a neuroscientist is looking at readouts of a number fMRI machines that are reading 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 thinks '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 all such cases, what makes or breaks the exhibition of IF is whether or not the information to the effect that the predicate is instantiated is coming in through an information link that is constitutively tied to one's I-Idea as part of its functional characterization. (As opposed to the information being collected in the CC of the Idea of some other object, with which an identity claim with the self is made.)

Evans will argue that one of the uses that the phenomenon of IEM is sometimes put to is doubly fallacious. That use is to discredit the notion that we manage to identify a subject at all in our self-

thoughts (Wittgenstein and Anscombe have made such arguments, and it is similar to the positions of Geach and Strawson that was discussed in Section 7.1). After quoting the Wittgenstein passage from *The Blue and Brown Books*, he points out an error that can be made on its basis. The phenomenon is that in certain cases of self-ascription, the information that is the basis of the ascription carries no identification information. Evans agrees that this is true. From this, it is concluded that in such ascriptions, nothing is identified. Evans argues that this is false.

One way to see Evans point here is as follows. A judgment based on information, including self-ascriptive judgments, identify an object as the object that instantiates the property in question. There are two ways this can happen. First, the deliverances of the information links (or set of cooperating information links) may provide information about the identity of the object 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 is instantiated gets treated automatically by the subject as information to the effect that the object corresponding to that Idea instantiates the property. But there is still an Idea of an object in play in the second case. 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 they are ascribing a property to nothing.

The second error is one Evans already mentioned and revisits at page 218: people take IEM to be tied exclusively to mental self-ascriptions. When these two errors are put together, the erroneous conclusion to the effect that in our mental self-ascriptions we are not identifying anything at all is reached. Thus we are told that if 'I' refers to anything, it refers to the subject of mental predicates, and since in such cases nothing is identified, there is no such subject, and hence 'I' does not refer. (This is essentially Anscombe's argument in 'The First Person'.)

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 is still not an asymmetry between mental and physical self-ascriptions.

### 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.*

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 bodily predicates. This would include things like 'My legs are bent', 'My hair is blowing around', 'I am being pushed'. Evans claims that in such cases it certainly appears to be the case that we have dedicated sensory channels whose deliverances we take to be of immediate, identification free relevance to ourselves.

To the suggestion that there might be deviant causal chains such that in fact the information coming in is not from the subject's own body, and thus in some sense a misidentification is made, Evans makes two counters. The first 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 is that if the proposal is that in normal cases an identity claim were involved, then it would not be clear what to make of the entities on either side of the identity operator. The suggestion Evans is arguing against here would be that the self-ascription that *I am F* is the product of two judgments: *b is F*; and *I am b*. Evans claims that the only Idea that could be in play for the entity *b* is a descriptive one along the lines of 'the body from which this information derives', and Evans claims that our normal self-ascriptive thoughts have nothing like form. And as far as the Idea of 'I' involved in the identity, Evans claims that if this Idea were such as to always require the kind of identity-dependent ascriptions envisioned by this proposal, 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 based on evidence. If this were not so, my identification of myself with my body would be the result of an inference based on a conclusion to the effect that only body *x* is so placed such that its causal encounters with its surroundings make sense of the pattern of sensory input I have (or on the action side: only body *x* is the one whose movements conform to the intentions I formulate). This sort of identification with one's body certainly seems to be quite unlike the way in which we actually identify ourselves with our bodies. (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 one's environment. For example, I self-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 next to a tree, but is it me who is next to a tree?" is anomalous at best.

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 his teacher Strawson's work in this respect in *Individuals* and *The Bounds of Sense*. I also discussed this a bit in my comments on Section 7.1 and Evans' remarks on p. 212.

More concretely, given that on Evans account we must have an adequate Idea of ourselves in order to entertain genuine 'I'-thoughts, and that an adequate Idea of ourselves is made available in part via a capacity on our part to identify ourselves with an entity (in the sense of having a capacity to understand an identity judgment to the effect that I am d) in the objective order, the theory described at the top of page 222 is the theory a mastery of which underwrites this ability. Is it, then, what makes our Ideas of ourselves adequate.

## 7.4 Mental self-ascription

*This section has five topics. 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 in operation in coming to beliefs 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 answer to my own psychological state proceeds by way of probing the world. 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 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: the resources that supply the content of the belief and the means by which one can discover that one holds the belief; and the conceptual mechanisms that make self-ascription of beliefs in general possible. The first part here is rather straight-forward: in ascribing a belief to myself, say in answering the question 'Do you believe  $p$ ?' what I do is *not* to look inwardly at a mental substance, but rather to bring to bear those cognitive resources that would be used to answer the question 'Is  $p$  true?'. If the yield of that procedure is a 'yes,  $p$  is true', then the answer to the first question is 'yes, I believe that  $p$ '.

But *forming* beliefs cannot be *sufficient* for the *self-ascription* of beliefs, of course. It is plausible that many agents (animals? Young children?) can have beliefs but not self-ascribe belief states. The latter is an achievement not guaranteed by the former. What else is needed is not an inward glance at mental state state that has been formed, but rather the concept of belief itself. 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 her own assessment to the effect that  $p$  is true is an instance of a belief state. The concept of belief must be the sort of thing that is in principle applicable to subjects other than herself. Evans closes this discussion with an allusion to Kant's famous pronouncement that the 'I think' must be able to accompany all my representations, but it is merely formal. By this Evans means to be giving voice to the idea that in absence of the concept of belief, a subject might form what are in fact beliefs, but would not herself be in a position to recognize that what she is doing is forming beliefs. From the subject's point of view, so



to speak, it is just manifest that the world is thus and so. The notion that she is holding beliefs that might be in error is not within her cognitive horizons. In such a case, the 'I think' has no grip. We as theorists looking at the subject can say, truthfully, 'She thinks that  $p$ ', but her self-ascription of that state is not possible. (Note: I think it is doubtful that this is what Kant had in mind, but that doesn't affect Evans' main point.)

**The second topic** is self-ascriptions of experiential states, and it runs from the middle of page 226 to the top of page 231. This account is similar to the account of self-ascription of belief states in a certain respect, in that it places at its core the reuse of capacities for representing the world. The structure of these capacities is as follows: the gaining 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 seemings (this was discussed in Chapter Five), and are non-conceptual. In forming a perceptual belief, one is applying concepts, *to the environment*, on the basis of this 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 on the content of the associated information state. For example when I see a stick partially submerged in water a certain information state is formed. On the basis of this information state, concepts such as 'bent' have their application induced, so to speak, in 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'. But again it should be noted that in neither case is the information state that to which the concept 'bent' is applied. In both cases it is applied to an (apparent) aspect of the subject's environment.

An analogy might help. Someone who wears contact lenses will make perceptual judgments on the basis of information states that are the way they are in part because of the operation of the lenses. The perceptual judgment 'That line is straight' when the subject is looking at a line drawn on paper does not concern the lenses, but some feature of the environment. Now the subject cannot perceive the lenses themselves (unless she removes them, of course, but assume she does not — if you don't like that example rephrase it in terms of the biological lenses in a subject with normal vision). But she can indirectly gain information about the lenses by simply using them. So if she wants to know if the lenses are smudged, she looks at some visual scene that she has reason to believe has only crisp contours, and if anything looks fuzzy or distorted, she can judge 'my lenses are smudged'. Similarly, my judgments to the effect that 'It seems to me as though ...' are arrived at not by applying any perceptual tools *at my own mental states* or *at the informational state*, but rather again *at the world*. But I use the deliverances of this outward glance to tell me about my informational state, exploiting the dependence between the specifics of the informational state and the perceptual judgments to which it inclines.

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, others appear not to be fallible, such as 'I seem to see red'. Some might try to exploit this to mount an objection to Evans' account,

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, experiential judgments involve two components: an informational state and a judgment based upon (though not *about*) that state. The competing account would have it that the information state is not distinct from the exercise of the concepts used in the judgment, but rather *just is* the disposition to apply the concept. If this were the case, then Evans' account, which posits a content-carrying information state independent of any judgments made on its basis, is incorrect. And the fact that there are infallible judgments such as 'I seem to see red' might appear to nail the case for the opposition, for infallibility is secured when the disposition to make a judgment employing a certain concept is identified with the information state itself. And it would also 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 disposition to make 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.' 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.

The second prong is that Evans points out that there is another, more plausible, explanation for this 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, but as not understanding what the concept 'red' means. So the infallibility of some psychological reports is not due to the fact that there is a fool-proof inward gaze, but is rather the result of the fact that the only ground 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 to 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.)

**Thirdly**, 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 presented to oneself as an object. The *illusion* that this gives rise to is the belief that the self does not figure in the content of anything that one is aware of. 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, because while the information link only carries information about the instantiation of predicates, thoughts based on this information employ an Idea that that information is directed towards. In this case, the surplus content concerns the content provided by the Idea itself, and the sorts of knowledge that allows one to conceive of oneself objectively. (One could imagine a transcendental argument here, whose starting point is the fact that our judgments have a certain kind of content, but the content of our informational input is less than this.)

**Fourthly**, 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 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 putatively turns to the issue of solipsism. He states that the phenomenon of identification-free knowledge needs to be understood correctly for it can be a pressure toward solipsism. This way of introducing the issue seems unclear, for nothing in the subsequent discussion seems in any way to be, or be an antidote for, pressure toward solipsism. Rather, it would seem that the discussion here might have been better framed in the following way: 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 solipsists has no self-concept). The relevant solipsists 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 the 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*. The solipsist might respond in the following way: 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 definition would 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 solipsists makes it seem as though a workable, non-circular definition of the self is being offered.

## 7.5 Some consequences

*This long section consists of three main parts. First, Evans discusses what he calls cognitive dynamics, which is a system of processes that modify the tense of self-ascriptions. So for example, one kind of memory is the result of a process by which one's judgment that I am F is retained and later automatically (i.e., without inference) becomes a belief that I was F. Second, Evans discusses another way that memory can operate, where what is retained is not a belief state but an informational state. At a later time a belief state may be formed on the basis of the retained information. Finally, Evans discusses at length the issue of whether the deliverances of memory of the second sort are immune to error through misidentification.*

The **first** issue is the cognitive dynamics of *belief* retention, and it runs from the beginning of the section at the top of page 235 to the top of page 239. I emphasized 'belief' because Evans will discuss two kinds of retention. In one case, an informational state at  $t_1$  produces a cognitive state (such as belief) at  $t_1$ , and the belief state is retained to  $t_2$ . This is the sort of process that is discussed in this first part of this section. The other situation is where information is received at  $t_1$  and is retained until  $t_2$ , and at  $t_2$  a judgment is made on the basis of the retained information. This is the process that will be discussed in the second part of this section.

The cognitive dynamics issue is similar in many respects to the topic of the first section of the Appendix to Chapter Six, where Evans discussed Dynamic Fregean thoughts. There, recall, the point was that there appears to be a temporally extended process such that the *same* thought might persist over time in a way that manifests a sensitivity to the subject's changing spatiotemporal position in an objective order. Given this, the beginning of this section can seem puzzling:

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]

Recall that the first section of the Appendix to Chapter 6, and Evans' article 'Understanding demonstratives' from which that material is taken, then there is a similar rule, so to speak, for the case of temporal indexicals. Evans' example there was the thought 'today is fine' and the continuation of the subject's grasp of just this thought as its verbal expression changes to 'yesterday was fine'. Thus, it would seem that if I remember being in a position, at time  $t$ , to assert 'today (now) is fine', I will be in a position, on the day after  $t$ , to assert 'yesterday (then) was fine'. The quote above might seem puzzling because it might seem that there is a rule of precisely the sort Evans says there isn't.

The point Evans is making though is the nature of the 'rule'. In the case of indexically identified times and places (*now*, *here*) and their more demonstrative-ish kin (*then*, *there*) there are multiple objects, so to speak, that are being kept track of against a background of possible competing objects. And on top of that there is a temporal element that systematically modifies time/tense. If I judge on Monday 'Today *is* fine' and on Tuesday think 'yesterday *was* fine', then there is one day I am tracking against a background of other days (Sunday, Tuesday, etc.). My thought tracks Monday even as 'today'

becomes Tuesday. Similar remarks hold for a location though of initially as 'here' that becomes a 'there'. And in both the temporal and spatial cases, tense is being modified accordingly.

But in the case of I-thoughts, there is only the latter mechanism, the systematic modification of tense. There are no competing objects that the self is being isolated from and tracked against.

That distinguishes 'I' from 'now' and 'here'. What about 'that'? Demonstrative Ideas also track a single object. Sort of. There are two respects in which the two kinds of Ideas differ. First, while a given demonstrative Idea tracks a single object, demonstrative Ideas as a type can track any number of objects: a few at any given time, and many at different times. Whereas one's I-Idea is designed to track one single object period. The second difference, closely related to the first, is that in the case of demonstrative Ideas the perceptual system needs to do work to maintain the ILs to the relevant object. In the case of one's I-Idea, nothing analogous to the *tracking* of the self is provided for.

In all the above cases, though, the beliefs that employ a single Idea, even if they span a temporal interval, are IF in that they don't rely on any identification (in the sense of  $[\delta_t = \delta_t]$ ) of what we might call different time slices of the object. But Evans points out (bottom of p. 237) that the sort of identification-freedom involved in judgments involving memory is separate from the IF status of the original judgment. For example, the judgment at  $t_1$  may be been identification dependent, but the memory at  $t_2$  of the judgment could still be IF. I might at  $t_1$  see a sunburnt person in a mirror and wonder "Someone is sunburnt, but is it I who am sunburnt?". But when I remember this at  $t_2$ , the *memory* would be IF in that it wouldn't normally make sense for me to wonder "Someone saw a sunburnt person, but was it I who saw a sunburnt person?" The IF status of the memory component will be discussed below. But the point is that there are two spots where identification links might occur in past-tense judgments.

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 **second 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*. The difference here is related to the difference between self-ascription of belief states and self-ascription of experiential states, discussed in 7.4. There, Evans claims that during perception there is an information state induced in the subject through sensation that can subsequently be conceptualized to yield a belief state, with self-ascription of *experiences* dependent upon the *informational state*, and self-ascription of *beliefs* dependent upon the *belief state*. In the sense 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 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 that is retained in memory from  $t_1$  to  $t_2$ , and 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. And as with self-

ascription of experiential states, that 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 at the bottom of page 248, 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 appear to 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). Call these *q-memories*. It is also possible that one might either suspect or know that one's own apparent memories are 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 exhibit 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 in the normal case such an inferential articulation is present. In the present case, from the fact that a subject can be aware that it is a possibility that she is q-remembering in such a way that an inference would be required 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 this bad move, but provides independent reasons for thinking that memory cannot, in general, be identification-dependent. Such an identity would be one of the form [I = that person], where that person is the person who 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, he 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 he's 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 his 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 that leave the identity of the subject open. Call this q\*-memory. Where q-memories still present themselves as being one's own, even though they may in fact not be memories of events one was involved in, q\*-memories are presented to the subject in a subject-independent way. 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 happened to someone else will arrive via inference, to conclusions such as 'someone, perhaps other than myself, was  $F$ ' on the basis of those memories. From this, it is concluded that the faculty of q-memory is actually q\*-memory, providing informational states whose content leaves the *apparent* identity of the subject unspecified — their content being simply 'someone was  $F$ '. Once this illicit slide from q-memory to q\*-memory is in place, it is all too easy to make a further slide to the effect that in the normal case, one's memories provide one with information whose content is such as to require an inference in order to establish that oneself is the subject of that experience.

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, imagining '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. The suggestion is that the difference between these two is comprised of some sort of surplus premise or inference to the effect that the one that one is imagining being in the West Indies is *oneself*. It is this surplus that the q\*-memory theorist appeals to as what is required for her account of memory. The account is that memory is to be analyzed as q\*-memory plus this surplus.

Evans' reply is to the effect that Williams' observation is correct about imagination, but that we cannot draw conclusions about memory on its basis. And Evans is also surely correct to point out that this gets things exactly backwards — the normal operation of memory provides states that presents the experiences as if they were the subject's, 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 normal memory is a composite of q\*-memory plus a surplus premise and inference; rather it is that q\*-memory is a composite of normal memory plus a special premise and inference (concerning the possible abnormal functioning of one's memory).

## 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 is unable to think about herself, if the subject lacks discriminating knowledge. Two kinds of case*

are discussed. The first are cases where the controlling conception of oneself is being fed information from more than one object. In this case there is not a single object that the information is coming from, and hence 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.

In a book full of novel, fascinating and counter-mainstream positions, this section is perhaps the most extreme. Having rejected the 'no ownership/'I' does not refer view, Evans now rejects its main rival, the Cartesian view to the effect that 'I'-thoughts have a guaranteed referent, the subject's own mind. The first sort of case 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. In cases where 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 of the result that there is no 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, involving substitution of information sources via q-memories and such, one's attempted 'I'-thoughts might be ill-grounded and lack an object. Nevertheless, if one is appraised of the situation, one could always retreat to a more limited sort of 'I'-thinking, one in which one abandoned the attempt to identify oneself prior to the onset of the disruptive processes, and restrict one's self-directed thought to times when the thoughts would be well-grounded. It would be possible, in this way, to have a sort of temporally modified by 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 will be claimed 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 qua spatiotemporally extended objective entity in the world — has been revealed to be non-existent, along with the rest of the 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, a mere theoretical possibility. It might be the case that such general elements of the charade (people's thoughts are sustained by brains) are modeled on facts of the reality in which the brain and its computer interfaces are maintained. But it might not.<sup>2</sup>

The next possibility is that the subject might 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). To this Evans points out that in the case at hand, the subject has no adequate Idea of 'this thought' to use to anchor the attempted description. Normally, an adequate Idea of a thought would individuate it with respect to the subject entertaining it.<sup>3</sup> We must assume that in this example thoughts are being individuated by something other than their contents, for if they are being individuated by their contents, then a description such as 'the thinker of this thought' is not guaranteed to have a unique referent, as indefinitely many subjects may be entertaining exactly thoughts with exactly that content. The example is easier with cases such as 'the experiencer of *this pain*'. And here it is the subject's capacity to have any inkling of the numerical identity of the pain that Evans is questioning, since normally numerical identity of pains is established via reference to the subjects that experience them (unlike qualitative identity, of course). But *ex hypothesi* it is the subject's identity that is in doubt.

---

<sup>2</sup> 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. And so guesses about there being a brain somewhere would be just that, mere guesses.

<sup>3</sup> 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*.

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

## 7.7 Conclusions

This concluding section is broken down into four numbered points. The first is by far the most substantive, and adds significantly to the prior discussion. The gist 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 some description unworkable, because no description can 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 sort of mysterious subjectivity that is impossible to reconcile with the objective order. He points out how his view, because it accounts for the phenomena in an entirely naturalistic way, lets us move past such views.

The third point concerns the fact that, on Evans' account, our 'I'-Ideas span the physical and mental. They are based on ways of gaining knowledge that are identification-free, and such that both bodily and mental ascriptions depend on these ways of gaining knowledge.

The fourth point concerns the dependence that our 'I'-Ideas have on certain empirical 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, is mistaken.

### 7.A1 Appendix Section 1

*In this section Evans introduces and defends what he calls the 'self-reference principle'. It is that in thinking about oneself self-consciously — the sort of thought that is associated with the first-person pronoun 'I' — one is bringing it about that one satisfies " $\xi$  refers to  $\xi$ ". Evans argues that this principle is part of an explication of our fully self-conscious self-knowledge. But it is not by itself the entire story. It does guarantee that the self one self-consciously refers to must be the subject of psychological predicates, but beyond that it determines nothing. In particular, that principle does not rule out that the self is a boldily thing.*

Evans begins by remarking on Anscombe's seminal paper 'The First Person'. As Evans reconstructs her argument, she claims that the import of the first person pronoun 'I' is not a device that the subject uses to refer to himself. The argument is that if this were true, then there would be only two ways to

analyze the referential potential. First, one could analyze it as “X refers to X,” for instance it is a device that Alice uses to refer to Alice. But this can't be right for the usual reasons (it would entail that it would be sufficient to license the use of ‘I’ that Oedipus refer to the slayer of Laius, which we know does not capture the correct sense, since Oedipus may not realize that he is the slayer of Laius). Or second, it gets explicated in some way using the first person pronoun, in which case the analysis is useless because circular. The conclusion then is that ‘I’ is not a referring expression at all.

Another suggestion discussed by Anscombe is that the first person pronouns is what one uses to knowingly and intentionally refer to oneself. But she dismisses this again because both ways of filling out the intention seem to fall into the same two options above, and hence not advance the topic.

But Evans says that the intention can be spelled out as: the subject intends to bring it about that he satisfy the concept-expression ‘ $\xi$  refers to  $\xi$ ’. It is this that is the topic of the section. In the paragraph bridging pp. 258 and 259 Evans' point is that while it is true that my intending to bring it about that I satisfy “ $\xi$  refers to  $\xi$ ” is equivalent to my intending to bring it about that “ $\xi$  refers to me”, (or even “I refer to me”) it is not the case that the *explication* works in the way Anscombe suggests. Evans' suggestion is that it is not that my intending to satisfy “ $\xi$  refers to  $\xi$ ” can only be explicated by appealing to my intention to satisfy “ $\xi$  refers to me”; but rather that it is the other way around. And because of this, the suggestion is not “useless”. It advances our understanding of the sort of thinking expressed by first person pronoun.

And he starts out, in the next paragraph, by pointing out that I-Ideas as he has so far discussed them do not automatically satisfy this principle. A creature could have an I-Idea as explicated in Chapter 7, and in using this Idea it might have thoughts about what is in fact itself. And not merely as one subject among others, but in a special way provided by the unique functional connectios to perceptual input and behavioral output. But though this way of thinking is special, it does not guarantee that in using this Idea the subject is aware that it is satisfying the predicate “ $\xi$  refers to  $\xi$ ”. Or to put it another way, it will be satisfying “ $\xi$  refers to  $\xi$ ”, but it can't *intentionally bring it about* that it satisfy “ $\xi$  refers to  $\xi$ ” because it has no inkling that it is in the business of self-reference.

What else is needed? In the next paragraph (bridging pp. 259-60), Evans brings up the idea that one thing that would be required is that the subject must realize that it is a thinker, or the subject of psychological properties. That is, what is required would be to beef up the analysis of I-Ideas -- a sort of upgrade (analogous to the cognitive dynamics discussed earlier) that included psychological self-ascription of the sort discussed in the chapter. So that in using an I-Idea to think “I was hot”, the subject also, sort of automatically, is disposed to engage in psychological self-ascription and judge “I believed that I felt hot.” Such self ascriptions would also be connected, in the usual sort of special way, to that same I-Idea. Then when the subject entertained I-thoughts, there would be the possibility of it self-ascribing self-directed beliefs.

In the next paragraph he compares this rather modest proposal with a more demanding one, which is that the sort of thought one would have to have would exhibit a sort of impossible-to-implement identity across the divide between subject and object. The suggestion is that Evans' own more modest proposal is sufficient.

Next he points out that the SRP can offer part of an explanation for why there has been a tendency to identify subjects with their minds. The suggestion is that this is because the SRP establishes an identity between the object referred to and the subject of the referring thought. And so it seems that the only thing it guarantees as the object of the thought is the subject qua psychological entity. But Evans points out that this is just an accident of the nature of the question; and that if the sort of reference in question were linguistic self reference, one could establish that one was essentially a speaker.

However, there is a route from the SRP to an asymmetry. The asymmetry is between treating any physical predicates as true of something which is not me (but is mine), *my body*; and treating my mental predicates as true of something which is not me (but is mine), *my thinking organ*. The SRP rules out the possibility of the second, but does not rule out the possibility of the first. Evans wants to frame this not as showing that we are in fact essentially mental beings and not essentially physical beings. Just that our essential physicality is not certified by the SRP.

The final point is just that the SRP cannot be the source of complete self-knowledge concerning what one is. That is, one can't conclude from the fact that it does not guarantee that one is a physical thing that one is not a physical thing.

## 7.A2 Appendix Section 2

*In this section, Evans discusses 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.*

This is an odd section. Evans first describes what he has called the functional component of the Idea types he has discussed — demonstrative, 'I'-Ideas, 'here'-Ideas, etc. The functional component has to do with the identification-free information links between these Idea types and relevant avenues of informational intake and behavioral expression. He next points out that this component by itself is not sufficient for thought about the implicated objects or places, for the functional component by itself does not guarantee satisfaction of Russell's Principle: it does not allow the subject to know which object in the objective order the thought is meant to concern. For that, additional resources are needed. These points are not new. Evans has been making use of them since Chapter 6. There, recall, much effort went into describing how the information link with the demonstratively identified object did two things: it played a role in the functional aspect of the demonstrative idea type, providing for the identification-free status of at least some of the subject's thoughts of the object; and it also allowed the subject to locate the object in objective space, thus allowing the subject to know which object her thought concerned.

There are two points Evans makes here. First, he points out that the functional level is not dispensable, the work it does cannot be done by the 'background' level. As best I can tell the idea is that the background level is, as the level that allows the subject to *know which* object is in question, the level that is responsible for our conception of the object as the specific type of object it is. And so the suggestion is that the functional stuff cannot be deduced from the general stuff. Understood this way Evans is in agreement with 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, and to those raised in this Chapter in relation to Nagel's argument about the gulf between the objective and the subjective.

The second point, articulated in the last paragraphs, is that these two components are closely connected. The way he had been describing the situation in prior chapters was that an Idea that is functionally characterized can, to be adequate, be capable of being judged to be identical to a fundamental Idea. But this treats these two things as distinct components. Whereas here he is suggesting that at least some of the resources that had been located in the 'background' level, the level that provides for the Idea's adequacy, have their source in the functional Idea itself. 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)

One might ask why is this material in Chapter Seven, and does the fact that it is shed any light on what it is supposed to mean? There's a clue at the very end, a worry about whether one's self-directed thoughts concern oneself or one's twin on Twinearth. Here is one thought. My twin and I will have, let us suppose, identical conceptions of the objective universe. Exactly because this conception is by definition bereft of the one sort of element that could distinguish the two: contextual elements. So if that is the only part that is responsible for the thinker's knowledge of which particular is in question, then there is a problem.

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 the x-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 what was called in Part 2 of the Appendix the 'background' aspect of our Ideas. This is the part that is responsible for their adequacy. The specific question is about what is involved in our Ideas being objective, since if they are fully objective then the connection between them and the functional Ideas becomes problematic in one way or another.*

Evans begins by pointing out that while his particular apparatus of the fundamental level of thought may seem unnecessary, the job that it is intended to do is entirely necessary. 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. Evans states that since our thought clearly is objective, 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. (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)

One point to make is 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*. These two tasks will arguably overlap in at least the domains Evans specifically mentions here – 'this'-thinking, 'I'-thinking, 'here'-thinking -- since it isn't implausible to think that discriminating knowledge of material bodies qua material bodies requires understanding what distinguishes them from all other material bodies in an objective spatiotemporal order. Nevertheless, the fact that Evans moves from 'discriminating one entity of a type from all entities of that type' to 'conceive of an entity objectively' can potentially shed light on his thinking on these issues.

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. If the map that is built up from these encounters is conceived of in a purely objective way, then the threat of massive reduplication becomes real. It is only because the map is conceived of as containing *these* objects — not just some set of objects that has a certain set of objective properties and relations that this set of objects exemplifies, but *these very objects* — that it is capable of imbuing the egocentric space that is coordinated with it with a workable kind of objectivity.

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 intimates that the 'objective' aspect of an Idea has some elements of the ground level functional part built into it in one way or another.

## 7.A4 Appendix Section 4

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. The discussion in this section of the Appendix and the previous few sections are probably addressing 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.

[Note: 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*).]