

Immunity to Error as an Artefact of Transition Between Representational Media

Jenann Ismael
University of Arizona
jtismael@u.arizona.edu

Introduction

It is often claimed that there is a range of self-ascriptions that are immune to error through misidentification relative to the first-person pronoun (IEM for short). Philosophical interest in IEM is usually traced to Wittgenstein's remarks in *The Blue Book*. He writes

"It is possible that, say in an accident, I should feel a pain in my arm, see a broken arm at my side, and think it is mine, when really it is my neighbours... On the other hand, there is no question of recognising a person when I say I have toothache. To ask 'are you sure that it is *you* who have pains?' would be nonsensical... And now this way of stating our idea suggests itself: that it is as impossible that in making the statement "I have a toothache" I should have mistaken another person for myself, as it is to moan with pain by mistake, having mistaken someone else for me."ⁱ

Shoemaker 1968 defines the susceptibility to error through misidentification thus, "to say that a statement '*a* is *j*' is subject to error through misidentification relative to the term *a* means that the following is possible: the speaker knows some particular thing to be *j*, but makes the mistake of asserting '*a* is *j*' because, and only because, he mistakenly thinks that the thing he knows to be *j* is what '*a*' refers to."ⁱⁱ

To say that a term is immune to error through misidentification is to say that this sort of mistake is not possible. It is recognized nowadays that this needs to be supplemented by relativity to grounds. Almost any judgment that is IEM when based on one kind of ground is subject to error when based on another. Coming to know *via introspection* that someone has a headache doesn't leave open the room for the possibility that I may have misidentified the sufferer. But coming to know by looking at a catscan of a brain evidently undergoing a migraine does, because I may be wrong about whom the catscan belongs to. And IEM doesn't extend to other ways of referring to oneself. If I know *via introspection* that someone has a headache, then I know that I have a headache. But I may not know that *JI* has a headache, even though I am *JI*. If I have amnesia, for example, or suffer from delusions about my identity. The "I" in "I"-thoughts is intersubstitutable *salva veritate* with other ways of referring to myself, but the validity of those substitutions is hostage to the requirement that they actually *are* other ways of referring to myself. And this is something I can be wrong about.

It might be thought that IEM shows that "I" is not a well-behaved referring expression. The suggestion was famously made by Lichtenberg that the appearance of reference is an illusion and that I-thoughts are not really thoughts about an object denoted by "I". He suggested that Descartes should not have said "I think" but "It thinks" ("Es denkt") or perhaps "There is thinking", to be understood in the sense of "it is raining". But the fact that "I" is intersubstitutable *salva veritate* with uncontroversially referential terms seems to show that it functions as a semantically significant constituent of thought. Therein lies the mystery; the fact that that I can pass directly from knowing *via introspection* that somebody has a headache to the belief that *I* have a headache without risk of error, where "I" exhibits all the characteristics of a genuinely referring term. It occupies an argument place that can take other values, it is intersubstitutable with other representationally mediated forms of reference to oneself, and it supports all of the ordinary argument patterns. The possibility of error is recognized and provided for in the integration of introspectively derived beliefs about myself into an inferentially articulated web of belief in which other ways of coming to know are recognized and in which one recognizes that there are other beliefs (those belonging to other people) that cannot be known in that way. The flow of information across this web is directed by the truth-preserving transitions that allow one to pass freely from one node of this web to another.

There's been an enormous amount written on the topic. It is connected to issues about self-knowledge, consciousness, Descartes' arguments for dualism, and the tendency to think of the self as a special sort of private object. I'm going to explain this phenomenon in terms that are a little more abstract than the terms in

which they are usually discussed. The more abstract perspective allows us to see IEM as an instance of a much more general phenomenon having to do with the interaction between representational media, and shows why it arises here.

In section I, I'll explain why it presents a puzzle for old views about representation. In section II I'll say why those views needed to be modified anyways to recognize reflexivity, but why reflexivity is only one of the pieces needed to solve the puzzle presented by IEM. In section III, I'll introduce the apparatus needed to describe the second piece and in sections IV and V, I'll show how the puzzle is resolved. In section VI, I raise a slightly different issue, a fully articulated world-view, from questions about how that world-view arises and how (if at all) it is justified. These are two different stories and the separation of issues is crucial to telling the semantic story.

I Why is identification freedom a puzzle?

There is some dispute about which self-ascriptions are properly classed as IEM. Some claim that only mental self-ascriptions are IEM, others claim that some non-mental self-ascriptions are IEM. And among mental self-ascriptions, we can distinguish self-ascriptions of experiences and self-attribution of intentional states. There is a question of whether these are all IEM, and whether the explanation is the same in each case. I'm going to bracket all of these disputes and look only at the central cases of self-ascriptions of experiences. Since it is the combination of the fact I can pass directly from knowing via introspection that somebody has a headache to the belief that *I* have a headache without risk of error and that "I" is intersubstitutable with uncontroversially referring terms that needs to be explained, I will take the explanandum to be the validity of these two transitions

1. {awareness of being in pain} → "I am in pain"
2. "I am in pain" → "JI is in pain"ⁱⁱⁱ

Pain is Wittgenstein's example, but the point goes equally for awareness of the presence of a table, the taste of cinnamon or the sound of thunder.^{iv} On the left of 1 we have awareness of the presence of pain, but it doesn't contain any reference to a subject to whom awareness is ascribed. On the right, we have the thought "I am in pain", which does ascribe the pain to an object.^v In 2 we have a transition from the thought "I am in pain" to "JI is in pain". The effect of 1 and 2 together is that information that one gets through introspection is fed into one's little mental folder on JI and integrated with information one gets about her through other channels. "JI" in 2 stands in for any name or definite description that refers to a person.

What is odd about 1 is that no evidence is required to make it, the "I" in the consequent is immune to failure, and yet it is not empty. When paired with the transition in 2, it yields a paradigmatic example of a straightforward matter of fact. The proposition that JI is in pain at a given time can get entered into the history books, other people assign truth-values to it and there can be evidence for and against it. The puzzle about IEM arises from the combination of real reference to an object and immunity to error through failure of reference.

Although some have drawn metaphysical conclusions about the nature of the self from IEM, it is really an anomaly for entrenched assumptions about the nature of representation. Shoemaker and Evans were the first to argue that IEM derives from the fact that reference to oneself in self-ascriptions like "I am in pain" is not mediated by representation, which means that the self is not singled out as the proper subject of those thoughts by satisfying something that has the form of a description. For, if reference to the subject is mediated by a representation, then it must be possible for the thinker to make a mistake about to whom the representation applies. But reference to oneself in the relevant I-thoughts is not mediated by a representation, and is therefore immune to that kind of mistake.

The reason that this presents a puzzle is that the deeply entrenched accounts of reference deriving from Frege don't make provisions for unmediated reference. They don't allow reference to particulars that is not mediated by a sense or individual concept, i.e., something that could be expressed as an identifying description. What it is on a Fregean account, to have a thought about an object is to have a concept of the object that stands in for it in thought. Anscombe, in "The First Person" was adamantly committed to the idea

that all reference is by way of mediating representations. If there is no possibility of error, then the appearance of reference is an illusion, and "I" is not after all a semantically significant constituent of thought. Here she is following Wittgenstein who reportedly held that "that in "I have toothache" the word 'I' does not "denote a possessor," and that "Just as no (physical) eye is involved in seeing, so no Ego is involved in thinking or in having toothache." vi In her words,

"This is perhaps the most important point, the identification of a presented object as oneself would have to go together with the possibility of misidentification, and it is precisely the absence of this possibility that characterizes the use of 'I' that concerns us."vii

Her argument was simple. No reference without representational mediation. Representational mediation entails the possibility of failure. No possibility of failure. Hence no reference. viii

II Reference Without Mediation

All of that was fairly shallow and rudimentary background. It is now widely recognized that there is something wrong with the view of reference embodied in the first premise. Studies of indexicals have signaled the need to recognize forms of reference that don't involve mediation by an individual concept. Indeed, it has become common to believe that indexicals are both essential to, and more fundamental than, reference that conforms to the Fregean model. More generally, there are good reasons to think that for there to be representationally mediated ways of referring to anything, there have to be representationally *unmediated* ways of referring to some. The latter serve as points of reference in relation to which objects identified in a representationally mediated way are picked out. The set of points of reference form a reference frame. In this sense, unmediated representation is first in the order of identification and ultimately grounds all other forms of reference. Strawson championed this form of argument and it was taken up and has been developed in different directions by many, myself included. ix If it is right, there must be some form of descriptively unmediated reference to particulars if there is any reference at all. It would seem that reference without a mediating representation is a condition of the possibility of any kind of reference. What grounds identification is ultimately something that can be identified non-descriptively. x

The need for models of reference that don't involve intermediaries has given rise to various adaptations of the pure Fregean model. In my view, a form of reflexive representation built on the model of the token-reflexive account of Reichenbach provides the unmediated representations that play this reference-grounding role. Provided that we get the semantics and content right, a token-reflexive account will supply a model for reference without semantic intermediaries. In these cases, the thought/utterance/ or representational act serves simultaneously as the vehicle of representation and object of reference, and is picked out *as such* by the rules of use for the expression. Since that is something that is guaranteed both to exist and to be present on the occasion of use, a competent user of token-reflexive expression cannot fail to know that the referent exists and cannot fail to recognize the present object as the thing that is referred to. xi So, for example, an utterance of the sentence "I contain four words" where "I" refers to the utterance in which it occurs generates immunities to failure of reference for "I".

Reflexivity won't by itself, however, explain the formal validity of transition in 1 because the immunities that flow from reflexivity on a token-reflexive account depend on the identity of the representational act (or state) and the object of reference. xii And the "I" in 1 doesn't refer to the thought "I am in pain". It refers to a new kind of object, a thing that is distinct from the thoughts ascribed to it, a thing that can support multiple thoughts, and is but one of a type of which there can be other instances. The troublesome feature of the phenomenon - and the one that really worried Wittgenstein and Anscombe (about which more, below) - was how we can introduce reference to an object that is distinct from the representational vehicle itself in a way that is immune to errors of both failure of reference and misidentification. And do it in such a way that it is exportable to the object position so that the subject can itself become an object, something that we can name and describe and to which we can ascribe properties. xiii Reflexivity suffices to understand how there can be awareness of experience. But this kind of self-awareness, as we might call it, is possible without awareness of one's *self* as a thing to whom those states belong. The conception of self as an object in this sense introduces a great deal of structure that the mere idea of reflexive awareness does not supply. Wittgenstein and Anscombe's discomfort comes from the structure that is introduced when self-awareness becomes awareness

of a *self* and the latter is treated as a new object of reference.^{xiv} The problematic transition for them is the inference from awareness of a state to awareness of an object to whom that state belongs, an inference that comes perilously close, in Anscombe's mind, to an inference to the existence of a Cartesian Ego.

III A helpful apparatus^{xv}

In this section, I introduce an apparatus to help us understand this transition that imposes just enough regimentation to highlight the featured relationships while abstracting from inessential detail. The central notion of a representational medium is familiar in cognitive science, information theory, AI, computer science, and engineering, and provides a way of talking about representation that spans different representational formats and vehicles. Anytime we can associate representational contents with the states of a system that can be used to send information from a source to a receiver, we have a representational medium. A representational medium is a collection of states with representational contents related by a dynamical process that effects transitions between them.

Almost any dynamical system can be treated as a representational medium. Partition the initial and final states and the dynamical process mediating them will induce a mapping from the former to the latter. If the initial states can be controlled, one can use the link to send messages to a receiver attuned to the final states. What information gets conveyed will depend on facts about sender and receiver, including but not limited to the use to which the information is put, the sensitivity of the receiver, the function or design of the systems of which they are a part. Some natural dynamical processes are selected for their informational properties and particularly amenable to description as representational media. We understand a great deal about the network of cell towers and satellites that litter our cities and skies, for example, by seeing them as information channels between telephones. We understand a lot about computers, about the way computers interface with one another, with fax machines, and with people by treating them as representational media. In natural systems, some processes are selected (in part) to play the role of conveying information as representational media. We understand a great deal about signaling pathways in cells, for example, by seeing them as information channels between cells. Description in informational terms will tell us a lot about why they have the physical properties they do, and we get a good high-level understanding of the global system dynamics. Like Mother Nature, engineers building artificial systems will designate certain processes specifically to carry information. The processes they choose will depend on their purposes. The media chosen for one purpose may have to interface with representational media chosen for another purpose, and the interactions between the media will provide us with a special sort of profile of the system that will, for some purposes, be especially revealing.

Representational media are characterized formally by the representational contents of their states and rules of transition. The apparatus can be applied in any setting in which we can associate representational contents with the states of a material regardless of whether the states possess their content intrinsically or derivatively, naturally or by convention, whether it is imposed by the viewer or inherent in the systems of interest. And it doesn't matter how the transitions are effected, e.g., by natural evolution, convention, or computational algorithm.^{xvi}

Complex systems of both natural and human design often employ multiple media and have cause to translate between them. A computer, for example, will employ multiple representational formats. A multinational corporation will employ multiple languages. The human brain is a representational system of tremendous complexity. It represents and re-represents information in formats for specialized uses, translating freely between them. Just as there are rules of transition among the states of a given medium, there are rules of transition *between* media. These transform the states of one medium into those of another. Philosophers tend to think of translation in linguistic terms. Translating from one language into another is an instance of translation between media, but more interesting examples include transformation of pictures into words, a collection of notes on a page into sound, or geometric into algebraic representations of two dimensional figures. Such transformations typically involve more radical change than the internal transitions that take us from one to another state within a medium. They preserve the structure or material properties of the original.

When you transform a visual impression into words, you change a lot more than when you transform “Tully was a great statesman” into “Cicero was a great statesman”.

The scope of a medium is a measure of the range of its representational states. Sometimes, when we compare the scope of different media, we have complete extensional overlap, a one-one mapping from representational states of the one into representational states of the other. In other cases, there is only partial overlap, and in some cases, the partial overlap is such that the scope of one medium is entirely included in the other. The scope of a map of Ireland, for example, is entirely included in that of a map of the world. The Tucson White Pages has restricted scope relative to that of the directory for the country as a whole. Restrictions on scope are always relative to a medium. M may have restricted scope relative to N, but not relative to O. When translating from a narrower to a wider-scope medium, we need to indicate the restriction on scope. There are different ways of formalizing the restriction.^{xvii} One way, which is particularly convenient when we have a wide scope medium interfacing with a set of media with narrower scope, is to introduce a parameter whose values effectively give addresses that locate the domain of the restricted scope medium in the one with wider scope. The representational state of the wide scope medium, then, contains an additional parameter whose value has to be given in translating from the narrow scope to the wide scope medium. That is what we do with area codes. To connect numbers listed in the Tucson white pages to numbers in the central directory, I need to add an area code. From the perspective of the central directory, the area code has the effect of identifying the source of the number from which the number was drawn.^{xviii}

So when we organize domain restrictions by parameterization, the translation then will involve specification of the value of a new parameter. The information-bearing states of one are more articulated than the other in the sense that there is an additional dimension of variation in content, represented by a semantic constituent that can take multiple values. This means that translation between them involves passing to a state whose representational content contains a larger number of adjustable parameters. The same phone line which requires seven numbers to identify in the Tucson directory requires ten in the directory for the country as a whole, and even more for a world-wide directory, which requires supplementation by a country code.

IV How to analyze IEM in these terms

The human mind employs multiple media, representing and re-representing information in different formats with various scopes. In practice, the introspectively available progression of thoughts, feelings, images, words, impressions and ideas that make up our conscious lives are woven together in a tangled mix.^{xix} But we can firm up the lines a little artificially to bring things into sharper focus and distinguish three media of special interest. There is the progression of content-laden qualitative states that carry information from the environment (Experience), the quasi-linguistic medium that we use for non-perceptual cogitation (Thought), and the public languages we use to communicate with one another (Language). I use capitals to signal that we've imposed some regimentation. Experience is imagistic, qualitative. It comes with representational content that can be expressed in discursive form, but it is not logically structured. Its representational states are more like images than sentences; they don't have logical parts and the transitions aren't inferential. In logically structured media like thought and language, rules of transition take the form of inferences. Some of these are deductively valid but in natural languages, the vast majority are material inferences. The material inferential structure of thought is soft and dynamic. It is a kind of repository of information about the world and changes as an agent's view of the world evolves. The material inferential structure of natural language is also soft and dynamic, though for different reasons. It acts as a kind of summary of distributed facts about speaker behaviour and it has a quasi-normative status over the behaviour of individual speakers.

The only thing that will matter about the contents of these media for our purposes is that Experience is 'egoless', which means that its representational states contain no articulated representation of the subject whereas Language and Thought do. Fully expressed contents in Language and Thought have an argument place for the subject that can take a range of values and that acts as a separate degree of representational freedom. There is no such degree of representational freedom in the contents of experiences. To experience pain is to be aware of a certain state, but there is no internal dimension of variation in that content of that

state that corresponds to different values for the subject. To *think* that I am in pain, by contrast, is to attribute pain to one person among any number who might have that property.

The puzzling feature of IEM involved a transition from a state whose content ‘ego-less’ to one whose content contains explicit representation of a subject, which is to say that it involves a transformation between media that articulates a parameter indicating domain restrictions from the point of view of the medium into which translation is effected. The translation from awareness of pain into a fully expressed content is effected in two steps. First, the additional parameter is made explicit by introducing a placeholder – “I” – occupying the subject position. Then its value is given explicitly in Thought by substituting a name or description, i.e., a term that has a semantic value in Thought.

1. {awareness of being in pain} → “I am in pain”
2. “I am in pain” → “JI [or any coreferential singular term] is in pain”

The “I” in 1 is immune to reference failure in two senses. Provided there are subjects at all, it cannot fail to refer due to the non-existence of its referent, nor can it fail to refer due to misidentification of the intended referent.

1 and 2 are rules that transform experiences into thoughts and thoughts into words. Together they determine how the information represented in experience *informs* thought. They have the status of translation rules, no different from the rules of transition that interpret the French word ‘chien’ in English, or interpret Fahrenheit temperature designations in Celsius. And like those rules, they are justified by extensional equivalence. The addition of a parameter is needed because of the greater representational scope of thought. Its value serves as a locator to identify the region of extensional overlap. So the semantic story is straightforward. 1 is a formally valid transformation from a narrow to a wide-scope medium that involves articulating a parameter. 2 is a substitution internal to the new medium involving substitution of coreferential terms. That substitution is governed by the rules of use for indexicals, in this case, “I” on any occasion of use can be replaced by a term that refers to the user. In this way, one passes from the beliefs licensed by inferences 1 and 2 *via* ordinary material inferences to beliefs about the way the world is.

V Self-location

The transition from a state of introspective awareness to a self-ascription of that state and then from a self-ascription to an objective characterization of a person transforms Experience into Thought. How that transformation is made determines how the information contained in experience is channelled over the web of belief. Transformations between media work in general by allowing licensed transitions between points of extensional contact. When you express (or translate) the contents of the representational states of one medium in another, the truth (or satisfaction) conditions of the old are given in the new medium. When moving from a less to a more articulated medium, that means that the truth conditions of the old will be reconstructed as implicitly relativised to a restricted domain whose value is given by a new parameter.

If we have an account of translation between media, then we can express the contents of whatever medium we are interested in by translation into the medium in use. The representational content of English sentences can be given by stating their truth conditions in French. But we could equally give the representational content of French sentences in English. Which of them we employ will depend on whether we are French or English speakers.^{xx} When *we* ask about the representational content of a medium, e.g., experience, we are looking for is a translation into Thought because for us, Thought is the medium of understanding. And we make the transitions in 1 and 2 intelligible to ourselves by showing that, when translated into thought, each of those transitions is truth preserving.

The reason that they are truth preserving is Experience has a place in the world represented in Thought, as part of the psychological history of a particular embodied subject. Passing from awareness of pain (or the presence of a table or any of the things I can know through introspection) to the self-ascription of that awareness is a licensed transition because it is an integral part of the view of the world embodied in thought that these experiences are part of the psychological history of an embodied subject. The substitution in 2, which does allow error, is justified by the contingent self-locating thought “I am JI”, which is shorthand for

the pure reflexive thought “these very thoughts and experiences are events in the psychological history of the embodied subject, JI”. This unpacking makes the source of the immunities to error very clear. The identification of the subject is immune to errors of misidentification because it is identified reflexively as the subject of *these very thoughts and experiences*. From the point of view of Thought, experiences belong to subjects, and the subject of these thoughts and experiences is identified implicitly *as* the subject of these very thoughts and experiences when it is identified as “I”. Coming to know *via introspection* that someone has a headache doesn’t leave open the room for the possibility that the sufferer may have been misidentified because the sufferer is identified *as* the subject of this pain, these very thoughts and impressions.

This rendering is very close to Anscombe’s explication. She writes

“These I-thoughts (allow me to pause and think some!) Are unmediated conceptions (knowledge or belief, true or false) of states, motions, etc., of this object here, about which I can find out (if I don't know it) that it is E.A.” xxi

Let’s pause and recap. The puzzle we started with was to explain how a *formally valid* transformation can introduce reference to an object in a manner immune to the kinds of errors of misidentification to which ordinary reference is susceptible and supports intersubstitution with paradigmatically referential singular terms. We explained why the transition preserves truth from the point of view of Thought even though it is justified in a quite different way than the deductive, and material inferences that are our models of formal validity. It is a substantive inference, one that depends on an understanding of how experience is situated in the world represented in Thought.

The first step in translating experience into thought involves the articulation of a parameter whose value is specified in terms that are immune to failure of reference because it is identified by an implicit, reflexive definition. “I” here identifies the subject as the person to whose psychological history these very thoughts and experiences belong. If there are subjects at all these thoughts and experiences fall in *mine*, because *I* just am (by implicit definition) that person – whomsoever she may be – to whom these thoughts and experiences belong. To put it a little differently, what each of us means by “I” is “the subject of these very thoughts and experiences”. This implicit, reflexive definition of a person is traded in for an explicit, non-reflexive characterization, justified by the self-locating thought “I am JI” in 2. This step identifies the subject in the language of the new medium and completes the translation. If I’m getting the information that someone is in pain through experience, then it is by its nature information about me, because the referent of “I” in this use is identified reflexively as that person about whom these states provide information. The reflexive element in the identification of myself by the first-person pronoun guarantees that there should be this connection between introspectively available states and myself, identified as the subject of those states. There are questions about what binds together the thoughts and experiences of a single person, bringing them under the purview of a single introspective point of view. And there are questions about what justifies the presumption that there are subjects at all, but if these thoughts and experiences belong to anyone, they belong to me. Compare, if I look outside and I see that it is raining, I know that it is raining here because ‘here’ identifies a place as a place about which I can get information by simply looking outside. Again, the immunities to error in misidentification arise because the referent of “here” is identified reflexively as the one about which one knows things in those ways.

VI The emergence of thinkers

The story I just told about the validity of the transition from awareness of pain to the thought “I am in pain” was told in the material mode, assuming the existence of a material world containing embodied subjects interacting in a spatially organized environment. At several points, I needed to make the qualification ‘if there are subjects’. I could have added ‘and if there is an objective world in which those subjects are arranged in the more or less the way we think they are.’ It is a story that a subject with a fully articulated world-view could tell to make the transitions in 1 and 2 intelligible to herself in her own terms. It is a story that could be told, that is to say, *in Thought*. But it is subject to the charge that it misses the whole point for at least a part of what concerned Wittgenstein (very likely) and Anscombe (quite explicitly) was that the transition from

reflexive awareness of, e.g., pain, to awareness of one's self as a *thing* to which pain is ascribed is far from an innocent grammatical transformation. What we are aware of in experience are states or Humean impressions, but no object to whom they belong. The transition from awareness of the presence of pain to "I am in pain" embodies in microcosm a transformation from what we might call 'self-awareness', i.e. introspective awareness of internal states, to awareness of one's self as an object in the world and brings with it the idea of self as substrate of thought, the bearer of mental states, the unity underlying the plurality. Anscombe identified this move as the source of the Cartesian error. In her view, this apparently innocent grammatical transformation brings with it the whole package of ideas that generates the mind-body problem.

Against this complaint, the story told in sections V -VI is powerless because it *presupposes* the whole package of ideas on which Anscombe is casting doubt. What does this package of ideas look like? It distinguishes one's sensory state from what those states represent, i It is the result of analysis that distills separable ideas of self and world out of experience. It distinguishes changes in experience due to changes in one's own state or position from changes due to differences in the environment. It separates information about the world from information about *me*. Self and world are conceived in this package as two independent dimensions of variation that conspire to produce experience. It is the result of analysis that distills separable ideas of self and world out of experience. The view of oneself as a spatially located embodied subject interacting with the environment through sensory channels is a lynchpin of the whole construction. This package of ideas is implicit in the employment of Thought as a representational medium. The emergence of Thinking is in this way, and for this reason, coincident with the emergence of a concept of self as an embodied subject.

And I don't think any non-question-begging justification for the employment of this package of ideas is possible in Thought because it requires an independent representation of the structure of the domain against which we can compare the structure imposed by Thought. So, I am in agreement with Anscombe and Wittgenstein (and Hume, before them) that what is given in experience – i.e., that of which we have reflexive (unmediated) awareness – is a progression of egoless states. I agree with Anscombe, moreover, that the transition in 1 is more than an innocent grammatical transformation. It brings with it a whole complex architecture that can't be derived from it, a world of embodied subjects interacting through linguistic channels. Is there a justification of this transition? That's an interesting and difficult question. Any justification given in Thought is question-begging for the reason we just gave, viz., it pre-supposes what it is supposed to justify. It can be motivated by talking about how a conception of an objective world is stabilized out of regularities in sensory streams, but there is little hope of a strict derivation.

There is a different question, however, that can be raised that does permit an answer in the terms employed by Thought. It asks why we represent the world at all in objective terms instead of just responding to sensory awarenesses like simpler animals? Why, that is to say, are there Thinkers? What natural advantage does interposition of the more articulated medium of Thought between experience and action hold and how did there come to be thinkers?

This is the story of the emergence of Thought. It is a story that I sketched part of in *The Situated Self*. Dennett has told the story in his way, and others have told it in theirs. However the story of the emergence of Thought is told, it comes with self-locating component that tells us how Experience is situated in the world represented in Thought. And the formal transitions in 1 coupled with the substantive self-locating identities in 2 will express that self-locating content. The more articulated world-view cannot develop without an internal understanding of how it relates to Experience. Reference to other locations is grounded in relations to one's own body, and one's own body is identified ultimately by its relation to Experience as the one that plays a certain role in *these* thoughts and experiences. In this way, all identification remains moored ultimately in unmediated, reflexive reference to events entirely internal to one's own psychological history. The result is that the connection between one's own location and all others is contingent, but undeniable.

This view of the special, reference-grounding role of self-reference was anticipated in the last sentence of Shoemaker's classic paper on IEM. He writes:

"There is, I think, a tendency to find the use "as subject" of 'I' mysterious and to think that it is perhaps not reference at all, because it cannot be assimilated to other sorts of reference... these other sorts of reference are possible only because this sort of self-reference, that involving the use "as subject" of 'I', is possible. There

is, I think, an important sense in which each person's system of reference has that person himself as its anchoring point, and it is important for an understanding of the notion of reference, and also for an understanding of the notion of the mental, that we understand why and how this is so.”^{xxii}

And it captures the truth in a theme that was central to Strawson, Evans, Brewer, Cassam and others that the self has to be conceived as an element of the world because the concept of world is inevitably the concept of something Self-containing, i.e., something that is occupied by *me*.^{xxiii} The connection is contingent because world and self are conceived as separate existences in a manner that allows us to conceive of a world unoccupied by selves, and to conceive of ourselves without bodies and world. But it is undeniable because identification of material particulars remains moored in and ultimately grounded by their relations to experience. Implicit definition is the key to this combination of ontological separability and *a priori* (contingent, but assured) connectedness.^{xxiv}

As problems about the mind go, there are easy ones and hard ones. Explaining how the semantic and cognitive peculiarities of first-person thought arise from the interactions between media with different scope is relatively easy. Telling a story about the emergence of Thought is hard, but not impossible. The trick is to tell that story in a way that doesn't lead to dualism. Whatever story one tells about the emergence of Thought, with its articulated argument place for the subject, however, has to be separated from the story of how it interacts with Experience once it is in place. The account of IEM in V and VI tells the latter story. It can, and must, make free use of all the structure inherent in Thought.

VIII Self as subject and self as object

How does the use of “I” as subject relate to its use as object? How do we associate the internal progression of thoughts and experiences with elements in an objectively conceived world of space-occupying bodies to which we ascribe material properties? How is that connection made in Thought itself? In the seminal passage from *Blue and Brown Books* cited above, Wittgenstein distinguishes the use of “I” as object which involves identification, and the use of “I” as subject which does not. The use of “I” as object occurs, for example, when I notice a wallet on the table and (checking my pockets) say “I am the person to whom that wallet belongs to”. The use of “I” as subject occurs when “I” is used to express awareness of an internal state. He writes

“One can point to the difference between these two categories by saying: The cases of the first category involve the recognition of a particular person, and there is in these cases the possibility of an error, ... On the other hand, there is no question of recognizing a person when I say I have toothache. ... it is as impossible that in making the statement "I have a toothache" I should have mistaken another person for myself, as it is to moan with pain by mistake, having mistaken someone else for me.”

The reason the topic was important for Wittgenstein, and indeed the reason that the notion of “I” has been a source of difficulty and intense interest since there has been philosophy is that it occupies the interface between two *prima facie* quite different referential orderings. I occupy space, I have ten fingers and toes, I have a mother and father and I move around the world. But I am also a subject of experience and thought, a thing that thinks, a thing to which these very thoughts and experiences belong. The difficulty of bridging the gap between these referential orders has given rise to very different reactions. Among them, Cartesian dualism which holds that there are two separate orders, equally real and contingently connected, the various forms of idealism which hold that there is only the mental order and that the material order is a construction or fictional projection out of it and more recently Dennettian anti-realism about the self which holds that the self is a fiction generated by brains in just-so explanations human behavior.

The concept of self as embodied subject bridges these two orders building in the needed connections between experience, thought, and action. The connection between the material world and the psychological events is made by the self-locating thought “I am so and so” against the articulated background of a world seen as occupied by embodied subjects. The thought “I am so and so” is further explicated in a manner that brings out its reflexive content as “These very thoughts and experiences are events in the history of a person that I

identify in objective terms as so and so'. This gives the concept of an embodied subject a central and indispensable role in Thought and explains why it is that embedded observers undergoing experiences are allowed to use the first person pronoun as an articulated constituent of thought and in an identification free manner that is immune to errors of misidentification. But it leaves open metaphysical questions about the status embodied subjects. Dualists and materialists and even Dennettian anti-realists can agree on the semantic explanation of IEM that I am proposing. They will not disagree that thoughts of the form "I am in pain" refer to an embodied subject (one that has a name and a place in the objective order. They will disagree on whether there are embodied subjects or whether an embodied subject is a (wholly) material thing.

Conclusion

Immunity to error through misidentification relative to the first-person pronoun presents a puzzle to ordinary ideas about reference because it involves a transition from a state whose content 'ego-less' to one whose content contains explicit representation of a subject in a manner that, on the one hand, supports intersubstitution with paradigmatically representational singular terms and, on the other, is immune to error. I suggested that two pieces are needed to solve the puzzle: (i) an understanding of how reflexive identification secures reference, and (ii) an understanding that reconstructing the contents of the states of a narrow scope representational medium in one with wider scope involves articulation of a new parameter. On this account, IEM is a formal product of the translation between media with different representational scope, together with the representationally unmediated character of reflexive identification. It has nothing in particular to do with *mental* representation and carries no implication of a special ontology for the *self*.

References

Anscombe, Elizabeth. (1975). In Samuel Guttenplan, ed., *Mind and Language*. Oxford: Clarendon Press, 1975.

Cassam, Quassim. 1999. *Self and World*. New York: Oxford University Press.

Evans, Gareth. 1981. "Understanding Demonstratives." In H. Parret and Jacques Bouveresse (Eds.),

— . *Meaning and Understanding*. Berlin: de Gruyter. Reprinted in Gareth Evans, 1985, *Collected Papers*. Oxford: Oxford University Press, pp. 291-321.

Floridi, Luciano. 2010. *Information: A Very Short Introduction*. New York: Oxford University Press.

Forbes, Graeme. 1989. "Indexicals." In D. Gabbay and F. Guenther (Eds.), *Handbook of Philosophical Logic, Volume IV*, Dordrecht: Reidel, pp. 463-490.

Frege, Gottlob. 1892. "Uber Sinn und Bedeutung." *Zeitschrift für Philosophie und Philosophische Kritik* 100. Translation by Herbert Feigl as "Sense and Nominatum" in Herbert Feigl and Wilfrid Sellars (Eds.), 1949, *Readings in Philosophical Analysis*, New York: Appleton-Century-Crofts, pp. 85-102. This translation is reprinted in Martinich 2001, pp. 199-211.

Frege, Gottlob. 1984. "Thoughts." In Frege (ed. B. McGuinness, trans. P. Geach and R.H. Stoothoff), *Collected Papers on Mathematics, Logic, and Philosophy*, Oxford: Blackwell, pp. 351-72.

Garcia-Carpintero, Manuel. 1998. "Indexicals as Token-Reflexives." *Mind*, 107: 529-563.

Ismael, Jenann. 2007. *The Situated Self*. New York; Oxford University Press.

— . "Me, Again". *Topics in Contemporary Philosophy, Volume 6: Time and Identity*, Keirn-Campbell, O'Rourke and Shier (eds.), Cambridge, MIT Press. 2008.

— . "How I Learned to Stop Worrying and Love Reflexivity" (ms.)

Kaplan, David. 1989a. "Demonstratives." In Almog, Perry, and Wettstein 1989, pp. 481-563.

1989b. "Afterthoughts." In Almog, Perry, and Wettstein 1989, pp. 565-614.

—. 1999. "Reichenbach's *Elements of Symbolic Logic*." German translation in Maria Reichenbach and Andreas Kamlah (Eds.), *Hans Reichenbach, Collected Works in 9 Volumes*. Frieder.: Vieweg & Sohn.

Perry, John. 1977. "Frege on Demonstratives." *Philosophical Review*, 86: 474-97.

—. 1979. "The Problem of the Essential Indexical." *Noûs*, 13: 3-21.

—. 2001. *Reference and Reflexivity*. Stanford, CA: CSLI Publications.

Reichenbach, Hans. 1947. *Elements of Symbolic Logic*. New York: Macmillan.

Strawson, Peter. *Individuals*. London: Methuen, 1959

Wittgenstein, Ludwig. *Philosophical Investigations*, London: Blackwell, 1953.

—. *Blue and Brown Books*, 1958.

J.J .C. Smart. *Philosophy and Scientific Realism*. New York: Humanities Press, 1963.

ⁱ Wittgenstein, 1958, pp.66-67.

ⁱⁱ Shoemaker, 1968, pp.7-8.

ⁱⁱⁱ I suppress temporal parameters. We can suppose for our purposes that these transitions take place instantaneously, though it is worth noting that the process of articulating the temporal parameters precisely parallels that of the articulation of the personal "I".

^{iv} First person avowals of the form "I believe that/judge that/will that/or assert that p" also exhibit immunities to failure of reference, but I leave them aside. There are differences that warrant separate treatment.

^v We can, if we like, distinguish the state of being in pain, which is an introspectively accessible qualitative state, from the thought "Pain!" or "There is a pain!", which brings that state under a concept. This is a thought of the kind one might have when one wakes up in the morning, feeling the signs of an impending migraine. Or when one is listening to a piece of music, one says to oneself; "there is the sound of an oboe, there is the French horn...".!". This distinction is a substantive one, but it doesn't concern us here. See TSS.

^{vi}The existence of a distinction between the two sides of the arrow can be contested. One might hold that there cannot be an introspectively available state without the corresponding judgment, as would be the case, for example, if introspective availability *consists* in the presence of the judgement. I remain neutral on this question.

^{vii} Shoemaker, *ibid.*, p. 561

^{viii} Compare Shoemaker, by contrast:

"If we consider the logical powers of first-person statements and the role played by the first-person pronoun in communication, nothing seems clearer than that in all first-person statements, including "avowals," the word 'I' functions as a singular term or singular referring expression. ..." (*ibid.*)

^{ix} 1959, p. 100. He traced it to Kant. See also Cassam, and references provided by him.

^x One might suppose that we know enough about ourselves to pick ourselves out by description, and then use ourselves as a basis for identification of other things. I know for example, that I'm the only person who has ever had this sequence of experiences in just this order. The reason this won't work is put nicely by Shoemaker; "in order to identify a self as myself by its possession of *this* property, I would have to know that *I* observe it by inner sense, and *this* self-knowledge, being the ground of my identification of the self as myself, could not itself be grounded on that identification" (*ibid.*, p. 563).

^{xi} There are different versions of token reflexive accounts. See Reichenbach,1947, 284ff, Smart., 1963, 132ff. It's not as obvious as it may seem to get the details of how reflexivity works right. See http://www.qsmithwmu.com/the_impossibility_of_token-reflexive_analyses.htm for difficulty with earlier attempts. The version in my "How I Learned to Stop Worrying and Love Reflexivity"

avoids this difficulty.

^{xiii}I don't make choices between talk of representational acts and representational states. Utterances and judgements are more naturally thought of as representational acts, but experiences or the states of a measuring device are more naturally thought of in the passive mode. We can translate between them by thinking of occupying state S at t as an act.

^{xiii} See, for example, Anscombe's discussion of Descartes' mistake in Anscombe 1975, and my discussion in Ismael 2008 and also Ismael 2007.

^{xiv} See Anscombe, *ibid.*, and Wittgenstein 1953.

^{xv} Representational media are also called sometimes called informational media. See Floridi (2010) for a very nice introduction.

^{xvi} The semantics tells us what they represent in extensional terms and the logic provides rules for transition that one can make freely, at least defeasibly, without new information or attention to context. In general, the rules of transition are truth-value preserving only provided we restrict range of contexts across which they preserve truth to the contexts in which the medium is used.

^{xvii} A fully general formal account of scope comparisons between media is possible, but unnecessary here.

^{xviii} Note that this does *not* entail that the parameter reflects anything internal to the states whose contents are being expressed. In typical cases, there is no way of expressing the relativisation in the medium with narrower scope, The new parameter is given meaning only in the new medium, which carries an enhancement of representational power.

^{xix} How we individuate media and which media we are interested in will depend on our purposes.

^{xx} The representational content of any medium is given in itself in deflationary terms (e.g., 'dog' refers to dogs, 'snow is white' is true *iff* snow is white), with allowable paraphrases of difficult vocabulary into vocabulary that is more precise or better understood.

^{xxi} Anscombe, *ibid.* p. 62. Anscombe's overall position is quite different than the one defended here.

^{xxii} Shoemaker, *ibid.*

^{xxiii} Self-containing, as used here, is distinguished from the reflexive adjective clause self-containing. It is not meant here to signal that the world contains *itself*, but that it contains me, this embodied intelligence, understood as a separate object of reference. The mapping of experience (obliquely, by way of construal as part of the psychological history of an embodied subject) into the more articulated framework provided by an objective world-view is integral to the content of the more articulated framework.

^{xxiv} In my view, the Strawsonian strategy of trying to parlay this into an anti-dualist argument doesn't quite work but for subtle reasons that deserve more than passing mention.