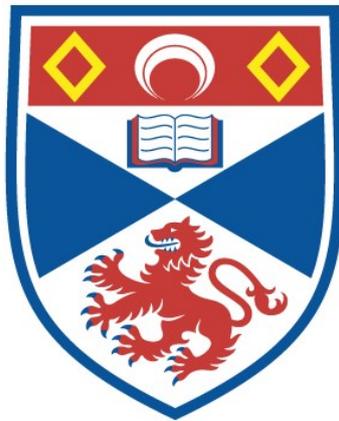


**IMMUNITY TO ERROR THROUGH MISIDENTIFICATION
AND THE TRILEMMA ABOUT THE SELF**

Annalisa Coliva

**A Thesis Submitted for the Degree of PhD
at the
University of St Andrews**



2001

**Full metadata for this item is available in
St Andrews Research Repository
at:**

<http://research-repository.st-andrews.ac.uk/>

Please use this identifier to cite or link to this item:

<http://hdl.handle.net/10023/12916>

This item is protected by original copyright

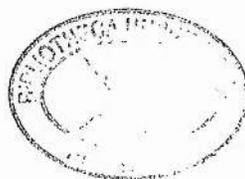
**IMMUNITY TO ERROR THROUGH MISIDENTIFICATION AND
THE TRILEMMA ABOUT THE SELF**

Annalisa Coliva

Submitted in fulfilment of the requirements for the degree of Doctor of
Philosophy in the School of Philosophical and Anthropological Studies

UNIVERSITY OF ST ANDREWS

20th February 2001



ProQuest Number: 10170791

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10170791

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

TL D838

Abstract

The thesis addresses the issues of error through misidentification and immunity to error through misidentification in relation to the problem of the first person. First, it provides an explanation of error through misidentification. Secondly, it shows that there are two possible ways of understanding immunity to error through misidentification. It is then argued that the first understanding of immunity to error through misidentification leads to what is labelled “the trilemma about the self”. That is to say, either we provide an explanation of immunity to error through misidentification, but we subscribe to two contentious metaphysical views about the self—the Cartesian and the Idealist; or else we hold the view that the self is identical with a human being, but we have no explanation of immunity to error through misidentification. It is then shown that in order to solve the trilemma, a different understanding of immunity to error through misidentification must be offered. After discussing various possible understandings of immunity to error through misidentification, a sound account of it is finally provided. Moreover, it is shown how non-inferential, introspection-based mental self-ascriptions can comply with it, in such a way that they turn out to be logically immune to error through misidentification.

Finally, by drawing on Evans’ and Peacocke’s accounts of the possession conditions of the first person concept—in which IEM I-judgements play a central role—it is shown that it is a concept of a human being who thinks of herself as such. Hence, our first person concept is firmly anti-Cartesian and anti-Idealist. As a consequence, it is maintained that not only is there no need to hold the Cartesian and the Idealist metaphysics of the self in order to explain why some I-judgements can be immune to error through misidentification, but it is also argued that one can no longer be either Cartesian or Idealist. For that would expose one to conceptual incoherence.

- (i) I, Annalisa Coliva, hereby certify that this thesis, which is approximately 70,000 words in length, has been written by me, that it is a record of the work carried out by me and that it has not been submitted in any previous application for a higher degree.

Date 20/2/2001 Signature of candidate.....

- (ii) I was admitted as a research student in October 1997 and as a candidate for the degree of Doctor of Philosophy in October 1997; the higher study of which this is a record was carried out in the University of St. Andrews between 1997 and 2001.

Date 20/3/2001 Signature of candidate.....

- (iii) I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the Degree of Doctor of Philosophy in the University of St. Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 20/2/2001 Signature of supervisor.....

- (iv) In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

Date 20/2/2001 Signature of candidate.....

Table of contents

Acknowledgements	p. v
Introduction	p. 1
SECTION I	
ERROR THROUGH MISIDENTIFICATION AND IMMUNITY TO ERROR THROUGH MISIDENTIFICATION AT THE ORIGIN OF THE TRILEMMA ABOUT THE SELF	
Summary	p. 16
Chapter One	
Error Through Misidentification	p. 18
1. Recognition-Based Judgements and Error Through Misidentification	p. 20
1.1. EM. A Comparison between Definitions: Shoemaker and Pryor	p. 27
2. Identification-Based Judgements and Error Through Misidentification	p. 31
2.1. De Re Misidentification and Which-Misidentification: Pryor	p. 34
3. Semantic Reference and Speaker's Reference vs. Error Through Misidentification	p. 38
4. Conclusions	p. 42
Chapter Two	
Immunity to Error Through Misidentification at the Origin of the Trilemma About the Self and the Forms of Its Solution	p. 46
1. The First Definition of Immunity to Error Through Misidentification	p. 47
2. IEM of I-Judgements and the Dilemma About the Self	p. 49
3. IEM of I-Judgements and the Trilemma About the Self	p. 55
4. Epistemological vs. Semantic Accounts of IEM	p. 59
5. The Forms of the Solution to the Trilemma	p. 64
5.1. Revisionist Accounts of IEM and Deflationist and Substantive Responses to the Trilemma	p. 68

SECTION II

A REVISIONIST ACCOUNT OF IMMUNITY TO ERROR THROUGH MISIDENTIFICATION AND A SUBSTANTIVE RESPONSE TO THE TRILEMMA: EVANS

Summary	p. 76
Chapter Three	
Evans: Perceptual Demonstrative Thoughts and the Revisionist Account of IEM	p. 80
1. Russell's Principle and the Generality Constraint	p. 82
1.1. Russell's Principle	p. 83
1.2. The Generality Constraint	p. 85
2. Demonstrative Identification	p. 87
2.1. Towards a Clarification of Our Concept of Perception: The Role of Information-Links	p. 87
2.2. Consequences: Russellian Thoughts	p. 96
3. Immunity to Error Through Misidentification as a "Corollary" of Identification-Freedom	p. 99
Chapter Four	
An Interlude: Evans on I-Thoughts and the Objectivity Requirement	p. 107
1. I-Thoughts	p. 108
2. I-Thoughts: Frege	p. 108
3. I-Thoughts: Perry	p. 114
4. I-Thoughts: Evans	p. 115
5. Evans and the Objectivity Requirement	p. 120
Chapter Five	
Evans: Immunity to Error Through Misidentification of I-Thoughts and the Substantive Response to the Trilemma	p. 124
1. IEM Relative to Physical Self-Attributions as "The Most Powerful Antidote to a Cartesian Conception of the Self"	p. 127
1.1. IEM of Bodily Self-Ascriptions Based on Somatic Proprioception	p. 130
1.2. IEM of Physical Self-Ascriptions Based on Outer Perception	p. 136
2. Mental Self-Ascriptions and IEM: Why the Cartesian Cannot Get Off the Ground	p. 139
3. Against the Cartesian Conception of the Mind I. The Case of Beliefs	p. 144

- 4. Against the Cartesian Conception of the Mind II.
The Case of Outer Perceptions,
Hume's Point and the Case of Inner Perceptions p. 148
- 5. Conclusions p. 154

SECTION III

REVISIONIST ACCOUNTS OF IMMUNITY TO ERROR THROUGH MISIDENTIFICATION AND DEFLATIONIST RESPONSES TO THE TRILEMMA

- Summary** p. 157

Chapter Six

Revisionist Accounts of IEM

and Deflationist Responses to the Trilemma:

Shoemaker, Wright and Pryor

p. 159

- 1. Shoemaker: Absolute vs. Circumstantial IEM,
De Facto vs. Logical IEM p. 160
 - 1.1. Shoemaker's Account of IEM of Non-Inferential
Mental Self-Ascriptions p. 166
- 2. A Revisionist Conception of IEM: Wright p. 174
 - 2.1. Wright's Explanation of IEM
of Non-Inferential Mental Self-Ascriptions
and His Response to the Trilemma p. 178
- 3. A Reminder: Pryor's *De Re* and Which-Misidentification
and the Corresponding Forms of IEM p. 181
- 4. Conclusions p. 187

Chapter Seven

I-Thoughts and Δ -Theories: Peacocke

p. 189

- 1. I-Thoughts p. 190
- 2. Peacocke's Account of IEM
of Non-Inferential Mental Self-Ascriptions p. 201
- 3. Peacocke's Account of the Epistemology
of First-Person Thoughts: Δ -Theories p. 204
 - 3.1. Representation-Independence p. 206
 - 3.2. Δ -Theories p. 210
- 4. Conclusions p. 215

SECTION IV
THE EXHAUSTIVE SUBSTANTIVE RESPONSE TO THE TRILEMMA

Summary p. 218

Chapter Eight
Immunity to Error Through Misidentification
of Mental Self-Ascriptions
and the Exhaustive Substantive Response to the Trilemma p. 220

1. EM and IEM p. 221

1.1. IEM as a Feature of the Grounds
of One's Judgement p. 227

2. The Extension of IEM I-Judgements p. 230

3. Logical IEM of Non-Inferential Mental Self-Ascriptions p. 232

4. Why There Is an A Priori Guarantee
That the Mental States One Is Immediately Aware Of
Are One's Own p. 236

5. Are There Counter-Examples to the View
that Non-Inferential Mental Self-Ascriptions
are Logically IEM? p. 240

6. Conclusions p. 245

Bibliography p. 248

Acknowledgements

When I first arrived in St. Andrews, as an ERASMUS undergraduate student from the University of Bologna, I did not know that I would have spent—with a one year interval—five years of my life there. However, when I started being supervised by Crispin Wright, I realised that I had an enormous chance of learning what doing research in philosophy means. Whether or not I have really learnt it, during the PhD and, before it, during the M. Litt., I have an intellectual debt towards him which I am not able to repay.

Other people have been extremely important in shaping my ideas with their comments and criticisms. In particular, I would like to thank José Luis Bermúdez who co-supervised the thesis during my first year and who has always been keen on reading the various versions of it I have written in the following two years and a half. I am also extremely grateful to Leslie Stevenson and Katherine Hawley who have commented on various versions of the thesis as members of the internal review committee for my doctorate. Finally, I have greatly benefited from discussions of the final draft of the thesis with Eva Picardi, Alberto Voltolini and Gianfranco Soldati.

I would also like to thank the University of St. Andrews for awarding me a top-up award and for providing a top-rate environment in which to pursue research.

Everyone knows that writing a PhD thesis is also a matter of endurance and stamina. I would not have had them without the continuous moral support of my friends, my parents and my husband. To all of them, my love and gratitude.

Introduction

Throughout the history of modern philosophy the problem of the first person has played a central role. Descartes, Locke, Hume, and Kant have all given different responses to the question: who is the I who thinks, feels and perceives? This question has been inherited by contemporary philosophy. Within the analytic tradition Frege, Russell and Wittgenstein—just to mention the fathers of that tradition—have all engaged with it.

This thesis starts precisely from the particular way in which Wittgenstein, in his middle period,¹ phrased that question. For he noticed that there is an important difference among our uses of the first person in speech and—we can add—in thought. On the one hand, it is possible, when one self-ascribes a physical property to make a mistake in identifying *who* has that property. Hence, it is possible for one to self-ascribe a physical property which is (or can be) true of someone else. On the other hand, when one self-ascribes psychological properties, then there is no question of wrongly identifying the subject who has them. In this latter case, the use of 'I' as *subject*—as it was labelled by Wittgenstein, as opposed to the use of 'I' as *object*—is guaranteed against a particular kind of mistake. In the literature on the topic the mistake in question has become known as *error through misidentification* and the corresponding impossibility of error as *immunity to error through*

¹Cf. *The Blue Book*.

misidentification.² Hence, according to Wittgenstein, when a subject self-attributes a psychological property she can (at most) be fallible about the property in question, but she cannot be wrong as to whether *she* is the one who has (or seems to have) it. By contrast, according to Wittgenstein, when a subject self-attributes a physical property she can be wrong both relative to that property and as to whether *she* is the one who has (or seems to have) it. Quite surprisingly, Wittgenstein concludes that while uses of 'I' as object are genuinely referential and refer to a person, i.e. to a human being or an embodied self, uses of 'I' as subject are not. That is to say, when 'I' is used as subject, it does not refer at all and, *a fortiori*, it does not refer to an embodied self.

Now, both the difference between uses of the first person that are liable to error through misidentification and those that are immune to it, and Wittgenstein's claim regarding the non referential role of the use of 'I' as subject, are in need of explanation. In fact, the difference between uses of 'I' which are liable to error through misidentification and those which are immune has been addressed several times in the literature, but quite seldom in a *comprehensive* way. That is to say, many philosophers have referred to it in their work and have also made attempts at clarifying it, but they have rarely considered it central to their reflections about the first person and the particular features of the first-personal use of psychological verbs.³ This

²These labels have been introduced in Shoemaker 1968, reprinted in Shoemaker 1994a.

³Shoemaker and Evans are certainly the ones who have made the most in giving a central role to error through misidentification and immunity to error through

thesis intends to remedy the situation. For it is only by seeing which role immunity to error through misidentification really plays within reflections about the self that it will be possible to understand Wittgenstein's claim about the non-referring role of 'I' when used as subject and it will be possible to assess it and show it to be mistaken.

Hence, in the first section, entitled "Error through Misidentification and Immunity to Error through Misidentification at the Origin of the Trilemma About the Self", the case is made for showing how the attempts at explaining immunity to error through misidentification of certain uses of the first person can lead to holding various positions about the self. To such an end, an exhaustive characterisation of error through misidentification is given in Ch. 1. It then becomes clear that error through misidentification is an *epistemological* phenomenon that depends on the *grounds* on which certain *judgements* (that, in turn, can be linguistically expressed) are made. In particular, if the relevant judgement is based on (a non-necessarily occurrent) belief in an *identification component*, which would constitute *the subject's rational ground* for her judgement, then the latter is liable to error through misidentification. That is to say, the subject who makes the judgement could have wrongly identified (or recognised) the object or person her judgement is

misidentification within their reflections on the first person. Wright and Peacocke, by contrast, have devoted attention to these phenomena, but they have not taken them to be pivotal to an understanding of various issues related to the first person. Pryor has focused on them in a systematic way, but he has not considered the relevance of these issues within the broader context of the reflection on the first person.

about. The relevance of insisting on the epistemological nature of error through misidentification becomes clear when this approach is compared to what can be called a *semantic* approach to the phenomenon at hand. According to the latter, error through misidentification consists in a *split between the semantic reference* of a certain term used to give expression to one's judgement *and the intended reference, or speaker's reference, of that term*. Hence, if X says "John has come to the pub", the person in the pub may not be the semantic referent of 'John', i.e. it could be someone else; yet, that person is the intended referent of 'John', when that name is used by the speaker on that particular occasion. Consequently, the subject can think and manage to say something true of the person presented to her, although the means used to give expression to her judgement are inaccurate. By contrast, we argue that, when error through misidentification occurs, the subject does not just want to refer to the person presented to her. Rather, she wants to refer to whom she takes to be the semantic referent of 'John', whom she thinks, mistakenly, to be the same as the person presented to her in the pub.

Having a clear characterisation of error through misidentification allows us, on the one hand, to dispense with a *semantic* understanding of *immunity to error through misidentification* (relative to the first person) which has been proposed in the literature on the topic. According to such a view, the first-person is governed by the *token-reflexive rule*, both at the level of language and of thought, i.e. it is governed by the rule according to which each token of the first person, either in speech or in thought, refers to its producer. Hence, uses of the first person can never exhibit a split between semantic referent

and speaker's referent. In short, since the first person is an *automatic indexical*,⁴ which hits its target in virtue of the token-reflexive rule, it can never be competently used by a subject in order to refer to someone else. Hence, on a semantic understanding of immunity to error through misidentification, *all uses of the first person would turn out to be so immune*. By contrast, we argue that although the linguistic meaning of 'I' is given by the token-reflexive rule—a fact this that, however, does not prevent 'I' from having a sense (cf. Ch. 7)—this is entirely compatible with the possibility for some uses of the first person to be affected by error through misidentification, once the relevant judgements are made on certain *grounds*. This issue is taken up both in Ch. 1 and 2 in connection with the semantic understanding of error through misidentification and of immunity to error through misidentification, respectively.

On the other hand, the characterisation of error through misidentification presented in Ch. 1 allows us to see more clearly one possible way in which *immunity* to error through misidentification can be understood. This is the main topic of Ch. 2. In fact, one may think that the I-judgements which are immune to error through misidentification are those which are based on a (non-necessarily occurrent) belief in an identification component, which, however, *cannot be wrong*. Such an understanding of immunity to error through misidentification is then closely parallel to our understanding of error through misidentification.

⁴"Automatic indexical" is an expression due to Perry 1997.

In fact, we argue that such an understanding of immunity to error through misidentification is at the origin of two influential views about the self, which both deny that the self who thinks, feels and perceives is identical with a human being. The thought is this: *misidentification of physical objects is always possible*. Hence, the referent of uses of the first person which are immune to error through misidentification *cannot be a physical object such as a human being*.

The Cartesian metaphysics of the self, which provides us with an *extraordinary object*, i.e. a mental substance which is present to the mind for each act of thought, and with *extraordinary conditions* for our encounter with such an object, i.e. a conception of the mind as a theatre where everything presented to the subject cannot be mistakenly identified, seems to be *prima facie* able to make sense of the possibility for some I-judgements to be immune to error through misidentification. However, the Cartesian metaphysics of the self and of the mind are problematic. For dualism is a difficult position to maintain. Moreover, the infallibility of the subject with respect to her mental states implied by the Cartesian conception of the mind can be shown to be mistaken. Finally, as Hume and Lichtenberg noticed, it is *prima facie* feasible to hold that no subject of sensation, perception and thought is presented to one as an object when one is aware of one's sensations, perceptions and thoughts.

As a reaction to Cartesianism, Wittgenstein, in his middle period, seems to have put forward the view that uses of the first person which are immune to

error through misidentification *are not about a person at all.*⁵ Hence, if *no object at all is given for recognition or identification*, when the relevant I-judgements are at stake, then the latter will be immune to error through misidentification *by default*. This position has become known in the literature as the *no-subject* or *Idealist view about the self*, and its offspring at the level of language as the *no-reference view about 'I'*. But this position too is highly problematic. For it would seem to imply that mental states are not owned. Moreover, it would be at odds with our linguistic intuitions, since we take 'I' to be a referential pronoun, from (the use of) which the existential generalisation follows.

Therefore, on the first understanding of immunity to error through misidentification, which calls for a *metaphysical account* of it, i.e. an account according to which I-judgements that are immune to error through misidentification are so because of the *particular nature of the self*, we reach the following *dilemma*. We do have a *prima facie* explanation of immunity to error through misidentification, but we either endorse the bad Cartesian metaphysics of self, or we endorse the bad Idealist metaphysics of the self and the bad semantics of the no-reference view about 'I'.

However, the situation is even more complicated than that. For, those who sympathise with the *primitivist* position about the self, according to which the

⁵Wittgenstein's claims in *The Blue Book* with respect to the first person have been interpreted in various ways. Yet, this thesis does not have exegetical ambitions and assumes the received view according to which Wittgenstein endorses the claim that uses of 'I' as subject are not genuinely referential.

self is identical with a human being, still owe us an explanation of how certain I-judgements can be immune to error through misidentification. Yet, as we have seen, on the first understanding of immunity to error through misidentification, such an explanation is not even in view. For such an explanation presupposes that there be infallible recognitions or identifications. However, since no physical object can ever be infallibly recognised or identified, if I-judgements are about a human being, then *they can never be immune to error through misidentification*. Hence, we are confronted with the following *trilemma*. Either we have a *prima facie* explanation of immunity to error through misidentification—but we either endorse the bad Cartesian metaphysics of the self, or the bad metaphysics and semantics of the Idealist position—or else we opt for a sound metaphysics of the self and a sound semantics of the first-person, but we have no explanation of immunity to error through misidentification. Consequently, the solution of the trilemma consists, at least in part, in offering *an alternative account of immunity to error through misidentification which makes it compatible with the reference view about 'I' and with a primitivist metaphysics of the self*. In fact, it also consists in showing precisely *which classes* of I-judgements are immune to error through misidentification and *why* this is so.

In order to give an alternative account of immunity to error through misidentification one can propose the following: one may hold that I-judgements which are immune to error through misidentification are those *not* based on *any* (non-necessarily occurrent) belief in an identification

component. This alternative way of understanding immunity to error through misidentification would certainly allow one to avoid metaphysical explanations of that phenomenon, since it would only appeal to the kind of *grounds* on which the judgement is made. Hence, it would replace the metaphysical explanation of immunity to error through misidentification with an *epistemological* explanation of it, which would also be in keeping with our *epistemological* understanding of error through misidentification. Yet, such an alternative account of immunity to error through misidentification has to be precisely characterised. Moreover, it has to be shown *which* I-judgements are immune to error through misidentification in this second way and *why* this is so.

Our understanding of the dialectic within which considerations having to do with immunity to error through misidentification are intertwined with an account of the metaphysics of the self and the semantics of the first person allows us to make a taxonomy of the most prominent views with respect to these issues which have been recently put forward in the literature on the topic. Hence, we call "*revisionist*" those accounts of immunity to error through misidentification that depart from the first account of immunity to error through misidentification presented in Ch. 2. Amongst them, we distinguish further those which aim at offering a *substantive response to the trilemma* and those which just offer a *deflationist response* to it. The difference among these positions is as follows. Substantive responses not only aim at making immunity to error through misidentification of certain I-judgements *compatible* with reference to a human being, but they also try to *rule out* the Cartesian and

the Idealist positions on *conceptual grounds*. As a consequence, if they are right, one *can no longer be either Cartesian or Idealist* about the self, because that would expose one to conceptual incoherence. By contrast, deflationist responses just aim at making immunity to error through misidentification of certain I-judgements *compatible* with reference to a human being. As a consequence, if they are right, one *need no longer be either Cartesian or Idealist* about the self, although these are still conceptually coherent, albeit problematic, positions. The aim of this thesis is to offer *a sound revisionist account of immunity to error through misidentification and an exhaustive substantive solution to the trilemma*. For, as it will become evident, substantive responses have so far been wanting with respect to the explanation of why introspection-based mental self-ascriptions are immune to error through misidentification.

Hence, in the second section, entitled "A Revisionist Account of Immunity to Error through Misidentification and a Substantive Response to the Trilemma: Evans", we present, amend and defend Evans' revisionist account of immunity to error through misidentification and his substantive response to the trilemma. We do so by introducing the basic elements of his theoretical framework and his account of immunity to error through misidentification first (cf. Ch. 3). We then introduce his account of I-thoughts, giving some historical background to it (Ch. 4). Finally, in Ch. 5 we consider the details of his proposal with regard to I-thoughts and their immunity to error through misidentification. In particular, we present and defend his extension of immunity to error through misidentification to I-thoughts based on somatic

proprioception and on perception of one's environment. These turn out to be at least *de facto* immune to error through misidentification. Moreover, according to Evans, these kinds of self-ascriptions must be at least dispositionally in place for one to be credited with the possession of the first person concept. Hence, we claim that Evans is right to say that our first person concept is a concept of an embodied self that is located in space. Hence, we agree with Evans that our first person concept is firmly anti-Cartesian and anti-Idealist.

We also argue that Evans is right in pointing out that only *non-inferential* introspection-based mental self-ascriptions are *logically* immune to error through misidentification, while inferential mental self-ascriptions are liable to error through misidentification. Yet, we argue that there is an *asymmetry*, which is in need of explanation, between *de facto* immunity to error through misidentification of some bodily self-ascriptions and *logical* immunity to error through misidentification of non-inferential introspection-based mental self-ascriptions.

Finally, we consider a different pattern of argument used by Evans in order to oppose the Cartesian and the Idealist on conceptual grounds. Roughly, it consists in a double application of the Generality Constraint. We argue that only one strand of it is successful and that it shows that in order to be able to self-ascribe mental properties subjects must *think* of themselves as embodied. Still, this argument does not establish the full anti-Idealist point according to which we have an *awareness* of ourselves as embodied when we self-ascribe mental properties on the basis of introspection.

In the third section, entitled "Revisionist Accounts of Immunity to Error through Misidentification and Deflationist Responses to the Trilemma", we turn to different accounts of immunity to error through misidentification and to various deflationist responses to the trilemma. In Ch. 6 we take into account Shoemaker's, Wright's and Pryor's papers on the topic. On the one hand, this allows us to clarify the difference between *de facto* and logical immunity to error through misidentification. On the other hand, it allows us to present and discuss other revisionist accounts of immunity to error through misidentification, which have been presented in the literature on the topic. Finally, it allows us to consider some possible explanations of why introspection-based mental self-ascriptions are logically immune to error through misidentification. However, we argue that, although these explanations present some elements that are relevant to an exhaustive explanation of immunity to error through misidentification, they are either partially wanting or incomplete.

In Ch. 7 we turn to Peacocke's account of I-thoughts which allows us to conclude the discussion of the nature of this class of thoughts introduced with Evans. We then argue that Peacocke's account of immunity to error through misidentification of introspection-based mental self-ascriptions is illuminating but that it rests on the assumption that one can only be introspectively aware of one's own mental states. We point out that we do not wish to take issue with this assumption, but that we wish to show why it holds *a priori*. We also take into account his more recent work on the epistemology of I-thoughts which is helpful in illuminating why when one self-ascribes a mental property

one is entitled to do so, even if, at least in some cases, one is not presented with oneself as an object, i.e. as an embodied entity.

Finally, in the last section, entitled "The Exhaustive Substantive Response to the Trilemma", a sound revisionist definition of immunity to error through misidentification is given. It is then shown why non-inferential introspection-based mental self-ascriptions are *logically* immune to error through misidentification. That is to say, it is shown why a subject cannot be introspectively aware of someone else's mental states and thus why the subject who self-ascribes the relevant mental state on those bases is necessarily the same as the one who has it.

In fact, it turns out that such an account of logical immunity to error through misidentification of the relevant self-ascriptions rules out the Idealist conception of the self (and its semantic offspring), since it maintains that no mental state can exist un-owned. Moreover, it dispenses with the Cartesian metaphysics of the self as unnecessary to an explanation of immunity to error through misidentification of the relevant I-judgements and as misguided on epistemological grounds.

Thus, the revisionist account of immunity to error through misidentification proposed, which is in keeping with Evans', together with the endorsement of his theoretical framework, allows us to maintain a substantive response to the trilemma. For immunity to error through misidentification becomes compatible with some bodily self-ascriptions. These, in turn, are at least dispositionally necessary in order for a subject to have the first person concept. Hence, our first person concept is firmly anti-Cartesian and anti-

Idealist. Moreover, our account of why non-inferential, introspection-based mental self-ascriptions are logically immune to error through misidentification allows us to supplement Evans' own position. Thus, we will finally be able to offer *an exhaustive substantive response to the trilemma about the self*.

SECTION I

**ERROR THROUGH MISIDENTIFICATION AND IMMUNITY TO
ERROR THROUGH MISIDENTIFICATION AT THE ORIGIN OF
THE TRILEMMA ABOUT THE SELF**

Summary

In this section we clarify and define the notions of (EM) *error through misidentification* (Ch. 1) and (IEM) *immunity to error through misidentification* (Ch. 2) and show how a certain understanding of IEM leads to the *trilemma about the self*.

To such an end, in Ch. 1

- we discuss recognition-based judgements and introduce EM (§ 1);
- we compare our understanding of EM with Shoemaker's and Pryor's (§ 1.1);
- we extend the class of EM-judgements to comprise not only recognition-based judgements, but also identification-based ones (§ 2);
- by so doing we disagree with Pryor (§ 2.1);
- we distinguish EM from the possibility of a split between speaker's reference and semantic reference (§ 3);
- we summarise our conclusions (§ 4).

In Ch. 2

- we offer a first definition of IEM which is closely parallel to our definition of EM (§ 1);

- we show how this definition forces one to explain IEM of some I-judgements by reference to the *particular nature of the self*. Two metaphysical views about the self, i.e. the Cartesian and the Idealist, are introduced as possible explanations of IEM. Finding them problematic, we claim that they confront supporters of the first definition of IEM with a *dilemma* (§ 2);

- the *trilemma about the self* is introduced: either we explain IEM, but we endorse either the Cartesian or the Idealist position about the self; or else we hold a *primitivist view about the self*, according to which the self is identical to a human being—and "I" refers to such an entity—but we have no explanation of IEM (§ 3);

- we consider *semantic solutions to the trilemma* (§ 4), finding them mistaken,

- we proceed to make a taxonomy of the possible solutions to the trilemma. We submit that only *revisionist* accounts of IEM, which depart from our first definition of IEM, offer the prospects of solving the trilemma (§ 5);

- we further subdivide revisionist accounts into *substantive* and *deflationist* ones, according to the kind of reply they give to the trilemma (§ 5.1).

Chapter One

Error Through Misidentification

In the *Blue Book*, in the context of his discussion of solipsism, Wittgenstein famously introduces a distinction between two uses of 'I'.¹ He writes:

There are two different cases in the use of the word "I" (or "my") which I might call "the use as object" and "the use as subject". Examples of the first kind of use are these: "My arm is broken", "I have grown six inches", "I have a bump on my forehead", "The wind blows my hair about". Examples of the second kind are: "/ see so-and-so", "/ try to lift my arm", "/ think it will rain", "/ have a toothache". 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, or as I should rather put it: The possibility of an error has been provided for (...). It is possible that, say in an accident, I should feel pain in my arm, see a broken arm at my side, and think it is mine, when really it is my neighbour's. And I could, looking into a mirror, mistake a bump on his forehead for one on mine. On the other hand, there is no question of recognizing a person when I say I have toothache. To ask "are you sure that it's *you* who have pains?" would be nonsensical. Now, when in this case no error is possible, it is because the move which we might be inclined to think of as an error, a 'bad move', is no move of the game at all (...). And now this way of

¹I am not going to engage in a discussion of the theme of solipsism in Wittgenstein. For a reconstruction of it, from *Tractatus* to the *Philosophical Investigations*, cf. Glock 1996, Hacker 1997, Pears 1987.

stating our idea suggests itself: that it is as impossible that in making the statement "I have toothache" I should have mistaken another person for myself, as it is to moan with pain by mistake, having mistaken someone else for me. To say, "I have pain" is no more a statement *about* a particular person than moaning is.²

Wittgenstein seems to be making two points: first, there are cases in which our use of 'I' can be based on a *mis-recognition* of ourselves, and cases in which this is not a possibility.³ Secondly, his examples may seem to suggest that *all and only mental self-ascriptions* are such that there is no possibility of misrecognising the subject they are about, namely oneself.⁴ The first point has

²BIB, pp. 66-67.

³A terminological remark. 'Misrecognition' is not an English word. Moreover, 'recognition' is a success word, i.e. a word which means the result of the successful process of recognising someone/something. However, here 'recognition' is taken in a slightly more etymological sense. That is to say, 'recognition' indicates the result of a process of *re*-cognising something/someone, i.e. a process in which a subject takes someone/something to be the same person/thing she has already been presented with. Now, such a process can go right or wrong. 'Recognition' will be used to indicate the result of a successful re-cognition, whereas 'misrecognition' will be used to indicate the result of an unsuccessful one, i.e. the result of a process in which the wrong object/person is taken to be the same as the one previously encountered. The reason of this terminological choice will become evident in the following when a distinction between identification and recognition will be introduced together with a distinction between misidentification and misrecognition.

⁴It is not obvious that Wittgenstein maintained the view that all and only mental self-ascriptions are immune to error through misidentification (cf. Garrett 1995),

become known in the literature as a distinction between I-judgements which are *liable to error through misidentification* (EM) and those which are *immune to error through misidentification* (IEM).⁵ We will presently see the rationale behind these labels. As for the second point, it states that IEM is a function of the *subject-matter* of the I-judgement. From these two assumptions, i.e. that there is a genuine distinction between EM/IEM for I-judgements and that it is a function of their subject-matter, Wittgenstein quite surprisingly seems to conclude that I-judgements which are IEM are *not about* a particular person.⁶ Both the assumptions and the conclusion need some explaining. In this chapter we will consider the first assumption, i.e. that there is a distinction between uses of 'I' which are EM and those which are IEM, while leaving the second assumption and Wittgenstein's conclusion for discussion in the following chapter.

1. *Recognition-Based Judgements and Error Through Misidentification*

In order to have a better understanding of EM and, consequently, IEM, we must have a grasp of recognition-based judgements, since it seems that the

nevertheless his examples seem to point in that direction. Be that as it may, I am not engaging in an exegesis of Wittgenstein's writings on this topic.

⁵These labels were introduced by Shoemaker 1968, reprinted in Shoemaker 1994a. page reference to the latter. Cf. Shoemaker 1994a: 81.

⁶Garrett 1995 has argued that Wittgenstein does not maintain that 'I' is not a referring expression. Again, I am not interested in offering an exegesis of this passage from Wittgenstein.

phenomenon we wish to explain arises precisely in connection with this kind of judgements. In fact, it is easier to bring out the features of recognition-based judgements by reflecting on judgements based on *misrecognition*. We will introduce an example first, and then we will consider the formal or general features that recognition-based judgements have, and what goes wrong in the case of misrecognition.

Let us suppose that a subject knows *a*, i.e. she is acquainted with her and can tell her apart from other persons and let us suppose that the subject enters a room and sees, sitting in a corner, a person that looks similar to *a*. This person is wearing a white shirt. So the subject thinks [*a* is wearing a white shirt].⁷ However, the person sitting in the corner is not *a*, but someone who looks similar to her. Hence, the subject's judgement is not mistaken as far as the predication component is concerned, but it is mistaken as far as the identification component is concerned. That is to say, although the subject is right in thinking that there is someone wearing a white shirt, she is wrong in thinking that it is *a*.

What we can say about this example is that a judgement affected by misrecognition arises out of circumstances in which the subject is presented with a person and takes *that person* to be *a*, when in fact that person is not *a*. On the basis of what she perceives she forms the judgement [*a* is *F*]. *If* the judgement is false, it is so *at least* because there has been an error in the

⁷Square brackets will be used to indicate propositional contents and concepts as opposed to sentences, names and predicates.

identification of the subject, namely that person is not *a*.⁸ Hence, we can say that error through misidentification occurs as a result of a misrecognition of the object the subject is presented with.

Notice, moreover, that the subject must already have a concept relative to *a* which she applies to *this person*, whom, however, is not *a*. Call this the *identification concept*. It is important to note that, when *recognition* (as opposed to identification proper, as we will see) is at stake, not only must such a concept be a *singular* one, i.e. a concept which is only true of *a*, but also that it must have been acquired by means of a *causal* interaction of the subject with the object.

Moreover, the subject must form a concept of the person/object perceptually given to her ready for recognition. Call this the *base concept*. This will be, characteristically, a perceptual demonstrative concept, i.e. a concept expressible by means of a complex demonstrative expression like 'that

⁸This sentence may be puzzling, but—since we are considering the general case—we have to leave open the possibility that [*a* is *F*] be true, although the subject's judgement is not based on the perception of *a*. So, for instance, it may be the case that *a* is behind the subject in the room and is in fact wearing a white shirt. However, the subject's judgement, based on a mistaken recognition of the person sitting in the corner, would be only accidentally true of *a*, who is in fact wearing a white shirt, but who is not the person presented to her for recognition. Moreover, we can imagine a situation in which the lighting conditions are bad and the person the subject is presented with is neither *a* nor is she wearing a white shirt but a light grey one. In that case the subject's judgement [*a* is wearing a white shirt] would be mistaken both relative to the identification component and to the predication component.

person/object', and formed on the basis of a *demonstrative individuation* of the particular perceptually given to her.⁹ Notice, moreover, that the subject does not know and, in fact, must not know that the object she is presented with and on which her base concept depends is different from *a*. With this piece of terminology, we can now say that a judgement based on misrecognition occurs when and only when the subject wrongly thinks that the base concept and the identification concept are coinstantiated by the object perceptually given to her. Since EM arises in cases of misrecognition, we can represent what happens when a judgement is affected by EM as follows:

$$\frac{[\text{That person is F}] \ \& \ [\text{That person} = a]}{\quad}$$

$$[a \text{ is F}]$$

The first conjunct depends on a *demonstrative individuation* of the object we are presented with, which gives us the base concept [That person] and involves a *predication component* relative to [F]; by contrast, the second conjunct involves an identity judgement to the effect that the base concept and the identification concept are taken to be coinstantiated. This is the *identification component* that leads to the recognition-based judgement [a is F] and it is the distinctive mark of this kind of judgement.

⁹A terminological remark. 'Individuation' will be used in the following to mean the discrimination of a particular, which is characteristically (though not necessarily), perceptually given to one and could be object of demonstrative reference. Hence, our use of 'individuation' has nothing to do with the fact that the subject possesses a *principium individuationis* of the object presented to her.

In fact, a recognition-based judgement could be exhausted by the identification component. For instance, when a subject thinks [That man is NN], she is recognising that man as NN, i.e. she is taking the base concept and the identification concept to be coinstantiated. What, however, is relevant for recognition is the fact that she must have already been *acquainted with NN* and that she is applying the singular concept [NN] as a result of a *new encounter* with him/her. What goes wrong in all cases of misrecognition is precisely the fact that the base concept and the identification concept are wrongly taken to be coinstantiated. Since the error occurs in the identification component, it is natural to consider the resulting judgement [a is F] as affected by *error through misidentification*.

However, it is important to note that [a is F] need *not* be the result of either an explicit or an unconscious inference. Rather, what we are saying is that a judgement such as [a is F] can be considered as recognition-based if it is *causally and rationally triggered* by a belief on the subject's part in the individuation/predication component and by a belief in the identification component, i.e. [There is something/someone which/who is F] and [That thing/person is identical to a]. The reason why the judgement is not seen as actually involving an inference containing an identification component is that it would not be truthful to the phenomenology of experience, nor required by theoretical reasons. Hence, we agree with Peacocke who writes:

Suppose you are able to recognize Michael Dummett (MD); and suppose you come on the basis of your perceptual experience to believe something about MD (...). The fact that in such an example the beliefs that MD is thus-and-so and that this man is MD may well

be acquired simultaneously does not prevent the former from resting on the latter. An experience can cause one to acquire a structure of beliefs simultaneously, a structure in which some beliefs are rationally dependent upon others, beliefs which would be abandoned if those others were abandoned. (...). There is no contradiction in a belief both resting upon an identity in the sense of being causally and rationally sustained by it, even though it did not initially result from a temporal sequential process of conscious inference.¹⁰

So the idea is that the belief in the predication and in the identification component need not be actually entertained in a “sequential process of conscious inference”. However, these beliefs must be *causally* and *rationaly* efficacious, i.e. it is *because* they are (maybe not occurrently) held that the subject makes the relevant judgement and, moreover, they constitute the *subject's rational grounds* for the judgement.

To sum up: what is distinctive of EM recognition-based judgements is precisely the fact that the identification concept has been acquired through an encounter with the object, and is deployed on the basis of a perceptual encounter with an object which, however, is different from the one which was at its origin. In other words, if the identification concept had not been acquired through an encounter with its object (demonstrative individuation) and/or it were not applied on the basis of a new perceptual encounter with an object (recognition), then the judgement [a is F] would not count as a recognition-based one. To put the point rather crudely: there would not be *re*-cognition if there had not been an individuation of the object in the first place and a new encounter with it (or with another one which looks sufficiently similar to the one

¹⁰Peacocke 1983: 143-144.

at the origin of our singular concept) later on. It has to be stressed that while these constraints on the conditions in which the judgement is formed must be met in order to have a recognition-based judgement that can eventually be affected by EM, this does not mean that the actual judgement should (necessarily) be considered as the result of an inference involving the two components. The idea is just that a new encounter with an object, plus memory traces left from a previous encounter with a possibly different object, can trigger the recognition which, in a conceptually endowed creature, immediately issues in the relevant judgement. This, in turn, is causally and rationally sustained by a (not necessarily occurrent) belief in an identification component and in a predication component.

We are now in a position to characterise error through misidentification as follows:

A singular judgement of the form [a is F] is affected by EM, relative to the subject, if the thinker has the singular concept [a] of an object *a* and applies [a] to the distinct object *x*, as a consequence of a misrecognition.¹¹

It is important to stress that a subject would be having a thought about *a*, which she mistakenly thinks to be the same as the object she is presented with, i.e. *x*, which she does not know to be different from *a*. Hence, when the

¹¹This characterisation will be revised shortly and this is why it only states a sufficient condition for error through misidentification.

judgement [a is F] is affected by EM, it must be considered as *about a*, which, however, is not the same as the object the judgement is actually *based on*.¹²

1.1 EM. A Comparison Between Definitions: Shoemaker and Pryor

In order to appreciate the consequences of our definition of EM, it is useful to compare it with the definitions offered by Shoemaker, who was the first to draw systematic attention to EM and IEM, and by Pryor who, to an extent, holds similar views to ours. Our definition of EM is compatible with Shoemaker's:

To say that a statement 'a is ϕ ' 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 ϕ , but makes the mistake of asserting 'a is ϕ ' because, and only because, he mistakenly thinks that the thing he knows to be ϕ is what "a" refers to.¹³

Obviously Shoemaker's definition should be amended in such a way as to account not only for sentences but also for *judgements* which are EM, and for the fact that we can *falsely* believe a to be ϕ , i.e. EM is compatible with there

¹²By contrast, Rovane (1987) and Christofidou (1995) seem to think that if a judgement is affected by EM, then it is about the object which is actually given to us in the particular episode in which recognition occurs. For a discussion cf. Ch. 2.

¹³Shoemaker 1994a: 82.

being also an error in the predication component.¹⁴ The reason it is preferable to consider judgements instead of sentences is that there are things we do have actual beliefs about but that we do not assert. Hence, talk in terms of actual belief or judgement covers this broader case. Moreover, it is customary to distinguish between actual beliefs and dispositional ones on the basis of whether or not the subject is actually entertaining a given belief. We will use 'judgement' to indicate actual belief and 'belief' to indicate dispositional belief. The reason EM is taken to affect judgements is that we are interested only in those actual beliefs the subject comes to entertain and, possibly, express. By contrast, we have maintained that the reason she entertains these mistaken judgements is her belief—which need not be occurrent—in a mistaken identification component, brought about, in turn, by an error in recognition due to a shortcoming of the subject's recognition system.¹⁵

According to Pryor, the kind of misidentification our definition highlights is labelled "*de re* misidentification". The salient feature of such a kind of misidentification is that it depends on misrecognition and hence on the fact that the subject has been acquainted with an object but, on a new encounter.

¹⁴A similar amendment to Shoemaker's definition can be found in Peacocke 1999, 269, although he opts for beliefs and thought constituents or modes of presentations, rather than for judgements and concepts.

¹⁵A description of human recognition abilities falls beyond the scope of this thesis. Notice, moreover, that there can be dispositional beliefs which are affected by EM, but we will concentrate only on actual ones.

mistakenly takes a different object to be the same as the one once presented to her. Pryor defines *de re* misidentification as follows:

[W]e have a case of *de re misidentification* whenever the following three conditions obtain:

- (i) There is some singular proposition about x , to the effect that it is F , that a subject believes or attempts to express. (...).
- (ii) The subject's justification for believing this singular proposition rests on his justification for believing, of some y , that y is F and that y is identical to x . (...).
- (iii) However, unbeknownst to the subject, $y \neq x$. (...).

I do not assume that whenever your justification for believing one proposition rests on your justification for believing other propositions, you will have *formed* or even *entertained* a belief in those latter propositions. Nor do I assume that you have *based* the first belief on a belief in those latter propositions.¹⁶

However, Pryor's considerations with regard to *utterances* affected by error through misidentification are slightly different from ours. In particular, he distinguishes between levels of *referential intentions*. Hence, he argues that *by referring* to the object perceptually given to us, we *aim* at referring to the semantic referent of the singular term (or what we take to be the semantic referent of the singular term). Consequently, our *basic* referential attempt makes us refer to the object perceptually given to us, whereas our *secondary* referential attempt is directed towards the semantic referent of the singular term (or what we take to be the semantic referent of the singular term). He then concludes that in the utterance the subject says something *about* the object perceptually given to her. But, as it is evident from Pryor's definition of *de re*

¹⁶Pryor 1998: 274-276.

misidentification, the *judgement* which is affected by EM and leads to an utterance which is, in turn, affected by EM, is *about* the semantic referent of the singular term (or what we take to be the semantic referent of the singular term).¹⁷

By contrast, since our starting point is the judgement, rather than the utterance, we agree with Pryor's claim that the utterance is about the semantic referent of the singular term used to express it, but we disagree with respect to his interpretation of the utterance as being *also* about the object perceptually given to one. For this would mean that one's utterance corresponds, at one and the same time, to two different propositions or, in our terminology, judgements that could also have different truth-values.¹⁸ Since it is not clear at all how this could happen, it seems safer to hold that the utterance just reflects the judgement. This, in turn, is affected by EM, which means that it is based on two

¹⁷Cf. Pryor 1998: 273-275. In particular, example 2, p. 275. A similar position is held by Shoemaker 1994a: 83: "Suppose that I am selling neckties, that a customer wants a red necktie, and that I believe I have put a particular red silk necktie on a shelf of the showcase that is visible to the customer but not to me. Putting my hand on a necktie on that shelf, and feeling to be silk, I might say 'This one is red'. Here it could be said that I have identified, correctly or incorrectly, the object I refer to in saying 'this' as the object 'I have in mind' (...). [In this case] I intend to refer to a certain red necktie I believe to be on the shelf, but there is also a sense in which I intend to refer, and do refer, to the necktie actually on the shelf (...)".

¹⁸In fact, supposing that the predication is correct, then, if the judgement is about the object perceptually given to one, it is true; whereas, if it is about the semantic referent of the singular term, it is false.

(not necessarily occurrent) beliefs, i.e. one in a predication/individuation component, triggered by the presence of an object, and one in a mistaken identification component, triggered by our recognition-system. Hence, in the standard case of an utterance affected by EM, the utterance (as well as the judgement) is *about* the semantic referent of the singular term but it is *based* on the object perceptually given to one, which is in fact different from the semantic referent of the singular term. This point will become important in the following (cf. § 3), where we will distinguish the phenomenon of EM from the phenomenon of the split between semantic reference and speaker's reference. There we will argue, amongst other things, that, in the latter case, the utterance, like the judgement, is *about* the object perceptually given to one. Hence, the difference between utterances which are affected by EM and those which just present a split between semantic reference and speaker's reference is that the former ones are about the semantic referent of the singular term, whereas the latter ones are about the object perceptually given to one, which is the speaker's referent of the singular term.

2. Identification-Based Judgements and Error Through Misidentification

We should now consider a further possibility in which error through misidentification is not supposed to affect a recognition-based judgement, i.e. a judgement in which the identification concept has been acquired through a causal interaction with the object; rather, it affects an identification-based judgement, i.e. a judgement in which the identification concept has *not* been

acquired through a causal interaction with the object. We will introduce this case with an example first, and then we will discuss the general features of it.

Suppose that a subject is at a conference and in the audience there is a woman who is very active in the discussion and is clearly an expert on the topic. Given the present context and the subject's background knowledge (she has read books written by her, she has a rough idea of her age and physical appearance, through testimony and inference, etc.), she makes a guess, and thinks [That woman is NN]. However, the woman is not NN, but someone else. Here we have a base concept [that woman] and an identification concept [NN]. In symbols, we can represent what happens thus:

[That woman = NN]

In this particular case we have only the identification component, but what we will be saying in the following does not rest on that, and a predication component could be added without any harm to the considerations we are going to make. The constituents of the identification component are a demonstrative concept (the base concept), expressible by 'That/this woman' and the identification concept expressible through either a proper name or a definite description, e.g. 'NN' or 'the such-and-such'. What is important to note is that the identification concept which is a singular one, i.e. a concept which is true only of a specific referent, *has not been acquired through a perceptual interaction with the object/person of which it is a concept*. In other words, it is not itself the result of a demonstrative identification. Hence, the judgement [That woman = NN] is not the expression of a recognition, in the sense we have characterised it in § 1. Rather, it is an expression of an *identification*, i.e. the

subject takes a particular to be an instance of a (singular) concept she has acquired other than through a previous encounter with its only instance. What identification-based judgements have in common with recognition-based ones is that they involve the exercise of two (singular) concepts, but while in the latter case both concepts depend on a perceptual identification of the object, in the former case only one does.

Notice, moreover, that this is not to say that proper names work like definite descriptions. For, in our example, the reference of 'NN' is fixed through a body of information the subject has about NN, but its semantic role is to refer to NN and not to any person who, in a logically possible world, satisfies the definite description(s) used to fix the reference of 'NN'.

Now consider the case in which the subject forms the judgement [NN is a fascinating person] on the basis of how she sees the woman she takes to be NN behave during the conference. In such a case we can represent what happens thus:

[That woman is a fascinating person] & [That woman = NN]

[NN is a fascinating person]

However, although there is someone who is a fascinating person, that someone is not NN. Thus, it seems as if, *prima facie*, the judgement [NN is a fascinating person] is affected by EM.

Yet, if we accept the possibility of a split between semantic reference and speaker's reference and, hence, that 'NN' can mean NN for the community and NN* for a particular speaker, then it is no longer obvious that error through

misidentification has occurred. After all, for the speaker 'NN' designates the person NN* who is perceptually given to her. So, we have a choice here: either we consider the judgement [NN is F] as affected by EM, but then we cannot allow for the possibility of a split between semantic reference and speaker's reference—for otherwise our identification component would not be mistaken, since, for the speaker, 'NN' means that woman (NN*) perceptually given to her: or else we think that there can be a split between semantic and speaker's reference, but then, by the same token, we cannot say that [NN is F] is affected by EM. In order to make a decision, however, we have to be clear about the difference between error through misidentification and the possibility of a split between semantic reference and speaker's reference. We will turn to this in § 3. In the mean time let us dwell on Pryor's account of EM of what we have called identification-based judgements.

2.1 *De Re Misidentification and Which-Misidentification: Pryor*

According to Pryor, we should distinguish between two *sorts* of error through misidentification (and consequently between two sorts of immunity to error through misidentification). On the one hand, there is *de re* misidentification, which we mentioned in § 1.1 and which is similar, although not identical, to our understanding of EM. On the other hand, there is what Pryor calls *which*-misidentification. The characteristic feature of this latter kind of EM is that *there is no recognition* of a particular involved, because the subject does not have an

identification concept formed on the basis of a previous encounter with an object. Pryor considers the following example:

I smell a skunky odor, and see several animals rummaging around in my garden. None of them has the characteristic white stripes of a skunk, but I believe that some skunks lack these stripes. Approaching closer and sniffing, I form the belief, of the smallest of these animals, that it is a skunk in my garden. This belief is mistaken. There is a skunk in my garden, but it is not the small animal I see.¹⁹

Pryor then defines *which*-misidentification as follows:

[Which-misidentification] occurs when:

- (i) A subject has some grounds G that offer him knowledge of the existential generalization $\exists xFx$.
- (ii) Partly on the basis of G , the subject is also justified, or takes himself to be justified, in believing of some object a that *it* is F .
- (iii) But in fact a is not F . Some distinct object y is F , and it's because the grounds G "derive" in the right way from this fact *about* y that they offer the subject knowledge that $\exists xFx$.²⁰

Hence, according to Pryor, we should realise that there can also be EM *when there is no mistaken identification component*. Consequently, according to Pryor, we should realise that there are *two altogether different kinds* of EM, i.e. *de re* and *which*-misidentification.

However, this conclusion is mistaken and brought about by a misleading way of interpreting the "ingredients" that can occur in an identification component. For, as we saw in § 2, an identification component can also involve, besides an indexical such as 'I', a name, and, for that matter, a (rigidified) definite

¹⁹Pryor 1998: 281.

²⁰Pryor 1998: 282.

description, which the subject has not acquired through a perceptual encounter with a given object. If we allow for this possibility, then Pryor's example too would involve an identification component. For it should be construed as follows: [There is a skunky odor], [This (small animal) = the animal which is the cause of the skunky odor I can smell], [This is a skunk], therefore [There is a skunk in my garden]. Here there is an identification component, i.e. [This (small animal) = the animal which is the cause of the skunky odor I can smell], which, however, is mistaken because the small animal I see is not the animal responsible for the skunky odor I can smell. Moreover, there is a mistaken predication because the animal I can see is not even a skunk.

Consequently, it does not seem right to conclude that there exist *two different kinds of EM* (and consequently of IEM). Rather, *there exist different kinds of identification components*. Namely, those which involve the exercise of two concepts which are both based on an encounter with the object(s) and those which do not (they can either involve one concept which is dependent on the subject's encounter with the object, or none at all). Hence, the generalised definition of EM is as follows:

A singular judgement of the form [a is F] is affected by EM, relative to the subject, iff the thinker has the singular concept [a] of an object a and applies [a] to the distinct object x as a consequence of either a misrecognition or a misidentification.

Here are some examples that should help to clarify our definition:

(a) [Kripke is wearing a white shirt];

(b) [Kripke is very active in the discussion];

(c) [Kripke is a knight].

(a) The subject is acquainted with Saul Kripke, she sees someone who resembles him and judges [S. K. is wearing a white shirt]. Here the mistaken identification component is [That person = S. K.] and both concepts depend on a causal (i.e. perceptual) connection of the subject with the object. Here EM arises because of *misrecognition*.

(b) The subject is not acquainted with S. K. But she is at a conference which she knows S. K. may attend as well. She sees someone who is very active in the discussion and forms the judgement [S. K. is very active in the discussion]. Here there is a mistaken identification component for it is not true that [That person = S. K.]. However, while [that person] is a singular concept the subject has because of her acquaintance with a certain individual, [S. K.] is a concept she has acquired otherwise than through a perceptual encounter with S. K. Here, then, EM arises because of *misidentification*.

(c) The subject is not acquainted with S. K. However, she thinks that S. K. is the author of *The Logical Basis of Metaphysics* and she knows that the author of LBM is a knight. Hence, she forms the judgement [S. K. is a knight]. Here the mistaken identification component is [S. K. = the (actual) author of LBM], where neither concept has been acquired by the subject through a perceptual encounter with either S. K. or the (actual) author of LBM. Here EM arises because of *misidentification*. Notice, however, that all these concepts can be considered as singular ones, provided that the definite descriptions be taken as *rigid*.

3. *Semantic Reference and Speaker's Reference vs. Error Through Misidentification*

The characterisation of EM proposed in the previous section should help us avoid the mistake of thinking of EM as equivalent to the *split between speaker's reference (SPR) and semantic reference (SR)*, which can arise only in cases of *referential* readings of singular *sentences*. If, by contrast, the two phenomena were one and the same, then we could easily explain the impossibility for *all* comprehending uses of 'I' to be affected by EM. For that would be a straightforward consequence of the fact that if a speaker masters the use of 'I' according to the token reflexive rule—any token of 'I' refers to its producer—then she cannot use 'I' to refer to someone else. Hence, her use of 'I' could never exhibit a split between SR and SPR and will always be IEM (similar considerations would apply to uses of the first person in thought).²¹

However, in order to clarify the difference between the split SR/SPR and EM, let us dwell a little on the former. Consider the following case, in which such a split is generated in the use of the proper name, despite the fact that most cases by means of which the split between SR and SPR is presented contain definite descriptions.²² A subject is at a conference sitting next to a friend of

²¹As we shall see in the next chapter, Christofidou 1995 maintains, on these grounds, that all uses of the first person are immune to error through misidentification.

²²Cf. Donnellan 1966, Kripke 1979. I actually think that the split between SR and SPR can be shown also in the case of complex demonstratives. Suppose I say 'That star is very bright', pointing to a luminous point in the sky which is in fact Venus. My

hers. The speakers are A and B, but the labels with their names have been swapped. So she says to her friend 'A is wearing a nice jumper', meaning the person who is behind the label with the name 'A' and her friend agrees with her. However, the subject is talking, and her friend understands the subject's *utterance* as being about the person who is in fact called 'B'. Thus, the subject's utterance is true, in a sense, but only if we read it referentially, i.e. as about B. i.e. about the person who is given to the audience *independently* of the name by which she is called. For the name 'A', which features in the sentence uttered by the subject, is just a means at the level of language to refer to the intended object, which, however, is demonstratively given at the level of thought.

Thus, the subject's *judgement* is in fact correct, being a singular judgement about an object demonstratively given to her. By contrast, it is the *sentence* that she uses to express it which is not appropriate because it features in the subject's position a name whose semantic reference is different from its speaker's reference. So, cases of split between SR and SPR are those which, once analysed, present a *dichotomy between the sentence and the judgement*. In symbols:

judgement could be true, although Venus is a planet. I could actually manage to make my audience understand my sentence and make them agree with me. But obviously the semantic reference of 'that star' would be none at all, whereas the intended reference or its speaker's reference would be Venus. Hence, it is just because 'That star' serves the only purpose of purporting reference to the intended object which is given independently of the expression I choose to use that I can manage to say

Judgement: [That person is F]

Sentence: 'A is F'

Notice, however, that if the context did not provide the subject with a name for the object she has individuated, her judgement would be fine in this case too. since she would have just thought, of the object perceptually given to her, that it has the property *F*. In symbols:

[That person is F]

In other words, if we abstract from the means of expression which can be supplied by the context, and we express the judgement the subject is making, we would express it thus: 'That person is F', which, if false, could only be false because the property *F* does not apply to the person perceptually given to the subject. That is to say, in our example the subject would be entertaining a *de re* judgement about that particular *x* which is given to her and to which she could refer in thousands of different ways, and her judgement could be false iff the property *F* were not true of that object.

However, one could object that, really, it is not only the case that the subject just says that *A is F*. Rather, she also comes to think that the person perceptually given to her is *A*. According to this point of view, at the level of judgement, the situation can be represented as follows:

[That person is F] & [That person = A]

[A is F]

something true of Venus. Thus, my sentence has to be read referentially and not attributively.

Hence, in this case there would be a mistaken identification component (and hence EM), which, in turn, would make it the case that the speaker's reference and the semantic reference of 'A' be different.

Yet, no matter how complicated things can get in actual life, it has to be stressed that in *basic* cases of a split between SR and SPR, the subject *intends* to refer to *the object perceptually given to her* and her judgement is about that very object. By contrast, when EM arises, as we saw, the subject *intends to refer to* and is forming a judgement about (what she takes to be) *the semantic referent* of the singular term she may use to give expression to her judgement.²³

To see this more clearly, consider that in at least some cases of a split between SR and SPR, if the subject were told that the person in question is not A, she could substitute 'A' which features in her sentence with a different means of expression obtained from the context, such as 'The person sitting behind the label A' or 'That person on the right of the person behind the label B', etc. But no such substitution could occur in the expression of the judgement [A is wearing a white shirt] which is affected by EM. This is so because the subject neither does nor wants to talk about the very person, whoever she is, who is perceptually given to her at the moment in which her judgement occurs, to which she could refer in thousands of different ways. On the contrary, the subject intends to talk about A whom she takes to be as the same person as the one perceptually given to her. In other words, contrary to most cases in which

²³Although EM and the split between SR and SPR can be combined, we are considering simple cases in order to make clear the nature of different phenomena.

there is a split between SR and SPR, the subject could not manage to express correctly her judgement by saying 'That person is F', because she is not aiming at *individuating* that person but at *recognising* the person perceptually given to her as a particular person A she has already been acquainted with.

4. *Conclusions*

The considerations of the previous section should help us make a decision about the example we introduced in § 2 (e.g. someone sees a person at a conference and takes her to be NN) showing either that EM is possible when no recognition is taking place or that a split between semantic and speaker's reference is possible, but then EM will not be occurring. In fact, it seems rather obvious that the ambiguity cannot be completely dispelled save by reference to the *intentions* of the speaker. If it is her intention primarily to talk about that very person perceptually given to her, then EM does not occur and we are just presented with a split between SR and SPR. By contrast, if the speaker wants to talk about NN, then, if the person presented to her is not NN, there is an error through misidentification, because the speaker would not substitute 'NN' with any other means of expression. Moreover, there would be an error through misidentification, although no recognition proper is taking place, because we have allowed for error through misidentification to occur also when the identification component is not itself the result of a recognition and involves singular concepts which have been acquired other than through a causal interaction with their referents. Hence, we see that referential intentions are

relevant to the determination of the *content* of one's judgement. However, once the content of one's judgement has been fixed, then whether or not the latter is affected by EM is solely a function of the *grounds* on which the judgement is based, i.e. of whether it is based on a wrong identification component.

We can now summarise all the relevant possibilities and their components as follows:

	Concept/Judgement	Error (Description)	Kind of Error (Name)
Individuation	[That person/object] The singular base concept is acquired through the perceptual demonstrative individuation of the particular.	No error is possible	
Predication	[That person/object is F]	The object/person does not have the property <i>F</i> .	Mistake of predication
Identification-based judgement	[That person/object = A] There is an identification concept [A], but this is not a concept acquired through acquaintance with its referent.	The base concept and the identification concept are not coinstantiated.	Both EM and a split between SR and SPR are possible, depending on the intentions of the speaker.
Recognition-based judgement	[That person/object = A] The identification concept [A] is acquired through acquaintance with its referent.	The identification concept and the base concept are not coinstantiated.	Error Through Misidentification.

Let us now bring our considerations to bear on the case of the first person. As noted above, if EM were equated with the split between SR and SPR, then we could explain why IEM occurs, relative to 'I' and other "automatic" indexicals

such as 'now' and most uses of 'here'.²⁴ For, if a subject masters the use of 'I', specified by the token-reflexive rule, then she cannot use it to name someone different from herself. In other words: it is the mark of the competent use of 'I' that a speaker does not use it to refer to someone else. However, in such a case *all* comprehending uses of 'I' would be IEM. So we would have lost the alleged genuine distinction between uses of 'I' which are EM and those which are IEM. To state it clearly, if EM is equated with the possibility of a split between SR and SPR, and IEM with its impossibility, then, given the *reflexivity* of 'I', all (competent and sincere) uses of 'I' would be IEM and, thus, we could not draw a genuine distinction between uses of 'I' which are IEM and those which are not.

However, if we distinguish between EM and the split between SR and SPR in the way proposed, then we can see how it is possible for some I-judgements to be affected by EM. Consider the case in which I am seeing a reflection in a mirror and on the basis of it, I form the judgement [I am getting fat]. Luckily for me, the person reflected in a mirror is someone else. So there is a chance that my judgement be false: I have mistaken someone else for me. Since my judgement is not well grounded, it is possible that in fact I am not getting fat, but someone else is. In this case, although [I] refers to myself and there is no split between SR and SPR, the *grounds* on which my I-judgement is based involve a mistaken identification component. Consequently, the judgement is affected by EM.

²⁴Cf. Perry 1997 for a discussion of "automatic indexicals". In particular, Perry 1997: 595-598.

We should describe this as a case in which there is the individuation of a person x , whom I take to be myself, when in fact it is someone else. I am unaware of that and I think about myself, although I mistakenly think I am the same as the person who is perceptually given to me in the reflection. The case can be represented as follows:

$$\frac{[\text{That person is } F] \ \& \ [\text{That person} = I]}{[I \text{ am } F]}$$

Using our terminology, we can say that the base concept given through the individuation component is fine, the predication may be correct, but the identification component goes wrong, because the subject mistakenly takes the base concept [That person] and the identification concept [I] to be coinstantiated. However, both the use of 'I' in the sentence which may give expression to the judgement and the use of [I] in the judgement would be correct: they do refer to the subject. Once more, it is important to note that this rules out the possibility of equating IEM with the impossibility of a split between SR and SPR. For in the example just taken into account, the latter split does not occur and yet EM arises. This concludes the discussion of EM in general and of EM relative to I-judgements; what we must still explain, however, is the possibility for I-judgements to be IEM.

Chapter Two

Immunity to Error through Misidentification at the Origin of the Trilemma About the Self and the Forms of Its Solution

In the previous chapter we saw how it is possible for some first-person judgements to be affected by EM. However, we must still explain how it is possible, at least for some self-ascriptions, to be IEM. In this chapter we will consider one possible explanation, which, however, will expose us to a trilemma about the self. To anticipate a little: either we have an explanation of IEM of mental self-ascriptions, in which case we are committed to either Cartesianism or Idealism about the self; or else we maintain that the self is identical with a human being, to whom the first person refers, but we do not have an explanation of IEM of mental self-ascriptions. Finally, we will consider some general features common to most solutions to the trilemma and we will offer a taxonomy of them, which shall give us a scheme of how to proceed in the following.

1. *The First Definition of Immunity to Error Through Misidentification*

Our problem is: how can it be that in the case of the first-person it seems impossible, at least in some cases, that there could be EM? In order to see why, let us consider what it would be for a singular judgement to be IEM in general. To such an end we must go back to the definition of EM.

A singular judgement of the form [a is F] is affected by EM, relative to the subject, iff the thinker has the singular concept [a] of an object *a* and applies [a] to the distinct object *x*, as a consequence of a misrecognition or a misidentification.¹

IEM would thus be defined as follows:

IEM_{diff} A singular judgement of the form [a is F], formed on the basis of a recognition or an identification, is IEM, relative to the subject, iff it is not possible for the thinker to apply the particular singular concept [a] to any other object but the relevant object *a*.

The motivation behind this definition of IEM is twofold: on the one hand, it is *prima facie* plausible to hold that—as the label suggests—immunity to error through misidentification can occur just in the case where *there is an identification* (or a recognition) which, however, cannot go wrong. On the other hand, singular judgements have often been considered to involve an identification or recognition of a particular. According to such a model, when

¹All the specifications and constraints introduced in Ch. 1 still apply here.

one makes a singular judgement one takes a particular to be an instance of a given singular concept [*a*]. Hence, if singular judgements involve either the identification or the recognition of a particular, then the possibility for the judgement to be IEM will naturally be understood as due to the fact that it is impossible for the subject to misrecognise the object which is given to her for recognition or identification.²

Given our terminology, we can read the above-mentioned definition of IEM as saying that IEM is guaranteed iff in the particular episode of recognition there is no possibility of wrongly taking the base concept and the identification concept to be coinstantiated. But when is it that we can guarantee that such a misapplication is impossible? It seems that, if recognition occurs at all, then IEM depends on the fact that the *object* individuated and which grounds our base concept cannot be mistaken for any other. Consequently, the nature of the object we are presented with and the conditions in which it is given to us must be such that no misrecognition or misidentification is possible.

²We will criticise this model at length in the following chapters. As we will presently see, the Cartesian may hold that when one self-ascribes a mental property one is in fact presented with one's self, i.e. the mental substance one is, which one identifies as oneself. In this case 'I' would function in a way which is similar to (at least some) uses of proper names, where an object presented to the subject is identified as NN.

2. IEM of I-Judgements and the Dilemma About the Self

As we saw in the previous chapter, according to Wittgenstein, there are some I-judgements that are IEM. Consider the recognition-based I-judgement [I am F]:

$$\frac{[\text{That person is F}] \ \& \ [\text{That person} = \text{I}]}{[\text{I am F}]}$$

Given our definition of IEM, it will be IEM iff there is no possibility of applying the identification concept [I] to the wrong person.

Now, the problem seems to be this. If a judgement is IEM, given IEM_{df1} , this means that it is based on an *infallible recognition* of the object thought about. However, we are never infallible in recognising physical objects. That is to say it is always possible for us to take the wrong object to be an instance of a certain identification concept. Indeed, it would be wishful thinking to believe that even in the case of objects and people we are most fond of or have long been acquainted with, we could be infallible in recognising them. That is to say, it is always possible, given infelicitous conditions of either the environment, or the subject or both, that the latter goes astray in recognising a certain object or person. Hence, if IEM is understood along the lines of IEM_{df1} , it must be the *peculiar nature* of the object thought about and the particularly *advantageous conditions* of our encounter with it that will guarantee that our judgement is IEM.

This position would square with Nozick's methodological claim that only a particular account of the *nature of the object thought about* could explain "why.

when we reflexively self-refer, we know it is *ourselves* to which we refer”.³ We will not discuss Nozick’s proposal. However, it is relevant to our purposes because it is symptomatic of an explanatory strategy of the phenomenon of IEM, which, in our view, would lead quite naturally to the Cartesian and the Idealist positions about the self; namely the strategy according to which an account of IEM has to be looked for in a suitable account of the *metaphysics of the self*. In other words, if IEM_{df1} holds, then IEM will be a phenomenon whose explanation is essentially *metaphysical*. Let us now see why assuming IEM_{df1} would lead to either Cartesianism or Idealism about the self.

According to the Cartesian ontology, there are two sorts of substance physical and mental. In the case of the first person we are presented in thought with our self, i.e. the mental substance we are, and we cannot be presented with anything but our self. For we cannot be introspectively aware of someone else’s self. Moreover, the fact that the self is presented in introspection guarantees that we are infallible in recognising it, because, according to the Cartesian conception of the mind everything that happens in it is transparent to the subject. Consequently, the Cartesian seems to be *prima facie* able to

³Nozick 1981: 90. As a positive claim about the nature of such a self he proposes the following: “the I is delineated, is synthesised around ... [the] act of reflexive self-referring. An entity is synthesised around the reflective act and it is the ‘I’ of that act” (Nozick 1981: 87). Hence, if the entity ‘I’ refers to is synthesised around the very act of reflexively self-referring, then there is no possibility for one to misidentify the entity one is referring to by using that very token of ‘I’. I thank John Campbell for drawing my attention to Nozick.

explain IEM of mental self-ascriptions, by making reference to the particular nature of the self and the peculiar conditions in which it is encountered.

By contrast, in the case of bodily self-ascriptions, a Cartesian will make sense of the possibility that they are liable to EM by saying that in those cases the subject will be misrecognising *her body*, i.e. the body to which her thinking self is connected and which is hers in virtue of being that body in which her thinking self is located, which can be cause of sensations, which is respondent to her will, etc. Using our terminology we can say that EM arises because of a perceptual encounter with a bodily entity which the subject wrongly takes to be an instance of the identification concept [my body].

However, there are, of course, huge problems with the Cartesian dualist metaphysics: how do the mental and the physical substance connect and interact? But, more importantly, what are the conditions of identity of the thinking self, so that it makes sense to say that in an episode of recognition, the subject is recognising that self as the same she encountered in another thought episode? Finally, the driving thought against Cartesianism about the self is that, in a given thought episode, the subject is not really presented with an object, i.e. she does not experience a mental substance at all. This is indeed Hume's point glossed by Lichtenberg by saying that from the *Cogito*, Descartes should have inferred that there is thinking, i.e. a conscious mental episode, but not that there is a substance which thinks. Hence, it is quite mysterious how the Cartesian could *actually* provide an explanation of IEM, given that, on closer reflection, her account of the nature of the self is deeply problematic.

The reaction to the Cartesian metaphysics of the self and its explanation of IEM motivated Wittgenstein's position in the *Blue Book*. For Wittgenstein endorsed Lichtenberg's objection to such an extent that, in the passage from the *Blue Book* we quoted in the previous chapter, he concluded by saying that first-person mental self-ascriptions are *not about* a person at all. The thought is as follows: on the one hand, no physical object is such that it cannot be mistakenly recognised and, on the other hand, Cartesian substances are myths. Yet, some mental self-ascriptions are IEM and this is a fact in need of explanation. Hence, the proposal is to deny that there is a self given for recognition when mental self-ascriptions are made. Thus, if *no object at all* is given for recognition, in the circumstances envisaged, then the first-person judgement cannot be based on a misrecognition. Consequently, it will be IEM *by default*.

The case is different for bodily self-ascriptions. For the subject can be presented with a different body, form the corresponding base concept, and take this and the identification concept [my body] to be coinstantiated, thus leading to a judgement that is EM. A consequence of this view, therefore, is that 'I' has a double meaning: when it is used as object, it is an expression that refers to *this person*; whereas, when it is used as subject, it is not. That is to say, not only does it fail to refer to this person, but it also fails to be a *referring* expression altogether. This, however, given Wittgenstein's overall position about language, is not to say that the first person pronoun has no *meaning*, for it has got a *use*. but simply that it is not used to *refer* to anything.

However, Wittgenstein's claims are problematic. In particular: how do we interpret first-person sentences containing a psychological predicate in the present tense, like 'I am in pain'? If 'I' is not referring to a person, then it is not a referring pronoun altogether. But this position has bad consequences. For one thing, from 'I am in pain' the existential generalisation 'Someone is in pain' would not follow. Moreover, it would be impossible to explain the truth-value link between 'I am in pain' and 'AC is in pain'. For we could not say that one of these sentences is true (or false) iff the other sentence is true (or false) because they are *about* the same subject. For 'I am in pain' would not be a sentence about a subject at all. Moreover, shall we conclude that mental states have no owner? Would that be tantamount to saying that mental states are floating around, without a subject who owns them? These are questions that would require an exegesis of Wittgenstein's views on the self, which falls beyond the scope of this work. However, it seems reasonable to hold that his views could, at least *prima facie*, be open to these quite problematic developments. Still, it is also important to note that Wittgenstein's views in the *Blue Book* are mainly motivated by a reaction against the Cartesian metaphysics of the self.⁴ If this is the case, then, whatever their developments might be, they would primarily be a symptom of the fact that the Cartesian metaphysics of the self is an attempt at solving theoretical problems, where the alleged solution, instead of solving them, creates new ones. Unfortunately, the Wittgensteinian response and its possible developments would not leave us in a better position.

⁴Cf. Blb, pp. 63, 69.

Consequently, on the view that IEM of mental I-judgements should be explained by reference to the nature of the particular object we are presented with in an episode of recognition and by reference to the peculiar conditions in which the recognition occurs, we seem to reach a dilemma: either 'I' refers to a Cartesian ego, or it does not refer at all. We will call these horns of the dilemma the Cartesian and the Idealist positions about the self. While the former label is clear enough, since it names the stereotype of the Cartesian position about the self, the latter label is slightly more mysterious. The reason why we are adopting it in the present context is that it conforms to Evans' terminology we will consider in the next chapters. Moreover, by using this label we will avoid the trouble of engaging in a close exegesis of Wittgenstein's position. Be that as it may, by "Idealist position about the self" we will understand the position according to which mental self-ascriptions are IEM because no self at all is given for recognition. Natural developments of this view include the no-reference view about 'I' and the no-subject view about mental self-ascriptions.

We can schematise the dilemma as follows:

	Self	IEM of mental self-ascriptions	Consequences
Cartesian	Salient object: a thinking substance which is present to the subject for each act of thought.	Given the nature of the self and the conditions of our encounter with it, we cannot mistakenly recognise it.	The self is essentially a thinking substance. 'I', when used as subject, refers to it. BAD METAPHYSICS
Idealist	No object	Since there is nothing to recognise, our use of 'I' as subject is IEM by default.	There is no self. 'I', when used as subject, is not a referring expression. BAD SEMANTICS and possibly BAD METAPHYSICS as well.

3. IEM of I-Judgements and the Trilemma About the Self

What we have seen so far is that under a possible model of explanation of IEM of mental self-ascriptions, according to which it should be explained by reference to the particular nature of the object encountered and the conditions in which the encounter takes place, we are confronted with a dilemma. That this, as it stands, is a real dilemma, seems to be indisputable: it is hard to see why one would be willing to live with the consequences of either of its horns. However, one may think that after all the dilemma does not really hold, because it is possible to provide an explanation of IEM which is compatible with the fact that 'I' refers to *this person*, i.e. a primitive entity with both physical and psychological properties, such as a human being.

To anticipate a little, the answer to this claim is that if we share the Idealist and the Cartesian assumption that non-inferential mental self-ascriptions are IEM—where IEM is understood along IEM_{df1}—then the appeal to the fact that the self is embodied is not going to provide an explanation of IEM. Let us expound on this.

We have seen that in the case of the Cartesian position, the nature of the self, together with the conditions of the encounter with it, can offer a *prima facie* explanation of why psychological I-judgements are IEM (according to IEM_{df1}). Similarly, according to the Idealist view, the absence of an object to recognise can explain why uses of 'I' as subject are IEM (according to IEM_{df1}). The question is now this: can the fact that the subject is an embodied entity explain, *as such*, the fact that when she judges [I think], or [I am in pain], she cannot misrecognise herself and thus form an EM judgement?

In fact, if one holds on to IEM_{df1} the mere appeal to the fact that the subject is a person is a complete non-starter. For, although this view would be preferable both at the metaphysical level and at the semantic one, it would not rule out the possibility that, in making a recognition, the subject could take another person to be an instance of the first-person concept. In fact, the reverse is true, namely the appeal to the fact that the subject is person can perfectly well explain the possibility of misrecognising oneself and forming I-judgements that are affected by EM. As a matter of fact, the view according to which the self is a person would be compatible with the fact that *all* I-judgements based on recognition are open to error through misidentification. This would be so because it is always an open possibility to apply the first-person concept to the wrong person. Hence, a

proponent of the view according to which the self is embodied; who also shares the assumption that mental self-ascriptions are based on a recognition or identification of oneself, should be prepared to endorse the consequence that *mental* self-ascriptions such as [I am in pain], or [I am thinking] be always liable to error through misidentification. But nobody would be prepared to say that this kind of mental self-ascription could be affected by error through misidentification. Hence, such a theorist, *on the basis of the assumptions we have made so far*, is in danger of denying that there is a genuine distinction between self-ascriptions which are EM and those which are IEM, because all self-ascriptions would turn out to be open to EM. But this, in turn, will contrast with the intuitive thought that at least some mental self-ascriptions are IEM. If, however, the supporter of the view that selves are persons tried to save the intuitive thought that at least some mental self-ascriptions are IEM, then it is obvious that—given the assumptions we have been making all along, i.e. that it is the nature of the self and of the conditions of our encounter with it that are supposed to explain IEM of mental self-ascriptions, understood according to IEM_{diff}—she would not be able to explain how it is possible for at least non-inferential mental self-ascriptions to be IEM. Indeed, the mere appeal to the fact that selves are embodied is of no help whatsoever.

So the dialectical situation seems to be this: the Idealist and the Cartesian seem, given certain assumptions, to be able, at least *prima facie*, to offer an explanation of the fact that mental self-ascriptions are IEM. By contrast, given those very assumptions, the supporter of the view that selves are persons does not seem to be in a position to explain it. So, for the latter to be able to solve the

problem of IEM of mental self-ascriptions, she must offer an altogether different account of IEM.

At this stage it seems fair to say that we have a trilemma: either we opt for a *prima facie* explanation of IEM of mental self-ascriptions, but then we endorse either the Cartesian position about the self or the Idealist one, and we end up having either a bad metaphysics or a bad semantics; or else we opt for a sound metaphysics and a sound semantics, but we have no ready at hand explanation of IEM of mental self-ascriptions.

We can visualise the horns of our trilemma as follows:

	Self	IEM of mental self-ascriptions	Consequences
Cartesian	Salient object: a thinking substance which is present to the subject for each act of thought.	Given the nature of the self and the conditions of our encounter with it, we cannot mistakenly recognise it.	The self is a thinking substance. 'I', when used as subject, refers to it. BAD METAPHYSICS
Idealist	No object	Since there is nothing to recognise, our use of 'I' as subject cannot but be IEM.	There is no self. 'I', when used as subject, is not a referring expression. BAD SEMANTICS and possibly BAD METAPHYSICS as well
Primitivist	A person, i.e. a primitive entity with both physical and psychological properties, such as a human being.	?	Sound metaphysics and semantics, but: NO EXPLANATION OF IEM OF MENTAL SELF-ASCRIPTIONS

4. Epistemological vs. Semantic Accounts of IEM

All we have been saying so far presupposes that there is a genuine distinction—which is in need of explanation, of course—between uses of ‘I’ which are EM and those which are IEM. But it can seem open to one to *deny* that there is any such distinction. In fact, as we anticipated in chapter 1 (§§ 3-4) one could say that all uses of ‘I’ are, in some sense, IEM. That is to say, any use of the first-person either in speech or in thought is governed by the token reflexive rule (TRR) according to which any token of it refers to its producer. Hence, no matter how mistaken a subject is, she cannot be mistaken about the fact that in making a first-person judgement she is correctly identifying herself.

The thought seems to be this: in so far as, in the case of the first-person, reference to the right object is guaranteed by TRR, then there is no possibility of error through misidentification. In other words, if, *per impossibile*, error through misidentification could occur, then the subject would not refer to herself by using the first person but to the person who is actually responsible for it seeming to her that there is someone who is ϕ -ing. Since this cannot ever be the case, given that comprehending uses of ‘I’ are governed by the token reflexive rule, then the possibility of EM is ruled out as well. Hence, all cases of alleged misidentification relative to the first-person are in fact cases of *mispredication*. This idea is clearly expressed by Andrea Christofidou, who writes:

Even if my memories and perceptions deceive me, and it is not I who climbed the steps of St. Paul’s but a friend, then I misidentify, not myself, but the person who climbed the steps; my use of ‘I’ refers correctly to me, and refers to me directly and unmediatedly. In

discovering that I did not after all climb the steps of St. Paul's, then I should say 'so I did not climb them', and then I should try to explain how I could have fallen into error. But the error has nothing to do with the self-reference of 'I'. If there is a problem with propositions such as 'I climbed the steps of St. Paul's', the mistake is in the predication component, not in the identification component. I have *misascribed* to myself a predicate, but the reference or identity of 'I' is unshaken: 'I' is immune to error through misidentification, whatever the predicate might be.

(...) [D]iscussions in this area have been vitiated by the failure to distinguish between two types of immunity: immunity to error through misidentification (which covers *all* uses of 'I' whether these involve corporeal predicates, or mental predicates, or both) and what I call immunity to error through *misascription* (which covers only certain mental predicates).⁵

In short, according to this position, if EM occurs, this implies that no reference to the semantic referent is made. Since, however, in the case of the first-person, TRR makes it the case that all uses of the first-person refer to oneself, then this means that EM can never affect first-person judgements. Consequently, in the case of first-person judgements (a) IEM is always guaranteed, and (b) IEM is a consequence of the *semantics* of the first-person, i.e. IEM is a consequence of TRR.

However, this style of explanation amounts to conflating IEM with the impossibility of a split, in the case of competent uses of 'I', between semantic reference and speaker's reference, which is indeed guaranteed by TRR. In other words, the latter impossibility, which we could also gloss, in positive terms, as the guarantee, in the case of the first-person, that not only does it

⁵Christofidou 1995: 227, fn. 7.

always refer, but it always refers to the right person, i.e. oneself, is indeed a consequence of the semantics of the first-person and, in particular, of the *reflexivity* of the first-person pronoun. But this, as we saw in chapter one, does not mean that in a particular case our first-person judgement cannot be prompted by a mistaken recognition and, hence, be affected by EM.⁶

Moreover, we have already seen in the previous chapter that it is a characteristic of *all* judgements which are affected by EM that the singular term used to express them does not refer to the object perceptually given for recognition but to its semantic referent. Hence, reference to the semantic referent is always compatible with EM, no matter whether the judgement contains the first-person or any other singular concept. In fact, what goes wrong, both in the case of judgements containing the first-person and judgements containing other singular terms, when affected by EM, is not the fact that the subject refers, by using a certain expression, to something different from the semantic referent. Rather, when EM occurs, the speaker mistakenly takes the term in question (and its associated concept) to be instantiated by the object perceptually given to her, which, however, is not its semantic referent.

A similar mistake, with respect to IEM, can be found in Carol Rovane's papers on the topic. She writes:

⁶Pryor seems to agree with our position. He writes: "[T]he claims "I am bleeding", "I have a broken arm" and so on, will *never* suffer from badly aimed reference. My use of "I" in these claims will always refer to the person my basic attempt aimed to refer to. However, these claims *are* vulnerable to *de re* misidentification". Pryor 1998: 301, fn. 22.

[When EM occurs] [w]e are inclined to say that I must have intended to refer to myself when I used the expression 'me' and that I thought that *in* referring to myself I was also referring to the person reflected in the mirror. Hence it is more natural to interpret my utterance as false of myself rather than true about the person reflected in the mirror. *Nevertheless, the alternative interpretation might be warranted in certain circumstances, and, when it is, we have a use of the first person that incorporates error through misidentification* (italics mine). We must grant, then, that, although uses of 'I' can be subject to referential error through misidentification, it is extremely rare that they are. There is always the strongest presumption that a speaker intends to speak of *herself* in using 'I'. With other referring terms, it seems, there is more scope for error through misidentification and, hence, more scope for supposing that a speaker intends to speak of something that is not the standard semantically determined referent of the term she uses.⁷

Here, again, the assumption is that if EM occurs, then no reference is made to the semantic referent of the singular term used to give expression to one's judgement. Rather, reference would be made to the object presented to one. However, Rovane endorses a radical Fregean semantics. Hence, according to her, reference is always mediated by a set of beliefs about the referent. In the

⁷Rovane 1987: 153-154. Rovane firstly generates the following puzzle, i.e. how is it that although I may have false beliefs about me, I am still referring to myself by using 'I' and not to the person who is at the origin of my beliefs? Moreover, when do my beliefs turn into beliefs about someone else? Then, in order to solve it, she goes on to suggest that this is so because we have mostly true beliefs about ourselves. By contrast, in the case of other objects or people we are more likely to have false beliefs about them, thereby failing to refer to them, despite our intentions. A similar problem is raised by O'Brien (1995: 240), in the course of the discussion of Evans, and traced back to discussions with John Campbell (p. 247, fn. 15).

case of the first-person, she thinks that most of our beliefs about ourselves are *true*. Consequently, this guarantees that, by using 'I', we do refer to ourselves. Hence, this rules out the possibility of EM in most uses of the first-person. Here again, although IEM does not apply to all first-person judgements but only to some or most of them, it is, nevertheless, a consequence of the semantics of 'I'. Namely, it is a consequence of the fact that the reference of 'I' is fixed by a set of mostly true beliefs.

However, if we are right in claiming that all uses of 'I' which are affected by EM refer to the subject, then this means that there is nothing special in the *semantics* of the first-person which is the key to a proper explanation of IEM of I-judgements. From this we can also conclude that a proper explanation of IEM of I-judgements will not advert to aspects of the *semantics* of, in our case, the first-person. Nor, for all we have been saying in this chapter so far, will it advert to the particular *nature of the self* and *of the conditions of our encounter with it*. Rather, it will advert to the *epistemological* aspects of some first-person judgements. For, as we have begun to see in the previous chapter and will see further on, EM and IEM have to do with the kind of *grounds* on which our first-person judgements are based. Hence, our task is to vindicate the idea that a proper understanding of IEM can only be reached within the confines of an *epistemological investigation*. Let us now consider in more detail the various accounts that are revisionist of the original definition of IEM and how they combine with possible solutions to the trilemma.

5. The Forms of the Solution to the Trilemma

The characterisations of EM and IEM offered in the last chapter and at the beginning of the present one allow us to make a taxonomy of the solutions which can be given to the problem of explaining IEM of some self-ascriptions. This, in turn, will be useful in order to give a taxonomy of the possible solutions to the trilemma.

Recall the definition of IEM we gave:

IEM_{df1} A singular judgement of the form [a is F], formed on the basis of a recognition or an identification, is IEM relative to the subject, iff it is not possible for the thinker to apply the particular singular concept [a] to any other object but the relevant object a.

In the following we will call a *conservative solution* to the problem of explaining IEM of some first-person judgements a solution that does not depart from this definition. As a consequence of holding this definition, conservative solutions are forced to appeal to the nature of the self and to the conditions of our encounter with it in order to explain IEM of mental self-ascriptions. As we have seen, the Cartesian and the Idealist solutions are conservative ones, but, unfortunately, they give rise to a dilemma. To repeat: we have at least a *prima facie* explanation of IEM of mental self-ascriptions, but we subscribe either to a Cartesian dualist metaphysics of the self or to the no-reference view about 'I' and, possibly, to an equally embarrassing metaphysics of the self, i.e. the no-subject view. Both horns, as we have seen, are problematic. By contrast, to insist that 'I' does refer to a person (i.e. to a human being) is not by itself

sufficient to explain IEM of non-inferential mental self-ascriptions, understood according to IEM_{df1}. Hence, the proposal that allows us to save both a sound metaphysics and a sound semantics, fails precisely to explain IEM, when it is defined as above. For this reason, at the end of the previous section, we concluded that, in fact, we are presented with a trilemma: if we hold IEM_{df1} then we can explain IEM of mental self-ascriptions, but we give up the prospects of having either a sound metaphysics or a sound semantics. By contrast, if we hold that the self is identical with a human being and that 'I' refers to it, we have a sound metaphysics and a sound semantics, but we do not have an explanation of IEM—understood according to IEM_{df1}—of some mental self-ascriptions.

In order to make an explanation of IEM compatible with the fact that selves are persons and the fact that 'I' refers to an embodied self, accounts of IEM that are revisionist of the original definition of IEM have been offered as well. In the following, we will call them "*revisionist accounts*" *tout court*. The main characteristic of these accounts consists in the fact that they offer a *re-definition* of IEM which is meant to show how it is possible for IEM to arise without giving rise to the trilemma.

Amongst revisionist accounts of IEM we can distinguish further between those which aim at offering a *substantive response* to the trilemma, and those which just offer a *deflationist response* to it. According to the former, it is not sufficient to show that a proper understanding of IEM is compatible with the fact that 'I' is a referring expression and that it refers to an embodied entity. Rather, we should also conclude that the Cartesian and the Idealist cannot even get off

the ground, as it were. This is so because—it is argued—in order to make use of the first-person concept at all we must think of ourselves as embodied entities, contrary to both the Cartesian and the Idealist positions. This result, in turn, is taken to be a consequence of the fact that, according to the proponents of this view, in order to have the *first-person concept* at all, the subject must be at least dispositionally able to make bodily self-ascriptions which are IEM, where this is thought to be incompatible with the Cartesian and the Idealist positions (we will see why in discussing Evans).⁸ In short, according to substantive responses to the trilemma, *we cannot be either Cartesian or Idealist*, for these positions, once properly analysed, are conceptually incoherent or unintelligible from the stand-point of the supporter of the view that selves are embodied.⁹

By contrast, according to deflationist responses to the trilemma, given that the redefinition of IEM offers a way of explaining why some mental self-ascriptions are IEM, and extends the class of IEM judgements to some bodily self-ascriptions and to certain uses of demonstrative expressions, it undercuts the specific motivation behind the trilemma. That is to say, since IEM becomes *compatible* with reference to material objects, like in the case of some demonstrative judgements, and with self-ascriptions of bodily properties, the alleged explanation of IEM of some mental self-ascriptions can no longer be considered a motivation for either the view that 'I' does not refer or it refers to a

⁸Cf. Evans 1982. But see also McDowell 1997.

⁹For a discussion, cf. section II.

Cartesian ego.¹⁰ However, deflationist responses to the trilemma do not take issue with whether the Cartesian and the Idealist uses of the first-person concept are intelligible or coherent. Rather, they maintain the following position: IEM of mental self-ascriptions is *not* enough motivation for either the Cartesian or the Idealist conception of the self, since IEM is no less compatible with the latter positions than it is with the position according to which the self is embodied. Hence, if there is a real issue about the nature of the self and the semantics of 'I', it cannot be motivated by considerations concerning IEM, at least not IEM alone. Rather, the Cartesian and the Idealist positions, if motivated at all, must be motivated by different sorts of considerations, which have to be further investigated. In short, according to deflationist responses to the trilemma, *we need not be either Cartesian or Idealist*. For we can explain IEM and hold that the self is identical with a human being. In fact, the deflationist and the substantive accounts can be combined. For instance, according to Evans, the deflationist account is the first step towards the substantive solution to the trilemma. However, they need not be so combined and, in fact, most of the time, they are not.¹¹ Let us now introduce these positions in more detail.

¹⁰This seems to be Shoemaker's, Wright's, Pryor's and Peacocke's position. For a discussion, cf. section III.

¹¹Consider for instance Peacocke 1999, Wright 1998, Shoemaker 1994a. For a discussion, cf. section III.

5.1 Revisionist Accounts of IEM and Deflationist and Substantive Responses to the Trilemma¹²

Three moves are common to deflationist and substantive responses to the trilemma despite their differences: first, a redefinition of IEM, secondly, the denial that IEM of self-ascriptions depends on their subject-matter and, finally, the observation that certain uses of demonstratives are also IEM. We will start by taking up this latter point.

Noticing that certain uses of demonstratives are IEM is meant to show that IEM has nothing to do with the *referential role* of the singular term featuring in the sentence used to express the judgement. Hence, this point is often made against the Idealist position, whose possible development consists in the no-reference view about 'I'. For—it is argued—although some uses of demonstrative expressions are IEM, nobody would deny that they do refer.¹³ Let us consider some of the examples used to make the point that certain uses of demonstratives are IEM. Consider, for instance, the case in which the subject sees a rug in front of her and judges [That rug is a Persian Nain], or she sees something approaching very fast and judges [That thing is approaching very

¹²A terminological remark on the use of 'response' which is made throughout this work. None of the authors we will consider in this work directly addresses the trilemma, for the obvious reason that it is our own claim that there is such a trilemma. However, from what they say we have tried to formulate various possible responses to it.

¹³Cf. Shoemaker 1994a, Wright 1998, Pryor 1998, Peacocke 1999, discussed in section III.

fast]. In both cases we have a perceptual-demonstrative judgement where a property is attributed to an instance of an observational concept such as [That thing] or [That rug]. There being a genuine perceptual-demonstrative judgement at all guarantees that the demonstrative concepts used to make it have a referential function.¹⁴ Moreover, since the judgement does not involve as part of the grounds on which it is based any belief in an identification component of the form [That F = the such-and-such], then “there is no room for error resulting from the falsity of such a belief”.¹⁵ In other words, these perceptual demonstrative-judgements contain only an *individuation component*. That is to say, they only involve the exercise of a single, demonstrative concept, on the basis of a perceptual discrimination of a particular. Hence, *they fall underneath the threshold of the distinction between EM/IEM_{dfr}*; for, if there is no identification component, given that no recognition is going on at all, then, *a fortiori*, there is no possibility of making a mistake of identification either. Hence, if error arises, it can only affect the predication component.

However, those who aim at offering a conservative account of IEM would think that this is a conflation, because if the conditions for the possibility of misidentification are not met, because no identification is going on at all, then there is no legitimate sense in which we could speak of immunity to error through *misidentification* either. After all—the thought would be—to notice that sounds are colourless, thus guaranteeing that they cannot be of two different

¹⁴We are assuming that perception is veridical and, hence, that [That thing] and [That rug] refer.

¹⁵Peacocke 1999: 287.

colours all over at the same time, would not help in explaining why a surface, which is necessarily coloured, cannot be of two different colours all over at the same time. Similarly, to notice that demonstratives can be used in such a way that the question of their being affected by EM does not even arise, because there is no identification component, cannot be used to explain why certain uses of 'I', which, at least *prima facie*, involve an identification component, are IEM.

By contrast, those who aim at offering a revisionist account of IEM argue that to notice that certain uses of demonstratives are IEM is important and indeed crucial to an understanding of IEM in general. For it helps to bring out the fact that any judgement is *IEM just in case it falls underneath the threshold of the distinction EM/IEM_{df1}*. That is to say, any judgement that is IEM is such not because an infallible recognition or identification is taking place, but because *no* recognition or identification is taking place at all. Consequently, the judgement would not involve as part of the grounds on which it is based any belief in an identification component. This would suggest a revision of the first definition of IEM along the following lines:

IEM_{df2} A singular judgement of the form [a is F] is IEM relative to the subject, iff it is not formed on the basis of a recognition or identification of the object thought about and, consequently, iff it does not involve an identification component.¹⁶

¹⁶Not all parties engaged in this debate give an explicit definition of IEM along these lines (Cf. Wright 1998), but most do (Cf. Evans 1982, Peacocke 1999).

An argument in favour of such a definition of IEM is the following one: if any judgement about a particular involves recognition (or identification) and recognition (or identification), in turn, involves fallibility, then all judgements can be affected by EM. Hence, there would be no IEM judgement, contrary to our initial intuitions. By contrast, if we wish to save our initial intuition that at least some I-judgements are IEM, then, by holding IEM_{df1}, IEM would only be guaranteed by maintaining contentious metaphysical views. Hence, in order to avoid endorsing bad metaphysical views, while preserving our initial intuition that there are IEM uses of the first person, IEM_{df2} ought to be preferred over IEM_{df1}.

At this stage we need not choose between the two accounts of IEM we have introduced in this chapter.¹⁷ What we should point out, however, is that it is not enough to give a correct definition of IEM in order to solve our trilemma. Rather what is more important, and indeed, more difficult, is to provide *arguments in favour of such a definition and explain how some first-person judgements can comply with it.*

Another point which is usually made by appealing to the fact that demonstratives can also be used to express judgements which are IEM is that in these cases not only is reference made, but it is made to *spatio-temporal objects*. Hence, IEM of some self-ascriptions is *compatible* with the fact that 'I' not only refers, but it refers to an *embodied entity*.

¹⁷We will come back to this issue in Ch. 8, where we will also revise this preliminary revisionist definition of IEM and will offer further arguments in its favour.

While deflationists can be content with this result, because they just aim to show that IEM is *compatible* with the position according to which 'I' refers to a person, substantive theorists want to show that some *self-ascriptions* of *bodily* properties are *also* IEM. Once this point is coupled with the thought that IEM *bodily* self-ascriptions must be at least dispositionally in place for one to have the first-person concept at all, then the substantive position can be fully fleshed out. For these two claims together imply that subjects must think of themselves as embodied in order to have the first-person concept. Consequently, the Cartesian and the Idealist by departing from such an use of the first-person are maintaining only *prima facie* intelligible positions about the self. But, in fact, they are surreptitiously altering the concept of the first-person, to such an extent that their views become unintelligible from our stand point.

Thus, in order to show that our first-person concept is of embodied entities, the second argument common to most deflationist and substantive accounts is introduced. This amounts to the *denial of the domain-specificity* of self-ascriptions that are IEM. That is to say, contrary to what we saw in the previous chapter, it is denied that *all and only mental self-ascriptions* are IEM. The denial consists of two steps: first, it is commonly noticed that there are some mental self-ascriptions which are based on inference involving an identification component and which are, therefore, subject to EM. Secondly, it is shown that also some bodily self-ascriptions are IEM_{df2}. In particular self-ascriptions of bodily properties based on proprioceptive information and on perceptual information derived from the environment are taken to be IEM_{df2}. We will consider these arguments in detail in section II.

Still, any response to the trilemma cannot hope to be fully effective unless it engages in a positive explanation of *why* non-inferential mental self-ascriptions are IEM_{df2}. That is to say, in order to provide a fully-fledged way out of the trilemma, not only must substantive and deflationist accounts endorse a definition of IEM which is revisionist of IEM_{df1}, but they must also explain why some mental self-ascriptions (as well as some bodily ones) can comply with it. To my knowledge, however, an exhaustive account of why some mental self-ascriptions are IEM_{df2}, within the framework of either a deflationist solution to the trilemma or of a substantive response to it has not yet been given. We can represent the situation as follows:

Revisionist Accounts
IEM_{df2}

Deflationist Responses

1. Some demonstrative judgements are IEM
2. Some bodily self-ascriptions are IEM
3. Some mental self-ascriptions are IEM
4. IEM is compatible with the fact that the self is embodied and with the fact that 'I' refers to such an entity
5. **We no longer need to be either Cartesian or Idealist in order to explain IEM**

Substantive Responses

1. Some demonstrative judgements are IEM
2. Some bodily and some mental self-ascriptions are IEM
3. The disposition to make bodily self-ascriptions which are IEM is necessary in order to have the first-person concept
4. Our first-person concept is incompatible with the conception of the self endorsed either by the Cartesian or the Idealist
5. **We can no longer be either Cartesian or Idealist**

Both Deflationist and Substantive Responses must explain why some mental self-ascriptions are IEM_{df2}

So, in the following we will analyse Evans' substantive response to the trilemma. Then we will move on to various deflationist solutions. Finally, we will provide both a sound revisionist definition of IEM and an explanation of why non-inferential mental self-ascriptions can comply with it.

SECTION II

**A REVISIONIST ACCOUNT OF IMMUNITY TO ERROR
THROUGH MISIDENTIFICATION AND A SUBSTANTIVE
RESPONSE TO THE TRILEMMA: EVANS**

Summary

In this section we consider the most thorough *substantive response to the trilemma* which has been given so far in the literature, namely Evans'. Such a response is presented in the context of a theory of *singular thoughts* (i.e. perceptual demonstrative thoughts, here-thoughts and I-thoughts). According to such a theory there are some kinds of thought which are not reducible to descriptive thoughts and which are *Russellian*, i.e. *object-dependent*. That is to say, these thoughts would not exist if the objects they are about did not. In order to present and evaluate Evans' response to the trilemma the exposition must proceed with some care, however.

To such an end, in Ch. 3

- we present some fundamental tenets of Evans' theoretical framework (§ 1), namely
 - Russell's Principle (RP) (§ 1.1) and
 - the Generality Constraint (GC) (§ 1.2);
- we discuss demonstrative identification (§ 2) ;
 - paying particular attention to the role information-links play within it (§ 2.1)
 - and showing how such links are the route to the claim that perceptual demonstrative thoughts are Russellian (§ 2.2);

- we then turn to Evans' account of EM and to his revisionist account of IEM (§ 3). According to Evans, IEM is a consequence of the fact that the judgement is *identification-free*. In turn, this is a consequence of the fact that IEM judgements are solely based on a way of gaining information directly from objects such as perceiving them. We propose some corrections to Evans' account, in particular we argue that an account of IEM should be, in the first place, an account of IEM *judgements*, rather than an account of IEM *knowledge*.

In Ch. 4 and 5 we turn to Evans' account of I-thoughts.

In Ch. 4

- we connect Evans' treatment of I-thoughts with his treatment of perceptual demonstrative ones (§ 1);
- we introduce further characteristics of I-thoughts by considering Evans' sources, namely, Frege (§ 2) and
 - Perry (§ 3);
 - we present another main tenet of Evans' account, namely the Objectivity Requirement (OR) (§ 5). According to such a tenet, for the first-person concept to be adequate it must be a concept of a creature who thinks of herself as embodied, as located in space, and as on a par with all other objects which are thought of as existing independently of being perceived.

- we turn to Evans' defence of the claim that our first person concept meets OR, and to his attack against the Cartesian and the Idealist. To such an end,
 - we present and defend Evans' claim that also bodily self-ascriptions are IEM when based on appropriate grounds (§ 1). Once this result is considered within Evans' theoretical framework, according to which in order for a subject to have the first person concept she must have discriminative knowledge of herself, it shows that the subject, by having discriminative knowledge of herself as embodied and located in space, has a concept of herself as an embodied self. We discuss:
 - bodily self-ascriptions based on somatic proprioception (§ 1.1) and
 - physical self-ascriptions based on outer perception of one's environment (§ 1.2). In both cases we argue that Evans' arguments for saying that the relevant self-ascriptions are IEM, because the nature of the information is such that they cannot be rationally based, at least *de facto*, on inferences involving identification components, are conclusive;
 - we consider Evans' attack to the assumption which seems to be presupposed by the Cartesian and the Idealist conceptions of the self that *all* mental self-ascriptions are IEM (§ 2). We agree with him that only *non-inferential* mental self-ascriptions are IEM. But we argue that he has not provided us with an explanation of why introspection-based mental self-ascriptions are *logically* IEM, i.e. they are such that they could *never* involve.

as part of the subject's rational grounds for them, a belief in an identification component;

- we then turn to a different pattern of argument used by Evans. This is meant to show that the ability to self-ascribe mental properties presupposes the ability to *think* of oneself as a human being. This pattern of argument makes use of GC. In (§ 3) we show how this kind of argument can work in the case of self-ascriptions of beliefs and

- in (§ 4) we show how this pattern of argument can work in the case of self-ascriptions of inner and outer perceptions. But, we argue that Evans' argument is not enough to show that all cases of mental self-ascriptions are based on *awareness* of oneself as an embodied entity;

- in § 5 we summarise our conclusions.

Chapter Three

Evans: Perceptual Demonstrative Thoughts and the Revisionist Account of IEM

In this chapter we are going to present Evans' theory of perceptual demonstrative thoughts and his account of immunity to error through misidentification. In fact, the chapter will be quite expository. This is for four reasons: first, in order to understand Evans' conception of EM/IEM we need some background about his overall approach to singular thoughts. This was first approached and brought out with vigour in his treatment of perceptual demonstrative thoughts. Secondly, a presentation of Evans' views on what he calls "demonstrative identification" will allow us to understand better the phenomenon which, in previous chapters, we have called "demonstrative-individuation". For we will understand better the conditions under which we can say that a subject has demonstratively individuated (or identified, in Evans' terminology) a particular perceptually given to her.¹ Thirdly, this will allow us to

¹We hope that our terminology is clearer than Evans'. For instance, as we will see, he defines immunity to error through misidentification in terms of identification-freedom. But this does not mean that there is no demonstrative identification of a particular in Evans' original sense of identification. On the contrary, it is a consequence of what Evans calls the demonstrative identification of a particular that the corresponding

see more clearly why, according to Evans, some uses of demonstratives are IEM. Fourthly, Evans' discussion of I-thoughts and of their being IEM is closely connected to his discussion of perceptual demonstrative thoughts (and here-thoughts as well). Consequently, an understanding of Evans' claims with regard to the first person will certainly benefit from an introductory discussion of his approach to perceptual demonstrative thoughts.

In fact, it should be kept in mind that Evans' theory of singular thoughts comes in a package. He was interested in giving an account of the conditions of the possibility of singular thoughts: perceptual demonstrative thoughts, here-thoughts and I-thoughts. According to him, despite some differences, which we will see in the following, all these forms of thought are essentially Russellian, which is *object-dependent*. Moreover, they are inter-related, since, according to Evans, what enables a subject to have perceptual demonstrative thoughts is the capacity to *locate* the objects of thought in *egocentric space*, i.e. the space which has its origin in a system of axes originating from the subject. This is possible only because of the subject's capacity to have here-thoughts. Since, however, here-thoughts are thoughts about positions in egocentric space, they can be entertained only by a subject who has a conception of her position with respect to other objects and has a conception of herself as occupying a region

judgement will be identification-free. As will become clear, "identification-freedom" means absence of an identification component, in our sense (cf. Ch. 1). Using our terminology, however, things are clearer. "*Identification-freedom*" is just what Evans means (but see in the following for a precise characterisation) and it is a consequence of the demonstrative *individuation* of a particular.

of space within which she can act. Thus, here-thoughts are dependent on I-thoughts. Notice, however, that Evans never tried to *reduce* any one of these forms of thought to any of the other ones. Rather, he just pointed out that they are inter-connected.

Hence, the structure of the chapter will be as follows: we will introduce some basic elements of Evans' theory of singular thoughts. Afterwards, we will turn to his account of perceptual demonstrative thoughts, paying particular attention to the role of informational-links. Finally, we will present and discuss Evans' account of immunity to error through misidentification in connection with perceptual demonstrative thoughts and we will propose some corrections to it which will be relevant in the following chapters, both for assessing Evans' own solution to the trilemma and for reaching a good understanding of immunity to error through misidentification.

1. *Russell's Principle and The Generality Constraint*

According to Evans, singular terms can be referring devices only if the appropriate kinds of thought are in place. In Evans' terminology the thoughts that allow one to make reference to an object constitute one's *Idea* of the object.² Moreover, for an Idea of an object to be *adequate*, not only must it be

²An Idea of an object is different from a Fregean sense in that it does not exist objectively, i.e. "independently of anyone's grasp of it". However, two people, by exercising two numerically distinct Ideas of an object, "may thereby 'grasp' the same Fregean sense. What this means is that they may think of the object in the same way.

based on a perceptual link with the object, but it must also give the subject *discriminative knowledge* of the object it is about and meet a *constraint of generality* so that it can enter into indefinitely many thoughts about that object. The former constraint constitutes Evans' interpretation of *Russell's Principle*, whereas the latter constraint is a consequence of his adherence to the *Generality Constraint*. We will consider them in turn.

1.1 *Russell's Principle*

Evans borrowed Russell's Principle (RP hereafter) from Russell and he phrased it as follows:

RP: a subject cannot make a judgement about something unless he knows which object his judgement is about.³

This principle has not always had good press, particularly because of the difficulty of determining what the "knowing which" requirement would amount to. In fact, it was a consequence of Russell's own interpretation of it that only sense-data could be known in the way specified by this principle. We shall not engage in a discussion of the interpretations of RP.⁴ However, it states a sound general condition upon thought about objects. For if a subject did not have *any* knowledge whatsoever of the object her thought is allegedly about, then it would

(And the way of thinking would be available even if no one ever thought of the object in that way.)". (Evans 1982: 104, fn. 24).

³Evans 1982: 89. Cf. also Russell 1912: 58.

⁴For a discussion, see Evans 1982, McDowell 1990, Millikan 1993, Sacchi 1999.

become altogether dubious that she could really think about it. Obviously, then, the problem, in order to make RP a viable principle, is to determine what sort of knowledge must be involved in *each* kind of thought about objects, or particulars, we may have. Hence, different forms of knowledge must be at work when we think about perceptually given objects, bearers of proper names we are not acquainted with, abstract objects, ourselves, places and times. In other words, RP is a principle which governs all our thoughts about objects. The substantive explanatory task, then, is to specify what it amounts to for each specific kind of singular thought we want to characterise.

Evans argued, on the one hand, that if, in the case of perceptual demonstrative thought, we understand “knowing which” as implying descriptive information about an object perceptually given to one, then a subject may have a perceptual demonstrative thought about an object in a very limited number of cases. By contrast, if we understand “knowing which” as involving no restriction whatsoever, then, by the very fact that one is causally interacting with the object, one would be credited with a thought about that object perceptually given to one. In order to make RP a “substantial principle”, in the case of perceptual demonstrative thought, Evans proposes to read “knowing which” as meaning that one must have “*discriminating knowledge*”. That is to say, “The subject must have a capacity to distinguish the object of his judgement from all other things”.⁵ We will come back to this in § 2.

⁵Evans 1982:89.

1.2 *The Generality Constraint*

Let us now turn to the Generality Constraint. Evans' version of it is indeed a developed version of Strawson's suggestion that, in order to block solipsism, we should realise that a subject can be granted with an understanding of the mental predicate used for a mental self-ascription just in case she is also able to apply that predicate to other subjects.⁶ Obviously, it is not our concern to determine whether Strawson's move against solipsism was successful. However, the gist of it is that we should see our mastery of a certain predicate *F* as the result of the ability to ascribe it to any subject *a, b, c...*, *n* of which the predicate may be significantly, "though not necessarily truly, affirmed".⁷ Evans' insight consists in generalising this principle and in bringing it to bear at the level of thought. Here is his formulation of it:

GC: If a subject can be credited with the thought that *a* is *F*, then he must have the conceptual resources for entertaining the thought that *a* is *G*, for every property of being *G* of which he has a conception.⁸

As a consequence, we should realise, according to Evans, that:

the thought that *a* is *F* [lies] at the intersection of two series of thoughts: the thoughts that *a* is *F*, that *a* is *G*, that *a* is *H*, ..., on the one hand, and the thoughts that *a* is *F*, that *b* is *F*, that *c* is *F*, ..., on the other (...).⁹

⁶Cf. Strawson 1959: 99.

⁷*Ibid.*

⁸Evans 1982: 104.

⁹Evans 1982: 209. Cf. Evans 1982: 104, fn. 21.

The main motivation behind GC is that of resisting the conception of thoughts as isolated occurrences in the subject's head, seeing them, instead, as possible only on the basis of various abilities, which can be exercised independently of one another on different occasions. It should also be stressed that, according to Evans, GC does not merely consist in a play of substitutions of subjects and predicates. Rather, it also consists in the ability to embed a thought within more complex structures. In fact, subjects should be able to reflect on their previous thoughts and recall them, thus using the past mode of thought. Similarly, they should be able to make them objects of suppositions about the future. Finally, they should also be able to embed them within various propositional attitudes and use them in prudential and moral reasoning.¹⁰

It is important to stress that GC should be subject to certain limitations. For, as Strawson noticed, in order to be credited with the mastery of a given predicate, one must be able to attribute it to any subject of which that predicate could be *significantly*, though not necessarily truly, predicated. Similarly, in the case of GC, we should impose the constraint that the various properties which a subject must be able to ascribe to an object, in order to be credited with a thought about that object, should be *relevant* properties, i.e. those which can be significantly attributed to the object in question. However, once we impose this constraint, GC seems quite plausible.

Hence, we see that according to Evans, one's Idea of an object is adequate, if we have discriminative knowledge of the object and can exploit such knowledge in indefinitely many thoughts about that object. Let us now turn to Evans' way of

¹⁰Cf. Evans 1982: 105, fn. 27.

giving substance to the idea that in order to think of an object we must have discriminative knowledge of it, by considering his discussion of demonstrative identification.

2. Demonstrative Identification

Having presented Evans' reading of RP and GC we are now in a position to appreciate better his account of demonstrative identification (or individuation, in our terminology), which is fundamental to an understanding of Evans' conception of IEM. It has to be stressed that Evans was interested in specifying the conditions in which a subject can be credited with a *perceptual* demonstrative thought about an object, i.e. with a thought about an object which "crucially depends upon the subject's currently *perceiving* that object".¹¹ First, we will dwell on his account of perception and the role it plays in grounding the relevant demonstrative judgements. Secondly, we will consider Evans' claim that perceptual demonstrative judgements are Russellian, i.e. object-dependent.

2.1 Towards A Clarification Of Our Concept of Perception: The Role Of Information-Links

Evans endorsed an account of the content of perception in terms of *non-conceptual content*. According to such a view, by being causally linked to an

object, subjects thereby acquire non-conceptual information about it. That is to say, they enjoy a mental state that has representational content, i.e. it represents the world as being thus-and-so, and has correctness conditions, i.e. it can be either veridical or non-veridical. However, this mental state does not depend, for its very existence, on the subjects' possession of the concepts relevant to its canonical specification.¹² Such non-conceptual information, in turn, is available at a personal level and can be relevant to action, thinking and reason-giving procedures.¹³ What, however, has still to be clarified is the

¹¹Evans 1982: 72.

¹²This is not the place to engage in a discussion of non-conceptual content. For an elaboration, cf. Peacocke 1992. For the opposite view, according to which experiences can have representational content just in case the subject has the conceptual resources relevant to its (canonical) specification, cf. McDowell 1994. For a liberalisation of non-conceptual content to experiences of subjects who do not have *any* conceptual resources, cf. Bermúdez 1994.

¹³I am representing Evans as having a straightforward conception of non-conceptual content. However, his pioneering remarks in this area are less sharp than I have represented them as being. Since it is not my main task to offer an exegesis of Evans, I am running the risk of over-reading his remarks. Yet, I do not think I am doing great violence to what he actually says. For what he says in Ch. 5 of *The Varieties of Reference* is that through our sensory system (but not only that) we receive information of objects. Hence, we get into a position in which we have a mental state with a certain representational content, which, in turn, is specifiable by means of an open sentence. Consequently, such a content would not have correctness conditions (cf. fn. 10, p. 125). However, Evans also claims the opposite, i.e. that the content can have correctness conditions (cf. pp. 226-227). Moreover, it is uncertain whether it constitutes

precise role of the information-link in the process of acquiring knowledge about the object which can eventually issue in demonstrative thoughts about it.

According to Evans, the mere information-link with an object is not enough to give us discriminative knowledge of that object. For, according to Evans, we must also be able to *locate* the object. However, in certain cases the ability to locate the object can consist in the practical ability to locate it relative to oneself and other objects by exploiting the spatial information afforded by perception. This, in turn, may suggest the idea that the information-link between the subject and the object is not only necessary but also sufficient in order for one to have a demonstrative Idea of an object. We shall presently see that this is not the case. Yet, let us clarify what has been said so far by means of the following example:

the content of a conscious experience either, although it can have some causal role in the determination of action (cf. p. 156, sq. For an opposite claim, cf. pp. 226-227). However, at least when these mental states are possessed by creatures endowed with *some* conceptual powers, then they become part of their conscious experience and can enter their acting, thinking and reasoning (cf. *ibid.*). Be that as it may, it is important to note that the conceptual powers which are needed in order for conscious experiences to have a certain representational content are not the ones which would be used in a canonical specification of that very content (cf. p. 159). Finally, it seems that concepts are brought to bear on the content of experiences through judgement (in fact, judgements are "reliably caused" by these mental states, cf. p. 227). Hence, it seems that Evans' position and Peacocke's (cf. Peacocke 1992) are quite similar. For, although it is disputable whether both of them ascribe non-conceptual representational content only to experiences of creatures who have *some* conceptual powers, they both

a subject is perceiving a glass in front of her. According to Evans, in this case the spatial information afforded by visual perception is enough (to first approximation)¹⁴ to enable the subject actually to *locate* the object in egocentric space, i.e. in the portion of space which has its origin in a system of axes originating from the subject. Hence, perception affords us discriminative knowledge of the object in question, which can issue in a demonstrative thought, such as [This glass is empty].

Moreover, according to Evans, in such a case, perceptual information is *unmediatedly* relevant to the determination of the truth value of a sentence containing a demonstrative expression, i.e. if one gets to know anything on that basis, then one does not know it through inference involving an identification component.¹⁵ That is to say, by being perceptually aware of an object (supposing the conditions of perception are favourable and the perceptual system of the subject is working reliably), the subject is immediately in a position to determine whether the object has a certain perceivable property. Thus, an information-link is a *necessary* condition for perceptual demonstrative

hold that the concepts which are needed to specify that content are not the ones which would figure in a canonical specification of the content of the experience.

¹⁴Provided that the subject has an adequate conception of space. Cf. below.

¹⁵Evans 1982: 146: "[A] subject who has a demonstrative Idea of an object has an *unmediated* disposition to treat information from that object as germane to the truth and falsity of thoughts involving that Idea. (When I say that his disposition is unmediated, I mean that it is not the product of any more general disposition to treat as germane to the truth and falsity of those thoughts information received from an object satisfying some condition, together with a recognition that the object satisfies that condition)".

identification. The further question is whether it is also a *sufficient* condition for it. That is to say, we have to ask whether the mere causal interaction with the object is enough to enable the subject to have an adequate Idea of the object.¹⁶

Unsurprisingly, given Evans' adherence to RP and GC, the information-link is *not* sufficient for having an adequate Idea of the object, i.e. knowing which object one's thought is about and being able to use such knowledge in indefinitely many thoughts about that object. In fact, according to Evans, there are two possibilities: either the information-link enables one to have knowledge of the spatio-temporal location of the object, or one knows what it would be for a *judgement* of the form [This = δt] to be true. That is to say, either the information-link suffices for enabling the subject to locate the object or it does not. If it does not, however, one can still form an adequate Idea of the object through perception if one knows what it would be for a judgement like [This = δt] to be true, where (δt) is a *fundamental* Idea of an object, i.e. descriptive knowledge of its location and of the fact that it is responsible for what the subject is perceiving.¹⁷

Let us illustrate the second possibility by means of an example. Consider the case of perceiving a man on television. Obviously, if the subject could rely *only* on the information-link in order to locate the object, she would say that the object is in/on the screen, which would be an inadequate thing to say. Thus the current perception gives the subject some knowledge only relative to the

¹⁶Cf. Evans 1982: 145-151.

¹⁷Cf. Evans 1982: 149.

immediate referent of 'This', namely a little figure on the screen. So, according to Evans, in order for the subject to form an adequate Idea of that object, she must be able to know what it would be for a judgement like [This = the man who is responsible for the sounds and images I am now perceiving] to be true. By knowing what that would be like, the subject can attain demonstrative-identifications of objects far away from her, and form adequate Ideas of them.

All this further work, however, is not done by being informationally linked to the object of perception. Rather, some further conceptual element must be at play. In particular, some conception of objects and their being spatially located in a public world that extends beyond the subject's perceptual reach must be at work.¹⁸ Thus, according to Evans, information-links between subjects and objects are not enough or sufficient in order to form an adequate Idea of the latter.

Notice, however, that, according to Evans, in this latter case, one should not conclude that, since sometimes perceptual demonstrative thoughts need to be connected to thoughts containing further conceptual elements (as in the case of [This = the man who is responsible for the sounds and images I am perceiving now]), perceptual demonstrative thoughts *as such* can be reduced to these descriptive thoughts (e.g. This glass = the glass I am now perceiving here). For, according to Evans, the former case, which could lead one to think that such a reduction could be possible, is in fact not a case of reduction at all. The idea is that even when subjects use some descriptive knowledge about the object they

¹⁸Evans 1982: 149: "It is when an information-link *does not provide the subject with an ability to locate the object* that a conceptual element is needed for identification".

are perceiving, they are in fact setting up a *link* between a demonstrative identification which gives them demonstrative knowledge of the direct object of perception and some more general conception of that object. But still, there is a purely demonstrative element related to the first term of the identity, namely 'This', which is non-eliminable, according to Evans.¹⁹ That there is a demonstrative, ineliminable element even in *mixed thoughts* like the ones we are considering can be brought out by reflecting on the fact that the subject is immediately sensitive to the possible variations the direct object of her perception can undergo, as in the case where it suddenly disappears, and she somehow reacts to its deeds. For instance, the subject would switch off the television, had the man disappeared from the screen, but she would not say that he is hiding somewhere, nor would she start an inquiry about what he really did in the place where the scene she was watching was shot.

Now we have to make two further complications of this model: first, we have to understand whether in the "easy" case of pure demonstrative identification it is really true that the information-links *alone* can enable the subject to *locate* the objects she is perceiving. Secondly, we have to understand whether *sortal* concepts are needed in order to have demonstrative identifications of objects and thus perceptual demonstrative thoughts about them.

In order to answer the first question, Evans embarks on a discussion of here-thoughts. For our purposes it is enough to recapitulate the gist of his discussion, which comes down to the following: information-links are *necessary but not sufficient conditions* for actually being able to locate an object in one's

¹⁹Cf. Evans 1982: 150-151.

egocentric space, i.e. the portion of space which can be captured by a system of axes originating in the subject, which is relevant to her perceptions and actions.

The idea is that, in order to have an adequate Idea of the space around one, i.e. an adequate Idea of the here in question, the subject must be able to conceive of her egocentric space as a part of public space. Hence, in order to have an adequate Idea of the location of an object within egocentric co-ordinates, i.e. in terms of 'here', 'there', 'to my left', 'to my right', etc., the subject must also be able to understand what it would be like to locate that object within public or objective co-ordinates, such as between X and Y, to the left of Z and the right of G, etc. To illustrate: suppose that a subject is in the Quad in St. Andrews, in the middle of the lawn. Suppose also that she is looking attentively at one of the benches. Further, suppose that she is able to locate it in egocentric space: in front of her, on her left hand-side, etc. But, according to Evans, for this location to be adequate, i.e. to exhibit a grasp of the category of place, the subject should also be able to locate the bench within a system of public co-ordinates, i.e. as between the Lower College Hall and School 1, etc. This, according to Evans, does not mean that she must *actually* be able to locate the bench within such public co-ordinates, but just that she should be so *disposed*, if she is to be credited with an adequate Idea of the location of the bench.²⁰ On this reading RP would amount to the following:

(RP1) In order for a subject to have a singular thought about an object perceptually given to her she must have discriminative

²⁰Cf. Evans 1982: 152.

knowledge of it. That is to say, she must *actually* be able to locate it in a system of egocentric co-ordinates and *dispositionally* able to locate it in a system of public co-ordinates.

Hence, although having an adequate Idea of an object is originally understood by Evans disjunctively (i.e. either one is practically able to locate an object in egocentric space, or one knows what it would be like for an object to be located in public space), it has to be born in mind that even in the case of the demonstrative individuation of an object made on the basis of perception, one's Idea of the object would be adequate only if the subject had the disposition to think of that object as located in public space.²¹

Finally, according to Evans, if RP1 is satisfied then we could have a perfectly adequate Idea of the object in question even if we had *no* idea of what *sort* of object that is.²² Let us exemplify: a subject could be on the beach in St. Andrews and find a mysterious object of a roundish shape, quite small, sort of brownish and wonder whether this is a stone or an egg. According to Evans, then, she would have a perfectly adequate Idea of that object, relative to [This] in the question [What is this?], for instance, even if she did not know what sort

²¹ Evans 1982: 173: "One has an adequate [demonstrative] Idea [of an object] in virtue of the existence of an information-link between oneself and the object, which enables one to locate that object in egocentric space. (That the Idea is adequate depends on one's ability to relate egocentric space to public space)".

²²Cf. Evans 1982: 178-179.

of object that is, provided that there is something like *discovering or finding out* what that object is.²³

2.2 Consequences: Russellian Thoughts

We have noted above that, according to Evans, perceptual demonstrative thoughts are irreducible to descriptive thoughts, perhaps containing “more fundamental” indexical-thoughts, such as here-thoughts (or I-thoughts). This, in fact, is the key to an understanding of Evans’ further claim that perceptual demonstrative-thoughts are *Russellian*, that is, *object-dependent*: were there no object there would be no thought either. For, according to Evans, if perceptual demonstrative thoughts are irreducible to descriptive thoughts and it is part of the conditions of possibility of these thoughts that there be an information-link between the subject and the object, then an object is needed in order to set up this link. Were there no object, there would be no information-link either, nor, consequently, any *perceptual demonstrative* thought at all. Hence, one’s thoughts would *seem* to have empirical content, but that would be just an illusion. Let us consider this claim in some more detail.

According to Evans, if there is no object and a subject is having an hallucination, then, if she forms the judgement [That a is F] she is not in fact

²³As we shall see in Ch. 8, the fact that actual knowledge of the sort of object one is presented with is not required for perceptual demonstrative thoughts may help to avoid the mistake of thinking that demonstrative thoughts always involve identification and, therefore, cannot be IEM.

having a singular thought at all because the necessary condition for singular thoughts, namely being informationally linked to an object, is not satisfied. Hence, according to Evans, the subject would suffer not only from a perceptual illusion but also from a *cognitive* illusion: she thinks she is thinking a singular thought, when in fact she is not thinking such a thought at all.

Consider the following situation: a subject is thirsty and sees a glass of water in front of her. That would lead to the thought [That is a glass of water] which can enter the following reasoning procedure: [I am thirsty], [That is a glass of water], [In order to drink it I should reach out for it], which can move the subject to act accordingly. Now consider the case in which she is hallucinating a glass of water in front of her and she is actually thirsty. From the *first-person perspective* it is plausible to say that all the reasoning and acting we envisaged in the former case will repeat identically in the latter case.

Now a broadly Fregean theorist can say that the subject was entertaining the same singular thought and that is why having a reliable perception or an hallucination makes no difference whatsoever from the first-person point of view. But Evans, as we have seen, denies that, in the case of hallucinations, there is any singular thought at all. Now, we can take Evans to be saying two different things: either there is nothing in the subject's head, so to speak, no thought whatsoever, as it were, or else there is some thought, but not the one the subject takes herself to be having. The former possibility is not very plausible, for we could not make sense of the subject's behaviour. That is to say, we could not even regard it as a kind of action, given that the belief which should issue in the formation of an intention that would turn merely bodily

behaviour into a piece of action would be lacking. The second alternative, however, is more promising. But what is there in her head? According to Evans, we should say that the subject entertains a thought which, phenomenologically, looks like a singular thought which is in good standing, but which, however, does not really have content (it has just an illusion of content). However, from a third-person perspective, we would say that the subject thinks that she thinks a genuine singular thought, although she is not thinking such a thought at all. This would reconcile our attribution with her subsequent behaviour, and also with her phenomenology, but would open a gap between the first-person and the third-person perspective. That is to say, it may well be that the subject is ignorant of the real status of her thoughts or beliefs, which is in fact evident from a third-person perspective. This would have consequences with respect to first-person authority. For we usually take subjects to know the content of their thoughts. For instance, if a subject tells us that she thinks that water is wet we would like to grant her with at least knowledge of what she is thinking about, namely water. But, according to the object-dependency thesis, we should conclude that in a world without water, the subject would not know even that much. For she would just be under the illusion of thinking of water. In this sense, therefore, she would not have authority over the contents of her thoughts.²⁴ Be that as it may, this is not a problem we will address here. All we need to point out is that there are ways, though quite costly ones, of reconciling the idea that if there were no object perceptually given to one, then one would merely have an illusion of

²⁴See Burge 1994 and Davidson 1994 for ways of reconciling externalism and self-knowledge.

content, with the observation that, however, the subject “feels” and acts as if she had a thought with a singular content.

To recapitulate and conclude: we have seen, so far that, according to Evans, perceptual demonstrative thoughts are object-dependent. Moreover, the perceptual link between the subject and the object, though a necessary condition for demonstrative thought, is not a sufficient one. For both Russell’s Principle and the Generality Constraint must be met. We have considered these conditions in general and we have then considered their application in the case of perceptual demonstrative thought. Let us now turn to Evans’ defence of the idea that some uses of demonstratives are IEM.

3. *Immunity to Error Through Misidentification as a “Corollary” of Identification-Freedom*²⁵

As we have anticipated in chapter 2, Evans’ account of IEM is *revisionist*. That is to say, it does not conform to the following definition:

IEM_{df1} A singular judgement of the form [a is F], formed on the basis of a recognition or identification, is IEM, relative to the subject, iff it is not possible for the thinker to apply the particular singular concept [a] to any other object but the relevant object a.

According to Evans, IEM is a corollary of a judgement's being identification-free. Let us see how he spells out this notion. He holds that we can distinguish between two ways of gaining knowledge about the truth of a singular proposition: there is knowledge that is *identification-dependent* and knowledge that is *identification-free*. Notice, however, that this terminology can be somehow misleading, for it will turn out that judgements which are identification-free are nevertheless based on identification. But the point is that this is demonstrative identification (or individuation, in our terminology), which is not based on knowing the truth of an *identity judgement*.²⁶ This brings us to Evans' definition of identification-dependence, which goes as follows:

When knowledge of the truth of a singular proposition, [*a is F*], can be seen as the result of knowledge of the truth of a pair of propositions, [*b is F*] (for some distinct Idea, *b*) and [*a = b*], I shall say that the knowledge is *identification-dependent*: it depends (in part) on the second basis proposition, which I shall call the *identification component*.²⁷

For example, knowledge of [This man is the chairman of the Department] can be seen as the result of knowing the truth of a pair of propositions such as [NN is the chairman of the Department] and [This man = NN].

²⁵Evans 1982: 183: "Now there is a general connection between identification-freedom (in the narrow sense), with its *corollary* (italics mine) of immunity to error through misidentification, and the presence of an informational component (...)"

²⁶Cf. fn. 1.

²⁷Evans 1982: 180.

Now, before moving on to the rest of Evans' account, we should point out that he does not spell out exactly what knowledge of the pair of propositions requires. In particular, it does not seem right to suggest, as Evans' passage may suggest, that one gets to knowledge of a certain singular proposition by *actually entertaining* a pair of other singular propositions of which one knows the truth. Rather, it is enough that one's judgement be rationally grounded on the non necessarily occurrent belief in that pair of propositions. Obviously, then, EM would arise just in case the identification component is mistaken.

Notice, moreover, that if identification-dependency is the route to the possibility of EM, it is preferable not to phrase it in terms of knowledge. For, EM arises precisely in the case where there is identification-dependence but the identification component is mistakenly believed to be true. So, we must have a formulation of identification-dependency which is clearly *compatible with the fact that a judgement be identification-dependent and also affected by EM*. If we avoid phrasing identification-dependency in terms of knowledge, then it becomes evident that it would not be right to suggest that if the subject is mistaken about the identification component she is not entitled to any belief whatsoever. For, obviously, she would not know the singular proposition [a is F], but she could still be entitled to believe [Something is F], i.e. she could still hold the predication component.

We can thus put forward the following definition of identification-dependency:

A judgement [a is F] is identification-dependent when it is causally and rationally based on a pair of not necessarily occurrent beliefs, [b is F] (for some distinct Idea b) and [a = b].²⁸

Such a definition of identification-dependency is then clearly consistent with the possibility of the subject's being mistaken (at least) about the identification component, thus entertaining a judgement which is affected by EM. Let us now turn to identification-freedom.

According to Evans, to first approximation, identification-free judgements are the ones that are *not* dependent on knowledge of such a pair of propositions. But this would make non-singular judgements identification-free. For consider the descriptive name 'Julius', i.e. 'Julius' =_{df} the inventor of the zip: in such a case, according to Evans, knowledge of the truth of [Julius is the inventor of something] would not depend on knowledge of the truth of [X is F] and [X = Julius].²⁹ In fact, it is not obvious that this is as much as true. For, after all, it is because the subject knows that Julius = the inventor of the zip, that she can know that Julius is the inventor of something. However, it is clear that Evans is interested in IEM of *perceptual* demonstrative judgements. Hence, the narrower notion of identification-freedom he is introducing, if not motivated by general theoretical reasons, would anyway be more adequate for the purpose of accounting for IEM of perceptual demonstrative judgements.

Evans' narrower notion of identification-freedom is defined as follows:

²⁸Of course the judgement is occurrent and the belief is dispositional.

²⁹Cf. Evans 1982: 180-181.

[K]nowledge of the truth of a singular proposition is *identification-free in the narrow sense* if (i) it is not identification-dependent and (ii) it is based on a way of gaining (perceptual) information from objects.³⁰

In this case too Evans must leave room for the singular proposition believed to be immune to a certain form of error, namely EM, while maybe mistaken as far as the predication component is concerned. Hence, to make this clearer, it is preferable to avoid talk in terms of *knowledge* of the truth of a singular proposition. Thus, it would be enough for the notion of narrow identification-freedom to read as follows:

The judgement [a is F] is identification-free in the narrow sense, if (i) it is not identification-dependent and (ii) it is based on a way of gaining (perceptual) information from objects.

Hence, the judgement [a is F] would be IEM in the narrow sense if (i) it is not based on an identification component and (ii) it is based on a way of gaining (perceptual) information from objects.

Notice that it is a consequence of Evans' theory about perceptual demonstrative-thoughts that they can be identification-free in the narrow sense. For, according to Evans, perceptual demonstrative thoughts are based on a way of gaining information directly from the objects. Consequently, the subject can know that [This is F] is true *directly*, i.e. without basing her belief on any identity proposition. In fact Evans writes:

³⁰Evans 1982: 181.

We have the possibility of knowledge which is identification-free (in the narrow sense) when a subject is in a position, on receipt from an object (place) of information warranting ascription to something of the property of being *F*, to ascribe that property to that object (place) without relying on an *identification* [fn. In the sense of an identification component; *not* in the sense required by Russell's Principle (...)] of the object (place) from which he receives the information.³¹

According to Evans, identification-freedom is a notion of the utmost epistemological importance. For it is at the basis of the possibility of knowledge, since either knowledge of a proposition is identification-free, or, on pain of an infinite regress, it must ultimately rest on knowledge of a proposition which is identification-free.³²

According to Evans, IEM is a corollary of identification-freedom in the narrow sense. For, if a judgement is the result of knowledge which is identification-free, then the subject immediately knows not only that a property is instantiated, but also that it is instantiated by a certain object perceptually given to her. Hence, identification-free knowledge does not leave room for wondering which object is instantiating the property in question. Consequently, it does not leave room for a mistake in individuating which object is *F*.

In fact, Evans' main point is that in *normal conditions* the deliverances of the subject's perceptual system will provide her with *immediate* knowledge both of the fact that some property is instantiated and of the fact that it is instantiated in

³¹Evans 1982: 183. *Ibid.* fn. 56.

³²Cf. Evans 1982: 181. Notice that this can hold despite our amendments because *true* identification-free judgements would be known and would be at the basis of the possibility of knowledge.

the object concerned (provided that she has an adequate Idea of the object in question).³³ Hence, if the subject's perceptual system is working in a reliable way, and conditions are normal, if it looks to her as if an object in front of herself is red, then she will have *ipso facto* knowledge of the truth of the proposition [That is red]. By contrast, if she is hallucinating a red object, then she will just have an illusion of content. That is why, according to Evans we should realise that some demonstrative judgements are IEM. In our terminology, we can see how Evans is driving home the point that if a demonstrative judgement is merely based on the individuation of a particular, then the informational-link the subject has with the latter is such to give her knowledge of [That is F] which does not rest on knowledge of any identification-component (provided that she also meets RP1 and GC). Therefore, her judgement will be IEM.

However, it would not make any difference to the issue of IEM, if we read identification-freedom as not mentioning knowledge. For it may perfectly well be the case that the subject's perceptions are inaccurate and she thinks that there is something red when there is something orange instead. But that would not mean that her judgement is not immune to error through misidentification relative to the subject. That is to say, if she has information (hence, there is an object) as if there were a red object, she *ipso facto* knows that there is

³³Evans 1982: 181: "The way of gaining information from something (an object or a place) with which such an Idea is associated will, in certain circumstances (*normal* circumstances) yield knowledge that some property, say that of being *F*, is instantiated; and, provided that the subject has an adequate Idea of the object (place) concerned, this will *ipso facto* constitute knowledge that *that object (place) is F*".

something, i.e. that particular object, but she can be mistaken with respect to whether it really is red.

It should also be pointed out that satisfaction of condition (ii) of Evans' definition of identification-freedom is not by itself sufficient for making a judgement identification-free. For, obviously, a judgement can be based both on a way of gaining perceptual information from an object *and* on an identification component (e.g. the case of [That man is the head of the Department] we saw above). Hence, satisfaction of condition (ii) is sufficient for identification-freedom only if conjoined with satisfaction of condition (i). In other words, identification-freedom in the narrow sense can be rephrased as follows :

The judgement [a is F] is identification-free in the narrow sense iff it is *solely* based on a way of gaining (perceptual) information from objects.

Obviously, it is a consequence of Evans' definition of identification-freedom in the narrow sense that not only can demonstrative thoughts (solely) based on perceptual individuation of a particular be IEM, but also that all other judgements (solely) based on a way of gaining information from objects are IEM. For this reason, as we will see in the following chapters, many I-judgements too are IEM, according to Evans. Let us now turn to Evans' account of them.

Chapter Four

An Interlude: Evans on I-Thoughts and the Objectivity Requirement

Evans' discussion of I-thoughts is so intricate and complicated that, for expository purposes, it is best to break it down in two parts. Hence, in this chapter we will present the basic features of Evans' proposal, and also give some historical background to it. In the next chapter we will consider Evans' more contentious claims about I-thoughts which are meant to provide both a revisionist account of IEM of relevant self-ascriptions and a substantive response to the trilemma. The structure of this chapter will be as follows: we will present Frege's account of I-thoughts, which is the ancestor of Evans' views. We will then consider the dispute between Perry and Evans about I-thoughts, which will allow us to clarify some background assumptions in Evans' treatment of I-thoughts in *The Varieties of Reference*, which we have not yet examined. Finally, we will expound Evans' reading of the Objectivity Requirement with regard to I-thoughts.

1. I-Thoughts

Evans' discussion of self-identification follows the general lines he has set forth for perceptual demonstrative-identification we presented in the previous chapter. Hence, I-thoughts turn out to be *object-dependent* and *irreducible* to any other kind of thought, whether descriptive or containing other allegedly more fundamental indexical elements (e.g. here and this). According to Evans, these features of I-thoughts are shown in action and reason-giving procedures. In addition, Evans maintains that the *information-links* between the subject and her body are *necessary but not sufficient conditions* for I-thoughts and that conformity to *Russell's Principle* and the *Generality Constraint* is also required in order for I-thoughts to be possible. The former requires the subject's capacity to locate herself (but recall the specifications introduced in Ch. 3) relative to the objects in her environment; whereas the latter requires one's Idea of oneself to enter into indefinitely many thoughts about oneself. In order to understand better Evans' account of I-thoughts it is necessary to look at its sources, namely Frege and, to a certain extent, Perry. We will now consider these in turn.

2. I-Thoughts: Frege

Evans inherits Frege's conception of I-thoughts, to a certain extent. Frege agreed with the widespread view according to which the essence of self-consciousness is self-reference, i.e. the ability to refer to oneself either in speech or in thought by means of the first-person pronoun. Since, however,

according to Fregean semantics, reference has to be mediated by the appropriate kind of thought about the referent, an inquiry into the nature of I-thoughts will automatically be an inquiry into the conditions of possibility of self-consciousness (there is more about this issue below).

Now, the reason it is difficult to attain a proper understanding of I-thoughts is that, if we are persuaded that a thought is made out of senses and, therefore, that 'I' must have a sense, then we have to account for this sense. Yet, this is far from being a straightforward matter. For, as Frege himself noticed, the sense of 'I' is never equivalent to the sense of any other coreferring expression. In order to see why, it is sufficient to run *Frege's test for difference of sense*.

Consider the case of Oedipus saying 'The killer of Laius must be brought to justice', when, unbeknownst to him, he is the killer of Laius.¹ Alternatively, consider the case of John Smith who speaks of the fortune inherited by the person named in a will, John Horatio Auberon Smith, without realising that he is talking about himself.² Then consider the case of a man who is seeing someone reflected in a mirror and says 'That man is sitting on a chair', when, unbeknownst to him, it is he himself who is reflected in the mirror.

In all these cases, the subject is intentionally referring to the person the definite description/proper name/complex demonstrative refers to and yet he can still fail to realise that he is talking about himself. Consequently he would assent to sentences such as 'The killer of Laius must be brought to justice', 'JHAS has inherited a fortune', 'That man is sitting on a chair', but he would not

¹Cf. Evans 1982: 206-207.

²Cf. Anscombe 1990: 135-137.

assent to the corresponding first-person sentences, although they are about the same referent. The different attitude the subject would take towards these sentences is explained, within a Fregean semantics, as due to the fact that the referent is presented differently. Thus, the mode of presentation or sense associated with 'I' is different from the mode of presentation associated with any other coreferring singular expression. In other terms, we cannot provide 'I' either with a descriptive sense, or with a sense reducible to the sense of any other expression which has the same reference as 'I'. Moreover, that there is a difference between the mode of presentation associated with 'I' and that associated with any other coreferring expression can be seen by considering the role I-thoughts play with respect to action and reason-giving procedures.³ Consider the case of Oedipus: before realising that he himself was the killer of Laius, he searches for the killer throughout his kingdom. Once he realises that he himself is the killer of Laius, Oedipus abandons the city of Thebes in order for it to regain its prosperity. Thus, Oedipus can justify his action by saying that he has done so because he wanted Thebes to regain its prosperity and because he knew that it was his presence that was causing trouble. Obviously, none of these justifications could have been used in order to explain Oedipus' behaviour before he realised *he* was the killer of Laius, although, before realising that, he could think of the killer of Laius, plan how to take action against him, and so forth.⁴

³Evans 1982: 207, Perry 1990, 1994. Cf. also Peacocke 1983: 127-sq.

⁴In fact this connection between I-thoughts and action was first noticed by Perry 1990, 1994.

This kind of considerations can lead one to conclude, with Frege, that:

Everyone is presented to himself in a particular and primitive way, in which he is presented to no-one else. So, when Dr. Lauben thinks that he has been wounded, he will probably take as a basis this primitive way in which he is presented to himself. And only Dr. Lauben himself can grasp thoughts determined in this way. But now he may want to communicate with others. He cannot communicate a thought which he alone can grasp. Therefore, if he now says 'I have been wounded', he must use the 'I' in a sense which can be grasped by others, perhaps in the sense of 'he who is speaking to you at this moment', by doing which he makes the associated conditions of his utterance serve for the expression of his thought.⁵

Notice that these private I-thoughts are nevertheless thoughts, according to Frege, and not ideas (in Frege's sense), that is, private representations of oneself, including sense-impressions deriving from one's own body.⁶ But how should we conceive of them?

Let us recapitulate the features of I-thoughts, according to Frege.

- (a) They are primitive;
- (b) They are private to their thinkers;

⁵Frege 1918, reprinted in Harnish 1994. Page reference to the latter. Frege 1994: 524.

⁶The difference between thoughts and ideas or representations (*Vorstellungen*) is crucial to Frege. The latter are private representations which need a bearer in order to exist. Frege does not deny that they exist, but he denies that communication and science would be possible if they were what we try to convey by means of our words. For ideas are necessarily private to their bearers and no science or common knowledge would be possible if we were actually talking about our own representations.

(c) They are incommunicable;

(d) They are objective.

On the one hand, (a) amounts to saying that they are not equivalent to any thoughts containing any other mode of presentation of the referent. That is to say, they are irreducible.

On the other hand, (b) allows for two possible interpretations: (b.1) there are as many I-thought types, i.e. as many modes of presentation of the self, as there are individuals. We can illustrate the consequences of this position by quoting Husserl:

The word 'I' names a different person from case to case, and does so by way of an ever-altering meaning. What its meaning is at the moment, can be gleaned only from the living utterance and the intuitive circumstances which surround it (...). In solitary speech the meaning of 'I' is essentially realized in the immediate idea of one's own personality (...). Each man has his own I-presentation (...) and that is why the word's meaning differs from person to person.⁷

(b.2) There is just one I-thought type, or mode of presentation of the self, which is exemplified in as many I-thought tokens as there are individuals. The idea has been clearly stated by Peacocke, who writes:

Suppose Peter thinks 'I am hungry' and Paul thinks 'I am hungry'. Since we are following the Fregean model, the modes of presentation they each express by 'I' must be different: for they determine different objects, Peter and Paul respectively. But of course when they think these thoughts, Peter and Paul think of themselves in the same *type* of way. It is this type which is denoted by '[self]'. The constituent of all Peter's first-person thoughts is called a token mode of presentation; it

⁷Husserl 1970: 315-316.

can be taken to consist of the type [self] indexed by Peter, that person himself. This token m.p. will be denoted by '[self_{Peter}]'. So it is token m.p.'s which are constituents of thoughts and which pick out particular objects. But one should not be misled by the word 'token': there is nothing relevantly unrepeatable about [self_{Peter}]. On many different occasions, Peter may have attitudes to thoughts which contain it as a constituent.⁸

It is not clear whether Frege held (b.1) or (b.2). As we shall see, however, (b.1) and (b.2) have a different bearing upon whether I-thoughts are shareable or not. In fact, I-thoughts conceived of as in (b.1) are not communicable because (c) they are private to their thinkers.⁹ What about I-thoughts conceived of as in (b.2), then? (b.2) in fact says that there is a unique mode of presentation of the referent each of us grasps when she is entertaining an I-thought. However, depending on how we think of this mode of presentation, the resulting I-thoughts will be communicable or will fail to be so.

Now, as far as (d) is concerned, according to Perry, once *communicability* is given up, then the *objectivity* of thoughts is impaired too. For, according to his reading of Frege, the former is a necessary and sufficient condition for the latter. By contrast, Evans points out that communicability is only a sufficient but not a necessary condition for objectivity.¹⁰ Hence, if communicability is impaired, this does not mean that objectivity is as well. We will consider whether

⁸Peacocke 1983: 108. Cf. also Peacocke 1983: 119-121.

⁹This would square with Frege's suggestion that in order for communication to be possible some "public" sense must be attached to 'I'. Hence, this would be indirect evidence in favour of the fact that Frege held (b.1).

¹⁰Cf. Evans 1985: 313-sq.

Evans succeeds in maintaining the objectivity of I-thoughts despite their incommunicability or non-shareability, after considering Perry's conception of I-thoughts and the Objectivity Requirement.

3. *I-Thoughts: Perry*

According to Perry, if we want to maintain the objectivity of I-thoughts, that is, their being intersubjectively accessible, then we should not follow Frege in equating the sense of a sentence with the thought it expresses. Consequently, we should not equate the sense of an expression such as 'I' with its contribution to the thought expressed by a sentence containing it. Perry's reading of (b.2) is to take the sense of 'I'—its *role*, as he names it—as a *function* from a *context* to a *value*—its referent. In the case of 'I' its role would be given by the token-reflexive rule—'I' refers to the utterer of this token—, which would lead us, varying from context to context, to its referent. Such a referent, in turn, would be a constituent, together with the referent of a predicate expression (*F*), of the thought expressed by a sentence containing 'I'. While Perry is departing from Frege's doctrine both with respect to senses (since Fregean senses take us directly to references, and not from contexts to references) and thoughts (given that P(erry)-thoughts contain objects and properties, and not modes of presentations of them), his account allows us to maintain the intersubjectivity of thoughts. For, by applying the token-reflexive rule, both the speaker and the utterer will grasp the same thought. Hence, for Perry the objectivity of I-thoughts is guaranteed by the fact that these thoughts are about a person, i.e. an

embodied entity which can be perceived by other people. Moreover, it is guaranteed by the fact that such thoughts are entertained by entertaining the *linguistic meaning* of 'I', which is publicly and intersubjectively accessible too. Let us now turn to Evans' criticism of Perry's account.

4. *I-Thoughts: Evans*

Evans' reading of (b.2) is as follows: there is one mode of presentation of the self as a type, which is entertained by each subject as a token when she thinks about herself in the first person way. Moreover, when one entertains an I-thought, one has a thought about an object, i.e. the *person* one is, which, however, is presented in the first-person way. Hence, the constituents of the thought are an object together with its mode of presentation and the mode of presentation of a property.¹¹

However, it is important to stress that, according to Evans, the mode of presentation of the self is not given by any linguistic rule governing the use of 'I' as a referring device. For—he argues—what has the token-reflexive rule to do

¹¹Cf. Evans 1985: 315-321. Notice that in Evans 1985 the thought is taken to be constituted by the object and its mode of presentation *together*. By contrast, McDowell's reading of *The Varieties of Reference* has made popular the reading of Evans' theory of singular thoughts as containing only senses, which, however, could not exist if their objects did not exist. The difference between Evans 1982 and Evans 1985 is interesting but not fundamental to our purposes.

with the subject's capacity to think of herself self-consciously?¹² The problem, in his opinion, is that such a rule "plays no part in an explanation of what makes a subject's *thought* about himself".¹³ In other terms, it does not seem to be plausible that, when a subject is entertaining an I-thought, she is thinking of herself as the utterer of that token of 'I'.

A first objection to the idea that the mode of presentation of the first person is to be equated with the linguistic meaning of 'I' is that a subject can entertain I-thoughts even if she is not uttering any sentence. Thus, it would not make sense to suggest that she is thinking of herself as the *utterer* of such an expression. A second objection would be that there are languages which do not have the first-person pronoun, but, presumably, speakers of those languages can nevertheless entertain I-thoughts. Consequently, these thoughts cannot be dependent on a linguistic convention.

In order to meet these objections, one could suggest a different token-reflexive rule, namely, that *I* refers to the *thinker* of this token of the first-person concept. But now, according to Evans,¹⁴ it would seem that the conception we have of ourselves is such that we take ourselves to be mere authors of thoughts.¹⁵ This would have unacceptable consequences for Evans, as we will

¹²Cf. Evans 1985: 320.

¹³*Ibid.*

¹⁴Cf. Evans 1982: 213, 252.

¹⁵It is not obvious at all that, contrary to what Evans maintains, the problem with the revised version of the token-reflexive rule is that one would think of oneself as the author of a certain thought. Rather, it seems that one should *already* be able to

presently see. Therefore, Evans concludes that the mode of presentation associated with 'I' must differ from the token-reflexive rule.

To recapitulate: Evans wishes senses to figure within the content of thought, thus giving rise to a form of "cognitive externalism".¹⁶ This is so because Evans was impressed by Frege's puzzle and, in particular, by the fact that one and the same speaker can believe and disbelieve, or be agnostic about, one and the same content, if it is presented in a different way. Moreover, he was impressed by Perry's insight that I-thoughts can have altogether different implications for action from thoughts with the same broad content, which, however, are not in the first-person mode of thought.¹⁷ Thus, modes of presentation are at least *cognitively* relevant. However, Evans does not think of the mode of presentation as either a definite description, or as an entity which mediates the access of the subject to the object, like a sense-datum. Hence, Evans rejects both a descriptivist position about sense and a sense-data theory. Rather, the notion of sense or of mode of presentation is designed to capture the intuition that any object is always known from or under a certain *perspective*, i.e. even the most

entertain the first-person concept. Hence, to think of oneself as the author of a certain token of [I] would be a *circular* explanation of the first-person concept, for it would presuppose what it seeks to explain. Cf. Bermúdez 1998.

¹⁶According to externalism the content of a thought is determined by the objects it is about. "Cognitive externalism" is a kind of externalism according to which the content of the thought is also determined by the way in which these objects are given to the subject.

¹⁷Two thoughts have the same broad content if they are about the same objects. For instance [I am F] and [AC is F] have the same broad content.

immediate access to an object, that is the perceptual contact with it, is always perspectival, depending on our position with respect to the object itself. Therefore, our access to any object, no matter how immediate it may be, i.e. no matter how non-mediated by either a conceptual repertoire or by any mediating entity, can never disclose to us the object with *all* its properties. Thus, Evans seems to think that the perspective from which an object is given to us is constitutive of the way in which we can think of that object, because the very content of our thought, i.e. that very object and some of its properties, are given to us according to that particular perspective. It has to be stressed, however, that the perspective from which an object is given is not a further ingredient, beside the object, which is part of the thought. Rather, the idea is that objects of thought are always given from a certain perspective and, as such, are constituents of thoughts.

However, the *aboutness* of the thought, i.e. its content, is always entirely determined by the object, without further ado, and it is the fact that various thoughts about X converge onto the same object which guarantees the possibility of communication, despite the various ways in which that very object can be given to different subjects or to the same subject on different occasions. That is to say, according to Evans, communicability is not secured by the fact that the subjects entertain the same modes of presentation of the referents. Rather it is secured by the fact that they talk about the same object, no matter how they think of it.¹⁸ Consequently, according to Evans, on the one hand, the

¹⁸Cf. Evans 1982, Ch. 11. Of course, it can be problematic to claim that convergence onto the same object guarantees successful communication. For if—as Evans

fact that the mode of presentation is a constituent of the thought does not impair the possibility of communication of that thought and, on the other hand, it allows for an explanation of the different attitudes a speaker may take towards one and the same content, when it is presented differently.

In the case of I-thoughts, however, the perspective from which the object is given is cognitively even more effective than in other modalities of singular thought (e.g. perceptual demonstrative thought). We can better see why this is so if we consider that the aboutness of the thought is entirely determined by the object the thought is about. Thus, the thoughts [AC is F] and [I am F] do not differ with respect to what they are about, i.e. they do not differ in broad content. But, after all, there is a difference between these two thoughts and it is a difference that can be reflected in beliefs, actions and justificatory procedures, as we have seen. This difference can be summarised as a difference between thoughts which are about oneself and thoughts which are about oneself *self-consciously*, i.e. which manifest the fact that the subject is aware that they are about herself. So, in the case of I-thoughts, the relevance of the notion of mode of presentation or sense of the first-person can be better appreciated as being introduced in response to the question "What makes a thought about myself a self-conscious thought?". Thus, an elucidation of the various ways in which we

maintains—sense does not determine reference and, moreover, speakers are allowed to entertain different senses, then it can look mysterious how they can actually be sure that they are talking about the same object. This, however, is a problem which does not directly concern us here.

have access to ourselves which can lead to I-thoughts becomes an elucidation of self-consciousness.

5. *Evans and The Objectivity Requirement*

We saw before that, according to Evans, a thought which is not communicable because its sense-constituents are not intersubjectively accessible can still be objective. But what does it mean to say that a thought is objective if not that it is (completely) communicable? According to Evans, the notion of *objectivity* has to be re-interpreted as meaning that, no matter whether thoughts are communicable, they are objective provided they are about objects *thought of as* “elements in the objective order of things”, i.e. thought of as objects amongst others, or on a par with others, and as located in space and time.¹⁹

OR: A thought about an empirical object is objective iff the object is thought of as an empirical object which is located in space and time and which exists independently of being perceived.

Obviously OR will be specified in a different way if thought about places, or times or abstract objects, etc. is at stake. However, when applied to I-thoughts OR means that we should conceive of ourselves as spatio-temporal continuants. Thus, Evans aims to develop a conception of I-thoughts such that they turn out to be primitive, i.e. irreducible to other forms of thoughts; private,

¹⁹Evans 1982: 256. Cf. also Evans 1982: 210, 211.

i.e. entertainable only by their respective subjects, although understandable also by others (once expressed); and still objective, i.e. about a subject who thinks of herself as a physical object located in space and on a par with other objects. That is to say, one should think of oneself as a physical entity which exists in space independently of one's current awareness of one's own location and of the spatial features of one's body. This is also why Evans cannot be entirely content with the version of the token-reflexive rule we discussed above, according to which we think of ourselves as authors of certain thoughts. For, if that is the case, then it is no longer *mandatory*, in order for us to have an adequate Idea of ourselves, to think of ourselves as embodied and located in space. Of course, the latter position would be compatible with holding that, in fact, we also do think of ourselves as embodied. But it would not make I-Ideas inadequate, if they were not comprising a conception of ourselves as embodied. We will come back to this issue in the next chapters.²⁰

However, why should one hold that for I-thoughts to be objective they must be about an object in the world, i.e. a person, who *thinks* of herself as such? It has to be acknowledged that there is not much in the way of an explanation of this tenet in Evans. According to Cassam, the Objectivity Requirement, as it is interpreted by Evans, stems from the conviction that, since we are persons, and I-thoughts are object-dependent, in order to meet Russell's Principle and the

²⁰ In Ch. 7 we will argue that Evans' characterisation of the first person concept is compatible with more "mentalist" ones such as Peacocke's.

Generality Constraint, we must think of ourselves as such.²¹ Hence, it is because we are persons that meeting RP requires us to be (at least dispositionally) able to locate ourselves and meeting GC requires us to think of ourselves by exploiting the category of person, thus satisfying OR.

Hence, Evans wishes to show that a theorist who assumes that we are human beings can claim that in order to have the first-person concept we must think of ourselves as persons. By so doing, such a theorist will be able to point out that the Cartesian and the Idealist make use of a *different* first-person concept, which is not the concept of an embodied entity. Hence, far from teaching us something about the first-person the Cartesian and the Idealist would just conceive of it in a different way, which, however, is irrelevant to an understanding of *our* first-person concept.

It should also be stressed that Evans is not obviously begging the question against the Cartesian and the Idealist by assuming that the self is identical with a human being. For the dispute between them is at the level of *sense*, not of reference. In fact, even on the assumption that selves are identical with human beings, if it turned out that one's first person concept did not depend on one's

²¹Cf. Cassam 1997a. Notice that Cassam only talks about RP. But GC is indeed necessary in order to get to the conclusion that since we are persons we must think of ourselves as such. For it is only when we entertain I-thoughts that conform to GC—such as [I was born in Milan in 1973] or [I will die]—that, in thinking about ourselves, we exploit a conception of ourselves as persons on a par with all others, i.e. as persons that exist through time and exist even when they are not able to perceive themselves, etc. There shall be more about this issue in Ch. 5.

conception of oneself as embodied, then the Cartesian and the Idealist uses of the first person would be intelligible. Thus, it seems fair to say that Evans makes use of an *indirect strategy* against the Cartesian and the Idealist which is meant to show that their use of the first person deviates from ours to such an extent that it becomes unintelligible from our stand point. Let us now turn to the details of Evans' proposal.

Chapter Five

Evans: Immunity to Error through Misidentification of I-Thoughts and the Substantive Response to the Trilemma

We have seen in chapter three that a theory of singular thought should specify what satisfaction of Russell's Principle requires in each case, i.e. in the case of perceptual demonstrative thoughts, here-thoughts, I-thoughts, thoughts about times and abstract objects. In particular, we have seen that in the case of singular thoughts about particulars perceptually given to us, satisfaction of RP requires us to be able to locate the objects relative to ourselves. Moreover, we have seen that, according to Evans, perception is often such as to enable us *immediately* to know the object's location relative to ourselves. The corresponding demonstrative thought is thus IEM, for it is not based on any identification (in our terminology) of the object, but on its mere perceptual individuation. Hence, if [This is F] is a IEM thought, then it can be mistaken as far as the predication component is concerned, but it cannot be mistaken as to whether it is *that* particular object which is *F*.

According to Evans, on the assumption that the self is embodied, satisfaction of RP in the case of I-thoughts requires us to be able to locate ourselves. As we shall see, according to Evans, this capacity consists in the practical ability to locate oneself as an object with a certain position and orientation in space on

the basis of the perception of one's environment and the awareness of the spatial features of one's body afforded by somatic proprioception.

It is important to stress that these abilities must be *dispositionally* in place. Hence, a subject who were in a state of sensory deprivation and were thus prevented from exercising such abilities could still be said to have an adequate Idea of herself provided that she retained the ability (manifested in the past) to locate herself on the basis of the perception of her environment.¹ Hence, Evans writes:

It is essential, if a subject is to be thinking about himself self-consciously, that he be *disposed* to have such thinking controlled by information which may become available to him in each of the relevant ways.²

On this view, the Cartesian and the Idealist, who think that there can be I-thoughts even if these thoughts are not dependent on an awareness of oneself as embodied and located in space, would confuse "the possibility of a non-activated disposition with no disposition at all".³ Hence, they would be mistaken in concluding that, since sometimes I-thoughts can fail to depend on an awareness of oneself as embodied and located in space, such knowledge is not essential to the very possibility of I-thoughts.

¹Evans 1982: 215: "A subject may be amnesiac and anaesthetized, and his senses may be prevented from functioning; yet he may still be able to think about himself, wondering, for example, why he is *not* receiving information in the usual ways".

²Evans 1982: 216.

³O'Brien 1995: 237.

We have also seen in chapter four that, according to Evans, I-thoughts must satisfy the Objectivity Requirement. So, for an I-thought to be objective, it must be the case that the subject thinks of herself as a person on a par with all others. Hence, on the one hand, she must think of herself as an embodied entity, which has the properties of physical objects. For instance, it cannot (wholly) exist at two different places at the same time, it exists through time and, finally, it exists even when it is not presently perceived either by others or by the subject herself. On the other hand, the subject must be able to conceive of herself as an entity with psychological properties that can be interpreted and made sense of also by others.

Now, it is important to stress that *for OR to be met, both GC and RP must be satisfied and satisfied in the appropriate way*. For, as we will see, satisfaction of RP in the case of first-person thoughts based on outer perception and somatic proprioception can at most give the subject a conception of herself as having a body which is *presently* perceived and thought of as located. But OR requires one to think of oneself as having a body located in space *independently* of the fact that one is now perceiving it and is presently able to locate it. Hence, satisfaction of OR requires GC to hold. For, according to GC, the subject should be able to entertain judgements like [I (am the person who) was born in Milan in 1973], [I (am the person who) was breast-fed for the first two months of her life], etc. Now, the relevance of these judgements is that they can only be entertained by someone who can make sense of, for instance, someone else's *testimony* about herself. Thus, they can only be entertained by someone who

takes herself to be a person on a par with all others, whose existence is independent of both her current perceptions and thoughts about herself.

The structure of this chapter will be as follows: we will present Evans' account of IEM of bodily and physical self-ascriptions, first. Then, we will consider his arguments against the claim that all mental self-ascriptions are IEM. We will then turn to his attack against the Cartesian conception of the mind and to his defence of the idea that by being aware of their own mental properties subjects must also conceive of themselves as embodied entities. We will agree with Evans that the deliverances of somatic proprioception and perception can give one immediate knowledge of oneself as embodied and located in space. But we will also claim that Evans does not explain why mental self-ascriptions which are IEM are so, and, moreover, are so in any logically possible world. Hence, his account should be supplemented with such an explanation. This, in turn, will be the topic of the following chapters.

1. *IEM Relative to Physical Self-Attributions As "The Most Powerful Antidote to a Cartesian Conception of the Self"*⁴

Recall Evans' definitions of identification-dependent and identification-free knowledge of the truth of a proposition we introduced in Ch. 3. Here they are:

When knowledge of the truth of a singular proposition, [*a* is *F*], can be seen as the result of knowledge of the truth of a pair of propositions,

⁴Evans 1982: 220.

[*b* is *F*] (for some distinct Idea, *b*), and [*a* = *b*], I shall say that the knowledge is *identification-dependent*: it depends (in part) on the second basis proposition, which I shall call the *identification component*.⁵

[K]nowledge of the truth of a singular proposition is *identification-free in the narrow sense* if (i) it is not identification-dependent and (ii) it is based on a way of gaining information from objects.⁶

Recall also our interpretation of Evans' definition we introduced in chapter three to make clear the compatibility between a judgement's being IEM and yet be mistaken as far as the predication component is concerned. Here they are:

A judgement [*a* is *F*] is identification-dependent when it is causally and rationally triggered by a pair of (not necessarily occurrent) beliefs, [*b* is *F*] (for some distinct Idea *b*) and [*a* = *b*].

A judgement [*a* is *F*] is identification-free if (i) it is not identification dependent and (ii) it is based on a way of getting information directly from objects.⁷

Hence, an I-judgement will be IEM if it is identification-free in our sense. By contrast, it will be liable to EM if it is identification-dependent in our sense.

⁵Evans 1982: 180.

⁶Evans 1982: 181.

⁷All the constraints on these definitions and the reasons behind them we introduced in chapter three are still valid here. Recall also that the latter definition can also be phrased as follows: A judgement [*a* is *F*] is identification-free iff it is solely based on a way of gaining information directly from objects.

According to Evans, there can be bodily self-ascriptions that are IEM. He writes:

There is a way of knowing that the property of ξ 's hair being blown by the wind is currently instantiated, such that when the first component expresses knowledge gained in this way, the utterance 'The wind is blowing someone's hair, but is it my hair that the wind is blowing?' will not make sense.⁸

Given our interpretation of identification-freedom this passage should read as follows: there is a way of coming to make the judgement [The wind blows my hair about], such that the claim 'The wind is blowing someone's hair, but I wonder whether it is my hair that the wind is blowing' is not a reasonable reaction to the epistemic situation the subject is in. Moreover, Evans claims that "the property of being immune to error through misidentification is not one which applies to propositions *simpliciter*, but one which applies only to judgements made upon this or that basis".⁹ Hence, these two claims together imply the following: there are particular *grounds* on which, if one judges [The wind blows my hair about], then the claim 'The wind is blowing someone's hair, but I wonder whether it is my hair that the wind is blowing' will not make rational sense.¹⁰ Thus, we have to investigate which *grounds* would lead to IEM bodily self-ascriptions.

⁸Evans 1982: 218.

⁹Evans 1982: 219.

¹⁰Precisely because Evans holds that identification-free and identification-dependent judgements are a function of the grounds on which judgements are based, it would be misleading to say, as he says, that the utterance 'The wind is blowing someone's hair,

1.1 IEM of Bodily Self-Ascriptions Based on Somatic Proprioception

According to Evans, we have “a general capacity to perceive our own bodies” which “can be broken down into several distinguishable capacities: our proprioceptive sense, our sense of balance, of heat and cold, and of pressure”.¹¹ According to him, when a judgement is based upon this way of getting information from one’s body, it is also IEM. He writes:

There just does not appear to be a gap between the subject’s having information (or appearing to have information), in the appropriate way, that the property of being *F* is instantiated, and his having information (or appearing to have information) that *he* is *F*; for him to have, or appear to have, the information that the property is instantiated just is for it to appear to him that *he* is *F*.¹²

First of all, notice that Evans’ way of putting his point here suggests that he himself was aware of the fact that a judgement that is IEM does not necessarily amount to *knowledge*.

More importantly, however, in this connection Evans himself considers the possibility of *deviant causal chains*: i.e. we can think of a possible world in which subjects receive proprioceptive information from someone else’s body. For example: “the subject’s brain [is appropriately linked] with someone else’s body, in such a way that he is in fact registering information from that other

but is it my hair that the wind is blowing?’ will not make sense. For the utterance makes sense, i.e. we understand it, but is not a *rational* response to the epistemic situation the subject is in.

¹¹Evans 1982: 220.

¹²Evans 1982: 221.

body".¹³ The further assumption which, however, is not explicit, is that one retains one's own body.

Let us consider the implications of Evans' claims. In fact, Evans' first claim quoted above just amounts to holding that proprioceptive information can fail to be veridical. Hence, it can lead to false bodily self-ascriptions. This, in turn, implies that what he calls proprioceptive information can be *misleading* about the issue of whether it is true that one's body is *F*. Now, it seems reasonable to hold that such misleading information can only be something like an *appearing* to one as if one's legs were bent. In such an event, one would just have the illusion that one's legs are bent.

However, the fact that Evans considers the possibility of deviant causal chains may suggest the thought that bodily self-ascriptions based on somatic proprioception are not IEM after all. In fact, one may think that in deviant causal conditions one's judgement, e.g. [My legs are bent] could be false (at least) because the subject mistakenly identifies herself with the subject from whose body she is receiving information. Hence, the judgement would seem to be affected by EM. Notice that, given our way of understanding EM, this would mean that the judgement should be seen as grounded in a belief in a predication component such as [The legs I am feeling now are bent] and in a belief in an identification component such as [My legs = the legs I am feeling now]. In other words, for the judgement to be affected by EM it should be the case that one's concept [The legs I am feeling now], formed on the basis of the deliverances of somatic proprioception, and the concept [My legs], formed

¹³*Ibid.*

through an individuation of one's body as a physical entity with fixed spatial boundaries, were wrongly taken to be coinstantiated.

One could then suggest that the alleged counter-example could be blocked by endorsing a *phenomenological* conception of one's body, i.e. by endorsing:

(0) my body =_α the body I am proprioceptively aware of

As a consequence of holding (0), the judgement [My legs are bent], when based on the deliverances of somatic proprioception, would result as being IEM in any possible world, because it would be based on an identification component which is true by definition.

However, whether or not a phenomenological conception of one's body is a feasible position to hold, Evans could not endorse it.¹⁴ For, as we saw in Ch. 4, he subscribes to the Objectivity Requirement, according to which for one's conception of oneself to be objective, it must be a conception of an entity which exists independently of being perceived and which has fixed spatial boundaries, e.g. it cannot wholly be at two different places at the same time.

However, this is not to say that Evans cannot reply to the alleged counter-example. For what the deviant case shows is that in certain conditions the judgement [My legs are bent] *could* be based on a belief in an identification component such as [My legs = the legs I am feeling now]. Now, according to Evans, we should ask under which conditions this would be so. For here the

¹⁴As a matter of fact, the phenomenological conception of one's body would have problems in its own right. For instance, there are parts of our bodies that we cannot access proprioceptively, such as some internal organs.

relevant thought is that one's judgement's being EM/IEM is always a function of the *grounds* on which it is based. In turn, we have seen that, according to Evans, the grounds of a judgement must be causally and rationally efficacious. Hence, they must be at the subject's disposal. So the alleged counter-example merely shows that if a subject is receiving proprioceptive information from someone else's body and she *knows* or, more mildly, *has reasons to believe* that she may be receiving information from a body which is physically disjoint from the body in which her brain is located and judged [My legs are bent] on that basis, then her judgement would be identification-dependent and, consequently, affected by EM.¹⁵ For the false (non-necessarily occurrent) belief in an identification component such as [My legs = the legs I am feeling now] would be part of the subject's grounds for her judgement.

However, this does not show that in *normal conditions*, i.e. when one is receiving proprioceptive information from the unique body in which one's brain is located, the judgement [My legs are bent], based on the deliverances of somatic proprioception, would be based on a belief in the relevant identification-component.¹⁶ For, true, here there is an assumption, i.e. that conditions be

¹⁵Evans 1982: 221: "In the second place, there are problems about the Ideas that would be involved in the supposed identification component (...). Such an Idea [i.e. 'the body from which I hereby have information'] would certainly be involved in one's thinking if one *knew* (italics mine) one was in the abnormal situation described (...)"

¹⁶Notice, in turn, that this means that for Evans genuine proprioceptive information counts as information received from one's unique body, i.e. the unique body in which

normal, but this is not an assumption that should figure as a belief in an identification component which is part of *the subject's grounds for her judgement*.

In order to see why, let us consider a similar argument which is commonly used in favour of direct realism in philosophy of perception. On the assumption that conditions are normal, i.e. that one's sense organs are working properly and the world cooperates, a subject who has a visual experience as if there were a bottle in front of her would be immediately entitled to the belief [There is a bottle in front of me now]. What is relevant here is the thought that in order for the subject to be entitled to her belief she does not have to ground it in a belief to the effect that conditions are normal, even less in a belief that certain deviant conditions are not obtaining.

one's brain is located. This position has the merit to give us an *objective* account of somatic proprioceptive information which would be defined as follows:

(0*) somatic proprioceptive information =_{df} information solely received from the body in which one's brain is located and which is about that body

Those who think that genuine proprioceptive information could be at work in deviant causal conditions too must clearly depart from (0*). But, then, they should provide us with an alternative account of somatic proprioceptive information. Such an account could not make reference to the identity of the body from which the information derives. hence it should specify the *intrinsic features* of somatic proprioception. This can be far from being a straightforward matter. For the sake of the argument, however, let us concede that such an alternative account could be provided.

This has a bearing on our case too. For if one is receiving proprioceptive information in the normal way, i.e. from one's unique body in which one's brain is located, then one is immediately entitled to form the belief [My legs are bent] on those bases. That is to say, if conditions are normal, one can form that judgement without basing it on a belief in an identification-component such as [My legs = the legs I am feeling now]. Nor does one have to make sure that conditions are normal in order to be entitled to that belief.

Hence, Evans can rebut the alleged counter-example by saying that in this world bodily self-ascriptions based on somatic proprioception are identification-free and hence IEM. So they are at least *de facto* IEM, because *de facto*, i.e. in this world, they do not involve any identification component as part of the subject's rational grounds for her judgement. This, however, does not mean that they could never involve such identification components in logically possible worlds where the subjects were appraised of the particular situation they may be in. However, *de facto* identification-freedom and, consequently, *de facto* IEM is all Evans needs in order to make the point that *our* first-person concept is immediately as of embodied entities, *contra* the Cartesian and the Idealist positions.

As a consequence of Evans' treatment of bodily self-ascriptions based on somatic proprioception, we have that in this world these self-ascriptions are open to just *one* form of error. That is to say, if one's proprioceptive system is not working reliably, then it could give one an *illusion* as if one's legs are bent. Think for instance of the sensation of falling down the bed which one may have when one is waking up. Hence, one would be under the illusion of falling down and form the corresponding judgement [I am falling down]. In such an event,

the subject may coherently wonder whether her judgement [I am falling down] is true or is just illusory, where this would be tantamount to wondering whether one's proprioceptive system is working reliably. However, this is not the same as to wonder whether one's proprioceptive system is giving one information about *someone else's body*.

The conclusion we can draw from Evans' discussion is that if a judgement is based on the deliverances of somatic proprioception, then it is (at least *de facto*) identification-free and, therefore, IEM. For, in our world, proprioceptive information is such as to give us information relative to the unique body in which our brain is located, which is immediately relevant to the formation of a first-person thought. If, then, Evans is right in holding that this kind of judgements must be at least dispositionally in place for one to have the first-person concept, then he is right in claiming that our first-person concept is firmly anti-Cartesian and anti-Idealist.

1.2 IEM of Physical Self-Ascriptions Based on Outer Perception

According to Evans, there is another way of gaining knowledge of our physical properties beside being proprioceptively aware of our bodies, namely, through *outer perception*. He writes:

I have in mind the way in which we are able to know our position, orientation, and relation to other objects in the world upon the basis of our perceptions of the world. Included here are such things as: knowing that one is in one's bedroom by perceiving and recognizing the room and its contents; knowing that someone is moving in a train

by seeing the world slide by; knowing that there is a tree in front of one, or to the right or left, by seeing it.¹⁷

Once again, Evans thinks that all these judgements are identification-free and, consequently, IEM.¹⁸ For, these judgements are based on the subject's perception of the environment. Hence it cannot be the case that on these bases the subject knows that there is a given object in front of someone, without thereby knowing that that object is in front of her, and, hence, that she is in front of it.

Moreover, according to Evans, these ways of forming first-person judgements are fundamental to having an adequate Idea of ourselves by the standards of the Objectivity Requirement. He writes:

I suggested that our knowledge of what it is for [I am δ_t] to be true, where δ_t is a fundamental identification of a person (conceived of, therefore, as an element of the objective spatial order), consists in our knowledge of what it is for us to be located at a position in space. (...) I argued that this in turn can be regarded as consisting in a practical capacity to locate ourselves in space by means of the kinds of patterns of reasoning that I have just described [i.e. locating oneself by perceiving other objects in one's environment]. It is this capacity which enables us to make sense of the idea that we ourselves are elements in the objective order; and this is what is required for our thoughts about ourselves to conform to the Generality Constraint.¹⁹

¹⁷Evans 1982: 222.

¹⁸Cf. *Ibid.*

¹⁹Evans 1982: 223.

Once again the idea is that since we are persons, in order for us to have an adequate idea of ourselves, we must be able to think of ourselves as such. Hence, part of this ability consists in thinking of ourselves as located in space, like all other physical objects. In turn, the ability to think of ourselves as located in space is manifested in our forming—or being disposed to form—I-thoughts which are based on the perception of our environment.

However, Evans himself considers the case of deviant causal chains in which the subject receives perceptual information from someone else's body. Yet, as before, considerations having to do with deviant causal conditions do not show that in normal conditions, i.e. when perceptual information is received through one's own sensory organs, the ensuing bodily judgement would be rationally grounded on an inference containing an identification component such as [I am appeared to as if there were a tree], [The person who is appeared to as if there were a tree = the person who is in front of the tree], therefore [I am in front of a tree]. At most, what can be said is that the judgement [I am in front of a tree], when based on direct perception of the objects in one's environment, could be affected by EM if it were formed by subjects living in a logically possible world in which they are hooked up to someone else's body and are appraised of the fact that they may be receiving perceptual information through the sensory organs of those bodies. However, in our world one may coherently wonder whether one's perceptions are veridical, hence, whether it is true that one is in front of a tree, but one cannot coherently wonder whether one's perceptions are giving one information about *one's own* position relative to the objects in the environment.

Hence, as in the case of somatic proprioception, we can conclude by saying that Evans is right in holding that, at least in normal conditions, outer perception immediately gives one knowledge about one's own position relative to the objects in the environment. Hence, bodily self-ascriptions formed on these bases are identification-free and, therefore, IEM. Thus, if Evans is right in saying that one must be disposed to entertain this kind of judgements in order to have the first-person concept, then we can conclude that he is right in saying that our first-person concept is firmly anti-Cartesian and anti-Idealist.

2. Mental Self-Ascriptions and IEM: Why the Cartesian Cannot Get Off the Ground

According to Evans, the Cartesian conception of the self, both at the level of reference and of sense is motivated by the thought that *all* and *only* mental self-ascriptions can be IEM. Hence, in order to undermine it, he tries to show that the assumption on which the Cartesian can get off the ground, i.e. that all and only mental self-ascriptions are IEM, is not safe. As we have just seen, he first convincingly denies that only mental self-ascriptions are IEM, by showing that also some bodily ones are. Secondly, he shows that not all mental self-ascriptions are IEM. In order to make this latter point, Evans puts forward the following example, which is meant to show that mental self-ascriptions can be affected by error through misidentification:

[I]t feels as if I am touching a piece of cloth, and my relevant information is restricted to seeing, in a mirror, a large number of

hands reaching out and touching nothing, and one hand touching a piece of cloth. Here it makes sense for me to say 'Someone is feeling a piece of cloth, but is it I?'²⁰

This example should be understood as follows: it can be the case that a subject *feels as if* she is touching a piece of cloth. She also sees, reflected in a mirror, one hand touching a piece of cloth. She surmises that that is her hand. Hence she infers that *she is feeling* a piece of cloth. However, she could be mistaken about the identification component, i.e. she could be wrong about the fact that the hand she sees reflected in the mirror is hers. Therefore, her judgement [I am feeling a piece of cloth] could be wrong because of an error in the identification component.

This example shows that there can be *inferential mental self-ascriptions* which are affected by EM. Still, if this happens it is because, first, there is a sensation, e.g. as if of touching a piece of cloth, which grounds a self-ascription which would be identification-free and hence IEM, i.e. [I feel *as if* my hand is touching a piece of cloth]. Secondly, error occurs because there is a wrong identification component, which, however, does not contain any psychological property. That is to say, the subject misidentifies the hand she sees reflected in a mirror and takes it to be hers, thus entertaining the belief [The hand of the person who is touching a piece of cloth = my hand]. Holding this identity to be true, then the subject is entitled to holding [I am touching a piece of cloth], from which the *factive* psychological self-ascription [I feel a piece of cloth] would follow. Hence, we see that, in this case, the *factive* psychological self-ascription,

²⁰Evans 1982: 219-220.

though partly dependent on a sensation, is also dependent on an identification component. So, while awareness of a sensation would issue in an identification-free self-ascription which would be IEM, the presence of an identification component which is liable to EM makes the resulting factive psychological judgement liable to EM as well. Thus, in the case of EM of psychological self-ascriptions we should conclude that their being liable to EM is a *derivative* feature of their being based on an inference.

Another, maybe clearer, example of a mental self-ascription which would be affected by EM is this. X meets regularly with a group of people who all have some neurotic behaviour they are trying to cure. They talk about their experiences and a psychoanalyst moderates the discussion. At some point X hears the psychoanalyst say 'You are afraid of dogs'. However, the psychoanalyst, by saying 'you' does not mean to refer to X, but to Y. Still, X understands the psychoanalyst's utterance as directed to her. Moreover, since she firmly believes what her psychoanalyst tells her, she forms the belief [I am afraid of dogs]. However, her self-ascription of that mental property would be affected by EM, for it would be grounded on the following inference, containing a mistaken identification component: [I = The person to whom the psychoanalyst refers by using 'you'], [The person to whom the psychoanalyst refers by using 'you' is afraid of dogs], therefore [I am afraid of dogs]. Again, this example shows that *not all* mental self-ascriptions are IEM. But, once more, it also shows that the self-ascription would be EM because it is based on an inference containing an identification component. However, we still have to

explain IEM of non-inferential mental self-ascriptions, for that will allow us to confront the Cartesian and the Idealist directly.

But, instead of embarking on this task, Evans simply *acknowledges* that self-ascriptions of propositional attitudes, sensations and psychological properties are IEM. He writes:

Presumably it goes without saying that both the ways of gaining knowledge of ourselves that I have discussed in this section [i.e. knowing one's mental properties such as certain beliefs or sensations] give rise to judgements which are immune to error through misidentification. When the first component expresses knowledge gained in one of these ways, it does not make sense for the subject to utter 'Someone believes that *p*, but is it I who believe that *p*?', or 'It seems to someone that there's something red in front of him, but does it seem to me that there is something red in front of me?'²¹

Of course what Evans is saying here is right. For it would not make rational sense to wonder whether the sensations and beliefs of which one is introspectively aware could be someone else's. But why is it so? Moreover, why is it the case that the corresponding self-ascriptions are *logically* IEM, i.e. IEM in any possible world in which the same self-ascription is made on the basis of introspective awareness of a certain mental state? True, in a sense we know why these self-ascriptions are IEM, for we know that they are not rationally grounded on inferences involving any identification components. However, this does not explain yet why these self-ascriptions are *logically* IEM, i.e. they are such that they could *never* (logically) involve any identification component. In

²¹Evans 1982: 233.

order to see the point more clearly, contrast the case of non-inferential mental self-ascriptions with the case of bodily self-ascriptions based on somatic proprioception. As we have just seen, the fact that a bodily self-ascription is made on the basis of the deliverances of somatic proprioception does not exclude that in *deviant* causal conditions (in which a subject is appraised of the facts) the judgement be also based on a belief in an identification component, which, if wrong, would make it EM.²² But it seems reasonable to hold that any self-ascription of a mental property, which is made on the basis of introspective knowledge of it, *could never (logically)* involve as part of its grounds a belief in an identification component such as [I = the person who is feeling this pain/thinking this thought] and would thus be logically IEM. This, however, is a difference that is in need of explanation.²³

Evans' discussion of IEM of mental self-ascriptions stops here.²⁴ Instead of pursuing it further, he concentrates on an attack against the Cartesian

²²There shall be more about the distinction *de facto*/logical IEM in Ch. 6.

²³A hint at explanation can be found in the following passage. Evans 1982: 229: "The judgement's being a judgement with a certain content can be regarded as constituted by its being a response to that state". This is developed in Peacocke 1983. We shall discuss it in Ch. 7.

²⁴In fact Evans discusses memory-based I-thoughts as well. He claims that if, on the basis of the deliverances of memory, one remembers being *F*, then this does not leave open the possibility of EM. Although quasi-memories are possible, this just shows that in peculiar conditions one could have an apparent memory of an event which, in fact, "embodies information deriving from the perception of that event by a person who is not necessarily himself" (Evans 1982: 248). Once again, Evans is right in claiming that if it

conception of the mind. We will present this attack briefly, following the order in which it appears in *The Varieties of Reference*. The interest of this discussion is that Evans presents another pattern of argument which is meant to show that in order to be able to self-ascribe any mental property we must conceive of ourselves as persons, i.e. as embodied entities. If this argument is successful, it would help Evans show that both the Idealist and the Cartesian conception of the self are un-intelligible. For, if an understanding of any mental predicate depends on our ability to self-ascribe it and this, in turn, presupposes an ability to *think* of ourselves as persons, then it follows that we must necessarily think of ourselves as physical entities which exist in space and time and are perceivable by others. Let us now turn to these arguments.

3. *Against the Cartesian Conception of the Mind I. The Case of Beliefs*

Evans wants to dismantle the Cartesian conception of the mind as a theatre, or a garden, in which the subject, situated in an epistemic advantageous position, can infallibly *perceive* objects and events going on within it. Evans tries to do so with regard to beliefs and perceptions. In other words, the aim is to show that what we call introspection is not like seeing something happening within

seems to me to remember being *F*, then the self-ascription of the apparent memory is IEM. Moreover, he is also right in saying that, in the usual circumstances, the content of the apparent memory is also IEM, relative to the subject. For a discussion, cf. Shoemaker 1994b, Peacocke 1983: 174-179, Pryor 1998.

ourselves.²⁵ Moreover, he wants to show that knowledge of our mental states necessarily involves thinking of mental predicates as satisfiable by persons other than oneself, and necessarily involves thinking of oneself as just one possible argument for these functions on a par with other possible ones, i.e. other persons. Indeed, as we will see, it is only this last part of the argument which, if successful, would be effective against the Cartesian and the Idealist.²⁶

The anti-Cartesian move made by Evans in his discussion of self-attributions of beliefs consists in following Strawson on the point that an understanding of 'I believe that *p*' goes hand in hand with an understanding of 'x believes that *p*', where x can be instantiated by someone other than the subject herself. By so doing, one should gain a conception of mental predicates as ascribable on the basis of various kinds of evidence. However, Evans goes on to claim that these diverse kinds of evidence have a bearing upon the *truth* of first-person *self-*

²⁵A similar enterprise is carried out by Shoemaker, especially Shoemaker 1996. Cf. Ch. 6 in the following.

²⁶In order to get rid of the Cartesian conception of the mind and of introspection. Evans, to an extent, follows Wittgenstein's solution to Moore's Paradox—'I believe it rains but it doesn't' and 'It rains but I don't believe it'—which consists in denying that one can know whether one believes that *p*, when *p* is an empirical proposition, by looking *within* oneself and finding out whether there is such an *object*, i.e. a belief, amongst the ones one can distinctly perceive within one's mental theatre. Rather, by finding out that it rains, one is immediately able to form the belief that it rains and, if conceptually endowed, is immediately able to self-ascribe such a belief. Cf. Wittgenstein PI, II, pp. 190-192. On Moore's Paradox, cf. Heal 1994, Malcolm 1995, Picardi 1995, Shoemaker 1996, Essays 2, 4.

ascriptions.²⁷ But this cannot be right. For one finds out whether someone else has the belief that *p* by looking at what she does (her verbal and non-verbal behaviour) and the kind of evidence on which third-person ascriptions of beliefs are based is such as to manifest an embodied subject. Yet, by Evans' own lights, one is often in a position to self-ascribe a belief *immediately*, without observing oneself. That is why, following a generalisation of Hume's point we will consider shortly, one can be led to conclude that awareness of one's own mental states is not also awareness of an object which instantiates those properties. Hence, one can also grant that an understanding of the psychological predicate 'believing that *p*' depends on the ability to ascribe it both to oneself and to others. But from this it does not follow that in order to understand that predicate one must also *think* of oneself and others as on a par.

To put the point more vividly let us make use of an analogy: it may be true that in order to display an understanding of the predicate 'big', subjects must be able to attribute it to all sorts of things, e.g. objects, persons, numbers, etc. and, of course, they would do that on the basis of various kinds of evidence. But from this it does not follow that they should think of, say, numbers and physical objects as entities of the same sort, or on a par. Similarly, it does not follow that in order to understand psychological predicates, which are attributable on the basis of various sorts of grounds, subjects must think of themselves as on a par with other persons. Thus, this strand of the Generality Constraint is not sufficient in order to show that they must think of themselves as embodied.

²⁷Cf. Evans 1982: 226.

Let us consider the other strand of the Generality Constraint, in which it is not the subject of the psychological self-ascription that varies, but the predication. The idea is this: it could be suggested that, in order to have an adequate Idea of oneself, the subject should be able to entertain judgements like [I = the person who can be said by X to believe that p on the basis of her utterances and deeds]. By so doing, an understanding of the mental predicate 'to believe that p ' would go hand in hand with one's ability to think of oneself as a person on a par with all others, i.e. as a human being who can be perceived and interpreted by other human beings. Hence, an understanding of a mental predicate like 'believing that p ' would involve the ability to *think* of oneself as a person who can be supposed by others to believe that p . Yet, this does not mean that the self-ascription of the belief in the first place is based on an *awareness* of oneself as a person. That is to say, it is not the case that 'I believe that p ' when formed on the basis of one's direct awareness of one's own mental states is also based on an *awareness* of oneself as an object, i.e. as an embodied entity. Hence, although the ability to self-ascribe a belief can depend on one's ability to *think* of oneself as a person, the self-ascription of the belief not only fails to depend on "perceiving" one's own mental states, but it also fails to depend on "*perceiving*" oneself as having a body.²⁸

²⁸As we will see, this may induce the thought that the I who thinks, feels and perceives is not given as an object. That is to say, one's ability to self-ascribe the relevant mental states is not grounded in one's awareness of oneself as an embodied self. We shall discuss this issue at length in the following and in Ch. 7.

4. *Against The Cartesian Conception of the Mind II. The Case of Outer Perceptions, Hume's Point and the Case of Inner Perceptions*

Evans' treatment of outer perception has by now become familiar, given the recent intense debate on the content of perception. But here is a quick sketch which summarises some of the points made in one of the previous chapters: perceptions have non-conceptual content, which represents the world as being in a certain way and can either be true or false. Judgements are then based on the non-conceptual content of perceptions, although they require the mastery of the relevant concepts.²⁹

According to Evans, then, one can gain knowledge of one's own internal informational states by going through "exactly the same procedure as he would go through if he were trying to make a judgement about how it is at this place now, but excluding any knowledge he has *of an extraneous kind*"³⁰ and by prefixing the judgement based on one's outer perception with the locution 'It seems to me as though...'. By so doing, one would produce in oneself "a cognitive state whose content is *systematically* dependent upon the content of the informational state".³¹ However, contrary to the Cartesian model of knowledge of one's own mental states through introspection, there would not be anything like "perceiving that state".³²

²⁹Cf. Ch. 3, fn. 13.

³⁰Evans 1982: 227.

³¹Evans 1982: 228.

³²*Ibid.*

Evans' further move against the Cartesian conception of the mind consists in undermining the idea that subjects can have infallible knowledge of their perceptual states. He writes:

[C]onsider a case in which a subject sees ten points of light arranged in a circle, but reports that there are eleven points of light arranged in a circle, because he has made a mistake in counting, forgetting where he began. Such a mistake can clearly occur again when the subject re-uses the procedure in order to gain knowledge of his internal state: his report 'I seem to see eleven points of light arranged in a circle' is just wrong.³³

Finally, we have to take into account Evans' reading of Hume's point, which is the following:

[W]hen I enter most intimately into what I call *myself*, I always stumble on some particular perception or other (...). I can never catch *myself* at any time without a perception, and never can observe anything but the perception.³⁴

Now, Evans argues that, once we dispense with the perceptual metaphor, as indeed we should do, Hume's point can be read as establishing the following conclusion:

[W]hat we are aware of, when we know that we see a tree, is *nothing but a tree*. In fact, we only have to be aware of some state of the world in order to be in a position to make an assertion about ourselves.³⁵

³³Evans 1982: 228-229.

³⁴Hume 1978: 252. Cf. Evans 1982: 231.

³⁵Evans 1982: 231.

Then Evans goes on to say that by being aware of a state of affairs in the world we are *ipso facto* aware of a *substantial and persisting self*.³⁵ Now, the problem arises with his way of arguing to that conclusion. For here Evans does not avail himself of ecological theories of perception, according to which outer perception and, in particular, visual perception, involves also a perception of some of one's bodily parts and, therefore, contains a representation of "a substantial and persisting self".³⁶ Nor does he avail himself of the point he has already made that by perceiving a state of affairs, one gets knowledge of oneself as located, and hence as a physical entity. Rather, Evans tries to make his point by arguing, first, that by just perceiving a state of affairs, we can gain knowledge of our mental state. That is to say, by perceiving a tree we can have knowledge of ourselves as the perceivers of this state of affairs. Secondly, Evans argues that we can only attribute a perceptual predicate to ourselves, if we are able to think of the first-order property, i.e. *seeing a tree*, as instantiated by other persons, i.e. other persisting selves, in conformity with one of the strands of the Generality Constraint. However, as before (cf. § 3), this cannot be carried on to establish that by so doing we must *think* of ourselves as persons. Nor can it be carried on to establish that we are *aware* of a substantial

³⁵ Notice in fact that the quotation from Evans above goes on as follows "Now this might raise the following perplexity. How can it be that we can have knowledge of a state of affairs which involves a substantial and persisting self, simply by being aware of (...) a state of the world?" (Evans 1982: 231).

³⁶Ecological theories of perception have been developed starting from Gibson's work. Cf. Gibson 1979.

and persisting self. For one thing, the kind of evidence on which these predicates are ascribed is *different* in the case of the first and the third person. In the case of the former, by *perceiving* a state of affairs, the subject is *ipso facto* able to attribute the predicate to herself; whereas in the latter case it is only on the basis of either observation of a person or testimony about a person that the subject can think of someone else as perceiving a state of affairs. Consequently, by Evans' own lights, it is only when these predicates are attributed to a third person that the bases on which this can be done *manifest a subject* with both physical and psychological properties. But, in our own case, that kind of knowledge is not available exactly because no observation of someone perceiving a tree is going on.³⁷ So, once again, this strand of the Generality Constraint, i.e. the strand according to which in order to be able to self-ascribe a mental property a subject must be able to ascribe it to others, is too weak to establish the intended points, i.e. that by attributing any mental property whatever to oneself one is *ipso facto aware* of oneself as a bodily entity, and that one can only attribute a mental property to oneself if one *thinks* of oneself as embodied.

Moreover, the second strand of the Generality Constraint, i.e. the one in which it is the predication that varies, would have similar problems to the ones we saw at the end of § 3. For, although an understanding of 'perceiving that *p*' may go hand in hand with an understanding of what it would be for one to be supposed by someone else to perceive that *p*, this does not show yet that self-ascriptions of perceptions are based on an *awareness* of oneself as an

³⁷For a discussion of this point cf. Ch. 6 and, in particular, Ch. 7.

embodied entity. Hence, according to Evans, self-ascriptions of outer perceptions are neither based on anything like “perceiving” one’s own mental state, nor are they based on anything like “perceiving” oneself as an object (unless we subscribe to ecological theories of perception, but Evans does not; or else we insist on the point already made by Evans that by perceiving a state of affairs we gain knowledge of ourselves as located in space and, hence, as physical entities). Still, they may involve *thinking* of oneself as a person and could be based on a mental state which represents the subject as located in space relative to other objects.

Considerations partly similar to the ones just exploited for outer perceptions apply to Evans’ treatment of reports of sensations as based on informational states of the subject’s body. He claims that by having pain in one’s leg (say) and so judging a subject thereby gains knowledge of being in an informational state which represents a part of her body as feeling awful. Hence, she would have knowledge of herself as embodied.³⁹ Thus, bodily sensations give one knowledge of oneself as embodied. Hence, although bodily sensations are not based on anything like observing oneself, they give one information about one’s own body. Therefore, contrary to the Cartesian and the Idealist claims it is not true that awareness of mental properties is never an awareness of oneself as embodied.

Still, even if bodily sensations give one knowledge of oneself as embodied, this is not yet to say that *all* sensations would be based on anything like

³⁹Evans 1982: 231.

“perceiving” oneself as embodied. In fact, it is reasonable to maintain that, for instance, sensing blue would not give one awareness of oneself as embodied.

Hence, we can conclude this section on Evans’ second pattern of arguments against the Cartesian and the Idealist by saying that he has not managed to show that *all* cases of psychological self-ascriptions are based on an *awareness* of oneself as an embodied entity.⁴⁰ For he has convincingly argued to that effect only in the case of self-ascriptions of bodily sensation. Still, he has shown that the ability to self-ascribe all sorts of mental properties requires the subject’s ability to *think* of herself as a person who can be perceived and made sense of by others. Hence, he has shown that our concept of the first-person is at odds with the Cartesian and the Idealist uses of it. Yet, he has not completely shown that their views cannot have a grip at all, because, as we have seen, he has not fully shown that self-ascriptions of mental properties are based on an *awareness* of oneself as an embodied entity. Hence, he has not managed to show how awareness of a mental state which does not manifest an embodied owner can ground a *self*-ascription of it, i.e. an ascription of that mental state to a human being.

⁴⁰ We will come back to this issue in connection with Peacocke’s discussion of representation-dependent and representation-independent uses of the first person. Cf. Ch. 7.

5. Conclusions

Evans is right in saying that bodily self-ascriptions based on either proprioception or perception of one's environment are IEM. Hence, if he is right in saying that one must retain the capacity, exercised in the past, to make these self-ascriptions in order for one to have the first-person concept, then he has managed to show that our first-person concept is firmly anti-Cartesian and anti-Idealist. Hence, his account of IEM offers the prospects of a sound substantive response to the trilemma.

However, an exhaustive explanation of IEM requires: (a) a satisfactory characterisation of the phenomenon at hand; (b) a determination of the class of (in particular) I-judgements which are IEM; finally, (c) an explanation of why these judgements are IEM. Now, Evans has accomplished (a) relative to a particular kind of IEM judgements, i.e. those which, *de facto*, do not involve an identification (as part of the subject's grounds for the self-ascription) and are based on a way of gaining information directly from objects (one's body included).⁴¹ He has also accomplished (b), since he has shown that, beside non-inferential mental self-ascriptions, there are some bodily self-ascriptions which, when based on appropriate grounds, are IEM and he has also shown that not all mental self-ascriptions are IEM. However, he has accomplished (c) only in part. For he has shown that bodily self-ascriptions can be IEM when based on a way of *gaining information directly from one's body or from the*

⁴¹But recall our amendments to his definition of identification-freedom and, consequently, of IEM (cf. Ch. 3 and § 2 above).

environment. However, he has not shown how *non-inferential mental self-ascriptions* can be IEM. Nor has he shown why introspection-based mental self-ascriptions are *logically* IEM, i.e. they are such that they could never be based on an identification component such as [I = the person who is enjoying *this* mental state], which, if wrong, could lead to EM. Hence, in the following we will try to supplement Evans' substantive response to the trilemma with a satisfactory account of IEM of these kinds of mental self-ascriptions, which will also show why they are logically IEM, i.e. it can never be the case that the rational grounds for them involve a belief in an identification component.

SECTION III

REVISIONIST ACCOUNTS OF IMMUNITY TO ERROR THROUGH MISIDENTIFICATION AND DEFLATIONIST RESPONSES TO THE TRILEMMA

Summary

In this section we consider the proposals of various authors in which *revisionist accounts of IEM* and *deflationist responses to the trilemma* are put forward.

In Ch. 6

- we consider Shoemaker's account of EM and his distinctions between *absolute vs. circumstantial* and *de facto vs. logical IEM* (§ 1);
- we claim that his account of IEM of non-inferential mental self-ascriptions, which is fundamental to his response to the trilemma, is partially wanting (§ 1.1);
- we turn to Wright's account of EM and to his revisionist account of IEM (§ 2);
- we then turn to Wright's account of IEM of non-inferential mental self-ascriptions, which is fundamental to his response to the trilemma, and claim that it is incomplete (§ 2.1);
- we turn to Pryor's distinction between *de re* and *which-EM* and to his corresponding distinction between *de re IEM* and *which-IEM*. We argue that since there is just one kind of EM, namely *de re EM*, there is just one corresponding kind of IEM, i.e. *de re IEM*. Hence, we conclude that his

revisionist account of IEM is partially wanting and no real response to the trilemma has been given.

In Ch. 7

- we discuss Peacocke's theory of I-thoughts and propose some amendments to it (§ 1). Moreover, we claim that it can be combined with Evans';
- we maintain that it explains IEM of introspection-based mental self-ascriptions only partially (§ 2), for it rests on the assumption, which is in need of explanation, that one cannot be introspectively aware of someone else's mental states;
- we then turn to Peacocke's more recent work on the epistemology of first-person thoughts (§ 3);
- we present his account of (RI) "representation-independence" which is typical to some I-thoughts in his view (§ 3.1);
- we turn to Peacocke's delta-account of the epistemology of RI I-thoughts (§ 3.2). We claim that it presupposes an explanation of why mental states are necessarily owned and owned by the same person who is introspectively aware of them. Drawing on Peacocke's own proposal we supplement it in order to provide an explanation of the former fact; we leave the explanation of the latter for the following section;
- we summarise our conclusions (§ 4).

Chapter Six

Revisionist Accounts of IEM and Deflationist Responses to the Trilemma: Shoemaker, Wright and Pryor

The conclusion we drew in the last section was that if Evans' theory of I-thoughts is correct, then the Cartesian and the Idealist positions are not only problematic, for the reasons we saw in Ch. 2, but also because their use of the first person is unintelligible from our stand point. However, although their views about the self are obviously contentious, they provide at least a *prima facie* explanation of why *non-inferential mental self-ascriptions are IEM*, according to IEM_{diff}. To repeat, the Cartesian will claim that they are IEM because of the peculiar nature of the self and of the conditions of one's encounter with it, i.e. the self is a mental substance which is presented to oneself for each act of thought and which cannot be mistakenly identified given the transparency of the mind. By contrast, the Idealist will claim that since no object at all is given for identification or recognition when the relevant mental self-ascriptions are made, it follows that there cannot be error through misidentification either.

In this chapter we will examine various *revisionist* accounts of IEM, i.e. accounts of IEM which are revisionist of IEM_{diff}, yet offer *deflationist* responses to the trilemma. That is to say, they do not try to confront the Cartesian and the

Idealist on conceptual grounds. Rather, they claim that there can be an explanation of IEM in general and of non-inferential mental self-ascriptions in particular, which is compatible with the fact that the self is embodied. Hence, they undermine the motivation to go either Cartesian or Idealist. In the following, we will assess whether they really do manage to give a satisfactory explanation of IEM of non-inferential mental self-ascriptions. Recall, in fact, that a satisfactory explanation of IEM of non-inferential mental self-ascriptions is the missing piece of an exhaustive substantive response to the trilemma. For, as we have seen, Evans does not really provide us with an explanation of IEM of this kind of self-ascriptions. Hence, if the deflationists offer us a sound explanation of this phenomenon, their views will help us supplement Evans' own proposal. However, as we will see, deflationists illuminate the phenomenon we want to explain only in part. Hence, we will have to integrate their views in important respects. This, however, shall be the topic of the last section of our work.

1. Shoemaker: *Absolute vs. Circumstantial IEM, De Facto vs. Logical IEM*

Recall Shoemaker's definition of IEM we introduced in Ch. 1:

To say that a statement 'a is ϕ ' 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 ϕ but makes the mistake of asserting 'a is ϕ ' because, and only because, he

mistakenly thinks that the thing he knows to be ϕ is what "a" refers to.¹

Obviously Shoemaker's definition should be amended in such a way that not only sentences but also *judgements* can be EM, and that we can *falsely* believe *a* to be ϕ , i.e. EM is compatible with an error in the predication component.² However, although Shoemaker offers a definition of EM which is essentially correct, he does not present a corresponding definition of IEM. Rather, he distinguishes between various *kinds* of IEM.

The first kind of distinction is between I-judgements which are *absolutely* IEM and those which are only *circumstantially* so. As the labels suggest, this distinction has to do with the *circumstances* in which the I-judgement is made. We can be more precise and say that the distinction has to do with the *grounds* on which the judgement is made.³ That is to say, the same content can be entertained on the basis of different grounds. For instance, I can judge [My hair is blowing in the wind] on the basis of the observation of my reflection on a glass window or on the basis of the deliverances of somatic proprioception. In the former is the case the judgement can be liable to EM, whereas, in the latter case, it is IEM.

However, according to Shoemaker, there are self-ascriptions which are absolutely IEM. In this connection, he seems to have in mind mental self-

¹Shoemaker 1994a: 82.

²A similar amendment to Shoemaker's definition can be found in Peacocke 1999: 269, although he opts for beliefs and thought constituents or modes of presentation.

³Cf. Evans 1982, Wright 1998, Garrett 1995, Peacocke 1999.

ascriptions. Yet, as we saw with Evans, also mental self-ascriptions can be liable to EM, when they are based on inference. Hence, also mental self-ascriptions can be circumstantially IEM, depending on which grounds they are based. The only case in which it seems to make sense to say that the self-ascription of a mental property is absolutely IEM seems to be the case of self-ascriptions of sensations. For it would be quite difficult to imagine that one may know that one is in pain or is appeared to redly on inferential bases. However, it has to be stressed that, if at all, these would be the only examples of absolute IEM.

Hence, Shoemaker's distinction between circumstantial and absolute IEM is potentially misleading, for it suggests that there are self-ascriptions, i.e. mental ones in particular, which are IEM no matter which grounds they are based on. In fact, it would be more correct to say that IEM is always a function of the grounds on which a self-ascription is based, and then notice that there are some mental self-ascriptions, like self-ascriptions of sensations, which seem to be (at least almost always) based on grounds that make them IEM.

Moreover, according to Shoemaker, judgements that are circumstantially IEM are so because they are ultimately based on a judgement that is *absolutely* IEM. Shoemaker illustrates the distinction between circumstantial and absolute IEM and this dependence-thesis as follows:

First-person statements that are immune to error through misidentification (...) [are] those in which "I" is used 'as subject', [and] could be said to have 'absolute immunity' to error through misidentification. A statement like 'I am facing a table' does not have this sort of immunity, for we can imagine circumstances in which someone might make this statement on the basis of having

misidentified someone else (e.g. the person he sees in a mirror) as himself. But there will be no possibility of such a misidentification if one makes this statement on the basis of seeing a table in front of one in the ordinary way (without aid of mirrors, etc.); let us say that when made in this way the statement has 'circumstantial immunity' to error through misidentification relative to "I". It would appear that, when a self-ascription is circumstantially immune to error through misidentification, this is always because the speaker knows or believes it to be true as a consequence of some other self-ascription, which the speaker knows or is entitled to believe, that is absolutely immune to error through misidentification; e.g. in the circumstances just imagined the proposition 'I am facing a table' would be known or believed as a consequence of the proposition 'I see a table in the centre of my field of vision'.⁴

Shoemaker rightly points out that certain physical self-ascriptions like [I am facing a table], when based on the direct perception of one's environment, are IEM. By contrast, if the same self-ascription is based on observation of the reflection in a mirror, it will be liable to EM. However, he thinks that bodily self-ascriptions that are circumstantially IEM are based on a self-ascription that is absolutely IEM. As the example he discusses suggests, Shoemaker wants to make the point that the relevant bodily self-ascription would be based on a mental one, which, depending on whether or not we allow for absolute IEM in the case of self-ascriptions of sensations, will be either absolutely IEM or, at least, circumstantially so. That is to say, since, in those circumstances, it would be based on introspective awareness of one's sensations, it would be IEM. Thus, in the case of bodily self-ascriptions, their being IEM depends, according

⁴Shoemaker 1994a: 82.

to Shoemaker, on the fact that they are based on a non-inferential, and hence, IEM, mental self-ascription.

Finally, it is important to be clear about what is meant by saying that a certain judgement is IEM “as a consequence of some other self-ascription (...) that is absolutely immune to error through misidentification”. In particular, we have claimed that one need not actually entertain the relevant self-ascription which is IEM. Rather, it is sufficient for it to be (not necessarily currently) believed, where such a belief would be part of the subject’s grounds for the physical self-ascription.

Moreover, Shoemaker distinguishes between *de facto* and *logical* IEM. He illustrates that distinction as follows:

[In the case of] ‘I was angry’, [said on the basis of memory in the ordinary way], a mistake of identification is impossible. It goes with this that the (...) past-tense, first-person judgement does not rest on an observationally based reidentification of the person referred to with ‘I’ (...). [But] one might ‘quasi-remember’ past experiences or actions that are not one’s own (...). To allow that this is [logically] possible is to allow that in a certain sense first-person memory judgements are subject to error through misidentification.⁵

A definition of quasi-memory can be found in Evans who writes:

[A] subject *q*-remembers an event *e* if and only if (i) he has an apparent memory of such an event, and (ii) that apparent memory in fact embodies information deriving from the perception of that event by a person who is not necessarily himself.⁶

⁵Shoemaker 1994b: 130-131.

⁶Evans 1982: 247-248.

Hence, according to Shoemaker, while non-inferential mental self-ascriptions are logically IEM, i.e. they are IEM *in any logically possible world*, bodily self-ascriptions based on proprioceptive grounds or on perception of one's environment as well as memory-based self-ascriptions are only *de facto* so. That is to say, according to him, in this world bodily self-ascriptions based on these kinds of grounds are based on a non-inferential mental self-ascription, such as [I feel as if my legs are bent] or [I seem to remember that *p*], but do not involve an identification component like [The body I am feeling now = my body] or [I = the person whose past I remember]. However, in a logically possible world, according to Shoemaker, the relevant self-ascriptions could involve such an identification component, which may or may not be right. Hence, these self-ascriptions are IEM only because, in this world, certain contingencies obtain, i.e. we receive proprioceptive information from our bodies and remember our own experiences.

However, as we saw with Evans, if it is maintained that EM and IEM have always to do with the subject's rational *grounds* for her judgement, bodily self-ascriptions based on somatic proprioception and self-ascriptions based on q-memories would be liable to EM just in case in deviant possible worlds the subject *knew* or *had reasons to believe* that she might be receiving information from someone else's body or might be q-remembering another person's past. For, if she did not know that, then she would have no reason to ground her self-ascriptions on a belief in an identification component.

To recapitulate: we have seen that Shoemaker offers an explicit definition of EM which, once some amendments to it have been made, is essentially correct.

By contrast, he does not offer an explicit definition of IEM. Rather, he distinguishes between absolute and circumstantial IEM and between logical and *de facto* IEM. Both these distinctions need to be amended, though. For the distinction between circumstantial and absolute IEM can rarely (if ever) make sense, although it has the merit to bring out the fact that EM/IEM is a function of the *grounds* on which a judgement is based. Moreover, the distinction between logical and *de facto* IEM has to be relativised to the subject's rational grounds for the judgement (or to her state of information) in a logically possible world. However, Shoemaker rightly recognises that IEM is compatible with the view that selves are embodied, contrary to the Cartesian and the Idealist claims.

Finally, Shoemaker also points out that circumstantial and *de facto* IEM depend, in a sense which we have specified, on some self-ascription that is logically IEM. In particular, these kinds of IEM depend on some non-inferential mental self-ascription. However, precisely because, according to Shoemaker, IEM of bodily self-ascriptions is a consequence of their being based on psychological non-inferential self-ascriptions, which are logically IEM, the latter are what must still be explained. Without such an explanation, Shoemaker's account of IEM does not represent a fully-fledged alternative to either the Cartesian's or the Idealist's.

1.1 *Shoemaker's Account of IEM of Non-Inferential Mental Self-Ascriptions*

Shoemaker himself points out that a satisfactory explanation of IEM of self-ascriptions depends on the viability of a satisfactory explanation of IEM of non-

inferential mental self-ascriptions. For, according to him, they are what must be explained in order to explain any other form of IEM, when self-ascriptions are at stake. For this reason, Shoemaker embarks on a discussion of *introspection*. To anticipate, on the one hand, he wants to claim that Hume is right in pointing out that in introspection one does not encounter a peculiar object like a mental substance which is identical to oneself, and is a bearer of mental properties. On the other hand, he wants to defend the claim that in being introspectively aware of one's own mental properties, one is *ipso facto aware* that they are instantiated in oneself. Consequently, although one is not quasi-perceptually aware of an object, when one is introspectively aware of one's own mental properties, one has nevertheless the right to *self-attribute* them. That is to say, one has the right to consider them as instantiated in a subject, which is identical to oneself. Let us consider Shoemaker's claims in more detail.

First of all, he points out that in introspection the subject is not aware of herself as a physical object.⁷ We saw how Evans tried to deny this point by appealing, first, to the fact that introspection should be conceived as involving also somatic proprioception, in which case the subject would be aware of herself as a physical entity. Secondly, Evans claimed that awareness of one's bodily sensations gives the subject awareness of herself as embodied. Hence, it gives her awareness of herself as a physical object. Here is how Shoemaker's reply to Evans' claims would go:

(...) [I]f my 'self' is a flesh-and-blood person, why shouldn't it be accessible to me (itself) in a way in which it is not accessible to

⁷Cf. Shoemaker 1994a: 85.

others, so that in knowing that what is presented to me is presented in this special way—from the inside, as it were—I would know that it can be nothing other than myself?

Now there is a perfectly good sense in which my self is accessible to me in a way in which it is not to others (...). I see nothing wrong with describing the self-ascription of such predicates as manifestations of self-knowledge or self-awareness. But it is plainly not the occurrence of self-awareness in *this* sense that has been denied by those philosophers who have denied that one is an object to oneself; e.g. it is not what Hume denied when he said: 'I can never catch *myself* at any time without a perception, and never can observe anything but the perception.' What those philosophers have wanted to deny, and rightly so, is that this self-awareness is to be explained in a certain way. They have wanted to deny that there is an experiencing or perceiving of one's self that explains one's awareness that one is, for example, in pain in a way analogous to that in which one's sense perception of John explains one's knowledge that John has a beard.⁸

Hence, the reply would consist in pointing out that at least awareness of one's own mental properties does not give one awareness of oneself as a physical object. This, in turn, can be glossed as the view that when one self-ascribes a mental property in a non-inferential way, the self-ascription is neither based on the identification of an object, nor on its individuation. For no self as an object is presented to one when one self-ascribes a mental property.

However, Shoemaker's claim is too strong. For, arguably, when one is aware of one's bodily sensations one is also aware of oneself as a physical entity with, for instance, an aching wrist. Surely he is right in saying that this kind of awareness is not based on an *observation* of oneself, which can disclose to one

⁸Shoemaker 1994a: 87.

one one's own properties, like having an aching wrist. Nevertheless, it seems wrong to deny that one would be presented to oneself as a physical entity and, therefore, as an object. Still, Shoemaker's claim is valid when it comes at least to self-ascriptions of beliefs and non-bodily sensations. For, when one judges [I believe that p] or [I am appeared to redly] one does not do so by being presented with oneself as embodied and by finding out that that self has the property of believing that p /sensing red, etc. Rather, one has a belief that p or a sensation of red and one is *ipso facto* entitled to self-ascribe it, without being presented to oneself in thought.

Secondly, Shoemaker proposes a *reductio* of the view that in introspection one is aware of an object, whether it be empirical or peculiar like a Cartesian ego. He writes:

An essential part of the explanation of my perceptual awareness that John has a beard is the fact that the observed properties of the man I perceive, together with other things I know, are sufficient to identify him for me as John. If the awareness that I am in pain had an explanation analogous to this, it would have to be that I 'perceive', by 'inner sense', something whose 'observed properties' identify it to myself. And if the supposition that the perception is by 'inner sense' is supposed to preclude the possibility of misidentification, presumably this must be because it guarantees that the perceived self would have a property, namely, the property of being an object of *my* inner sense, which no self other than myself could (logically) have and by which I could infallibly identify it as myself. But of course, 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. Yet, if it were possible in this one case for my self-knowledge not to be grounded on an identification of a self as myself, there seems to be no reason at all why this should not be possible in other cases, e.g. in

the case of my knowledge that I feel pain or my knowledge that I see a canary. Thus the supposition that there is observation by inner sense of oneself—where this is something that is supposed to explain, and therefore cannot be simply equated with, the ability to self-ascribe those predicates whose self-ascription is immune to error through misidentification—is at best a superfluous hypothesis: it explains nothing that cannot be just as easily, and more economically, explained without it.⁹

The point made by Shoemaker can be put as follows: if knowledge that I am in pain depended on some form of perceptual awareness of my self, it would also involve an identification of that object as my self. Hence, I should be able to entertain identity judgements of the form [I = this self]. However, knowledge of the first component, i.e. knowledge relative to [I], cannot be grounded on some previous identification, on pain of a vicious infinite regress. Therefore, one must already have knowledge of one's self which is not based on identification. Hence, there is no reason to suppose that knowledge that, for instance, I am in pain cannot be independent of any identification of a self as myself. Thus, appeal to the object-perceptual model, in the case of introspection, is, at best, superfluous.¹⁰

Although Shoemaker's argument is correct, it is not conclusive because it overlooks the following possibility. Although no *recognition* or *identification* of one's self is involved when one is aware of feeling pain, it could be the case that an object be presented and, hence, *individuated*. Indeed, this would square

⁹Shoemaker 1994a: 87-88.

¹⁰For a discussion of the object-perceptual model of introspection, cf. also Shoemaker 1996, Ch. 10.

with the intuition that when one self-ascribes a bodily sensation, one is also presented to oneself as an embodied entity. However, this possibility would not square with the phenomenology of self-ascriptions of beliefs and non-bodily sensations, for, in those cases, the embodied self does not seem to be presented in introspection.¹¹

Let us now turn to Shoemaker's positive suggestion. To repeat, he wants to show that, although we are not aware of an object when we are aware of our mental properties, we can, nevertheless self-attribute them, i.e. attribute them to a self, which is necessarily one's self. Here it is:

I think that the main source of trouble here is a tendency to think of awareness as a kind of perception, i.e. to think of it on the model of sense-perception. I have been denying that self-awareness involves any sort of perception of oneself, but this should not be taken to mean that in making a judgement like 'I feel pain' one is *aware* (italics mine) of anything less than the fact that one does, oneself, feel pain; in being aware that one feels pain one is, tautologically, aware, not simply that the attribute *feel(s) pain* is instantiated, but that it is instantiated *in oneself*.¹²

The idea is this: introspection should not be construed on the basis of the object-perceptual model. However, awareness of one's mental properties is *ipso facto* awareness of those properties as instantiated in oneself.

¹¹Cassam 1997b makes the same point. Consequently, a materialist theory of the self, in his view, should show, *contra* Hume and Kant, that, when subjects self-attribute all sorts of mental properties on the basis of introspection, they are also introspectively aware of a material self, i.e. of an embodied self.

¹²Shoemaker 1994a: 89.

The trouble with this claim is that it is made in terms of *awareness*. And awareness seems to imply the presentation of something. Hence, saying that in being aware of a mental property one is “aware (...) that it is instantiated *in oneself*” is in danger of suggesting that, in introspection, one is, after all, aware of one’s self as an object. However, since awareness seems to involve the presentation of an object, this interpretation would be in danger of slipping into the object-perceptual model again. Alternatively, Shoemaker’s point could be understood as meaning that one is not aware of one’s self as an object (empirical or peculiar like a Cartesian ego, it does not matter). Indeed this seems to be the right interpretation, given Shoemaker’s purposes. However, it is totally unclear how simple awareness of one’s mental properties is *ipso facto awareness* that they are instantiated in oneself, when no self as an object is given or presented in introspection. That is to say, the problem is to explain how we can justify, once Hume’s point is taken on board, the passage from the awareness of, for instance, there being a tree, which would lead to the judgement [There is a tree] to the self-ascription of it, i.e. the judgement [I believe that there is a tree]. In other words, once Hume’s point is endorsed, the problem arises of explaining how it is possible to justify the ascription of a mental property to a self, i.e. one’s self, thereby avoiding the Idealist position.

At this stage, however, Shoemaker also mentions the fact that it is *tautological* that if one is aware of a mental property, it is *ipso facto* aware that it is instantiated in oneself. Equivalently, he claims that there are certain predicates “each of which can be known to be instantiated in such a way that knowing it to be instantiated in that way is equivalent to knowing it to be

instantiated in oneself".¹³ But the question still is: how should we develop this suggestion in such a way that it is no longer mysterious why the relevant self-ascriptions are logically IEM? That is to say, what makes it the case that, contrary to other predicates, some mental ones are such that to know in a certain way that they are instantiated does not (logically) leave open the question of who instantiates them, i.e. oneself?¹⁴

Finally, Shoemaker mentions the fact that from Hume's point, which is phenomenological, and consists in claiming that in introspection we are not aware of a self who has the relevant mental properties, it would not follow that there is no self at all.¹⁵ For the former is a correct *epistemological* claim, which, however, would not license the latter *metaphysical* conclusion. This observation seems a sound one, but, in order to carry some weight, it must be placed in a wider context. We will do so in Ch. 7 and 8.

To conclude, Shoemaker's way of phrasing his positive suggestion (i.e. there is no self presented in introspection when one self-ascribes a mental property and yet one has the right to ascribe it to oneself) is, at times, infelicitous and, in any case, partially incomplete. However, it is an interesting and important one, on which we will elaborate at length in the following two chapters. Hence, we can conclude by saying that Shoemaker has provided us with a good understanding of EM, useful distinctions between various kinds of IEM (once some amendments to them have been made), good arguments in favour of the

¹³Shoemaker 1994a: 90.

¹⁴The fully-fledged answer to this question will be given in Ch. 8.

¹⁵Shoemaker 1994a: 85-86.

“elusiveness of the self” in introspection (at least in some cases),¹⁶ but with an insufficient explanation of IEM of non-inferential mental self-ascriptions, which could seem highly favourable to the Idealist position. Therefore, he has failed to provide us with a satisfactory deflationist account of IEM of mental self-ascriptions which is capable of dissolving the trilemma about the self.

2. A Revisionist Conception of IEM: Wright

Recall Wittgenstein's passage in the *Blue Book* we quoted at the beginning of Ch. 1. There it looked as if he were drawing the line between self-ascriptions which are liable to EM and those which are IEM on the basis of their *subject-matter*, i.e. their being either mental self-ascriptions or bodily ones. By contrast, Crispin Wright, in his discussion of Wittgenstein's passage, follows Evans and Shoemaker in pointing out that this need not be the case. Although, on the one hand, it can be a matter of exegetical dispute, whether Wittgenstein held such a naive view, on the other hand, no matter whether he held it or not, this would indeed be wrong—so it is argued.¹⁷ For, according to Wright, it is not so much the subject-matter of a sentence (or a judgement) that decides whether the use of 'I' in it is immune to error through misidentification, rather the *grounds* alone on which a subject attributes to herself all sorts of predicates can decide the issue. For, as we have already seen, the same judgement, e.g. [My hair is

¹⁶This expression is due to Cassam 1994: 1-18 and Cassam 1997b: 1-6.

¹⁷Cf. Wright 1998: 19, but also Evans 1982, Ch. 7 and Garrett 1995: 351.

blowing in the wind], can either be based on the *observation* of oneself in a glass window, or on somatic proprioceptive *feelings*. If the former is the case, then one could mistake someone else's hair for one's own, in which case the use of the first-person pronoun would be affected by error through misidentification. By contrast, according to the latter option, there is no room (at least *de facto*) for the self-attribution to be affected by error through misidentification. Hence, Wright is making the point that IEM has to do with the circumstances, and, more specifically, the *grounds*, on which the judgement is formed.

In particular, however, Wright maintains that uses of 'I' as subject, i.e. uses of 'I' that are IEM, are those for which, if one's grounds for the self-attribution were defeated, they could not survive as grounds for the corresponding existential generalisation. By contrast, the use of 'I' as object would be such that, if one's grounds for the self-attribution were defeated, the existential generalisation would nevertheless continue to be supported by the grounds which justified the judgement.¹⁸

Let us illustrate Wright's claim with reference to our previous example: if the judgement [My hair is blowing in the wind] is based on seeing a reflection in a shop window, and the observation is defeated, maybe because someone else is reflected in the shop window, then the observation of the reflection would nevertheless survive as a ground for the existential generalisation [Someone's hair is blowing in the wind]; whereas this would not be the case if the judgement were based on proprioceptive feelings. That is to say, if I am a victim of a

hallucination and hence it is not true that my hair is blowing in the wind, then my proprioceptive feelings could not survive as grounds for the existential generalisation [Someone's hair is blowing in the wind]. Nor would those feelings survive as grounds for a third-personal ascription if, in deviant causal conditions in which the subject is receiving proprioceptive information from someone else's body, she did not know that. For she would not have any reasons to consider those feelings as grounds for a third-personal ascription.¹⁹

Therefore, the conclusions drawn by Wright can be summarised as follows: (1) self-attributions which are IEM cut across the physical and the psychological domains, since they depend on the kind of grounds on which they are made. Such grounds, in turn, (2) can allow or fail to allow for the existential generalisation once they have been defeated. Consequently, the ensuing judgements will be EM or IEM respectively.

Hence, we have seen that Wright gives an elegant account of the fact that IEM is a function of the grounds on which the judgement is based. In the case of bodily self-ascriptions, however, he thinks that their being IEM is only a

¹⁸Cf. Wright 1998: 19.

¹⁹Wright's account of EM and IEM is compatible with ours. For, if the judgement is based on a belief in an identification and in a predication (or individuation) component, the former can be defeated by undercutting evidence and the latter survive. Hence, the existential generalisation would follow from the predication or individuation component. By contrast, if the judgement is not based on any identification component, then if the subject's grounds are defeated, then they cannot remain as grounds for an existential generalisation.

derivative feature of their being based on a possible mental self-ascription. Here is what Wright writes:

[T]he *source* of IEM in non-psychological first-personal claims is always their being based on avowable psychological matters (...). The IEM of non-psychological self-ascriptions, when they have it, is presumably to be viewed as an inheritance from their basis in an underlying possible avowal.²⁰

To recapitulate: Wright explicitly recognises the fact that IEM is not a function of the subject-matter of a sentence or a judgement. In particular, he points out that some bodily self-ascriptions are IEM when based on somatic proprioceptive feelings. In turn, these feelings could give rise to a non-inferential mental self-ascription. However, precisely because of Wright's account of IEM of bodily self-ascriptions, the quest for an explanation of IEM of non-inferential mental self-ascriptions is even more pressing. For either IEM of non-inferential mental self-ascriptions is what must be explained in the first place, or it is what must be explained anyway in order to provide an explanation of IEM of bodily self-ascriptions. Hence, the Cartesian and the Idealist are not defeated yet, because, at least *prima facie*, they do provide us with such an explanation. Let us now turn to Wright's account of IEM of non-inferential mental self-ascriptions.

²⁰Wright 1998: 20.

2.1 Wright's Explanation of IEM of Non-Inferential Mental Self-Ascriptions and His Response to the Trilemma

Wright's observations about EM/IEM are placed in the wider context of his discussion of Wittgenstein's philosophy of mind. In particular, he is interested in various kinds of avowals. Thus, he writes:

It is not the fact that IEM is not the exclusive property of psychological claims which entitles us to bracket the phenomenon in the present context, but rather the reflection that avowal's exhibition of it is a *derived* feature, as it were—effectively a consequence of respects in which we have already noted the distinction from third-personal psychological claims. Specifically it is a consequence of their being groundless while the corresponding third-personal claims demand evidential support. For if an avowal, 'I am ϕ ', did not exhibit IEM, then its defeat would be consistent with the subject's retention of an entitlement to the corresponding existential generalization—'Someone (else) is ϕ '—which could then be asserted *groundlessly*. But to suppose that such a claim could be both admissible and groundless would clash with the original asymmetry.²¹

The basic thought is clear enough: IEM of mental self-ascriptions is a derivative feature of the fact that these are avowals, i.e. they are neither based on inference nor on observation of oneself. For, if avowals were not IEM, then, given Wright's definition of IEM, their defeat would sustain a third-personal psychological claim. This, in turn, would be groundless, i.e. neither based on inference nor on observation. However, all third-personal psychological claims

²¹Wright 1998: 20.

are either based on inference or observation. Consequently, avowals must be IEM on pain of contradiction.

Even so, however, a number of questions are still open. For instance, we are told that IEM is a consequence of the fact that avowals are groundless (in the specific sense meant by Wright). But what we want to know is how it is possible to make *self*-ascriptions which are neither based on inference nor on observation. In other words, as we saw with Shoemaker, what we want to know is how it is possible that one's awareness of one's own mental properties licenses an ascription of them to *oneself*, when no object is in view (at least in some cases); and, secondly, why one cannot be mistaken in identifying the subject to which these properties belong. As long as there is no answer to these questions, then no real advance has been made in order to explain IEM.²² However, Wright is certainly right in suggesting that an explanation of IEM of non-inferential mental self-ascriptions has to be looked for in a wider domain, i.e. the domain of one's justification to move from awareness of one's own mental properties to the relevant self-ascriptions. We will come back to this in the following chapters.

Finally, let us concentrate on Wright's attack against the no-reference view. He writes:

[T]he idea that Wittgenstein's 'as subject' uses of 'I' are somehow shown to be non-referential by their having IEM should have been

²²Of course, since Wright's main aim is not to explain IEM he cannot be blamed for not having provided us with a fully-fledged explanation of why non-inferential mental self-ascriptions are IEM.

strangled at birth by the reflection that a similar immunity is characteristic of many *demonstrative* claims, in which case there is of course no question but that reference to an object is involved. If I see an object hurtling towards us and say, 'That thing is approaching very fast', there is no way in which that claim can be defeated (...) yet my original grounds for it survive as grounds for the claim 'Something is approaching very fast'.²³

Thus, Wright suggests that demonstrative judgements can be IEM and yet nobody would deny that they are genuinely referential. Consequently, IEM is compatible with the fact that the singular term used to give expression to one's judgement be genuinely referential.

However, as we have noticed in the Ch. 2, it is *prima facie* true that first-person judgements which are IEM license the existential generalisation, contrary to the no-reference view. Hence, there is no obvious need to make this point by arguing indirectly from the demonstrative case. Thus, if the no-reference theorist is so much as in a position to try and defend her claim, she must be arguing on different grounds. In other words, she must be struck by something which is so relevant as to make her ignore or override ordinary semantic considerations. In fact, when we look closer at the no-reference theorist's position, it seems obvious that the route to this peculiar claim has to be viewed in the *epistemology* underlying non-inferential mental self-ascriptions. That is to say, the basis of the no-reference view is Hume's denial—endorsed by Lichtenberg, to an extent—that we are introspectively aware of a subject who has got mental properties, when we perceive or think. As a consequence, also

²³Wright 1998: 20.

the surface grammar of the phrases used to give expression to that awareness is reinterpreted in such a way that 'I' is taken to be similar to 'it' in 'It rains'. Hence, epistemological considerations make the no-reference theorist believe that from the presence of a grammatical subject it cannot be inferred that there is someone who has the property in question.

Hence, Wright is right in saying that IEM is not incompatible with the fact that the uses of 'I' as subject are genuinely referential. However, on closer reflection, it is clear that the no-reference view and the no-subject view are consequences of Hume's point. Once more, it becomes evident that the solution to our problems cannot but be found in a better understanding of the *epistemology* of first-person mental self-ascriptions and that reminders about the referring role of the first-person are not enough to settle the worries in this area. In particular, we need an explanation of why, from the mere awareness of mental properties, in which no proper object is in view (at least in some cases), subjects are entitled to self-ascribe those properties and why this ascription cannot be mistaken relative to the first person (no matter whether it is a self-ascription of a sensation, an outer perception or of a propositional attitude).

3. *A Reminder: Pryor's De Re and Which-Misidentification and the Corresponding Forms of IEM*

As we saw in chapter 1, Pryor puts forward a distinction between two kinds of EM and, consequently, of IEM. Recall the definitions offered by Pryor:

[W]e have a case of *de re misidentification* whenever the following three conditions obtain:

- (i) There is some singular proposition about x , to the effect that it is F , that a subject believes or attempts to express. (...).
- (ii) The subject's justification for believing this singular proposition rests on his justification for believing, of some y , that y is F and that y is identical to x . (...).
- (iii) However, unbeknownst to the subject, $y \neq x$. (...).

I do not assume that whenever your justification for believing one proposition rests on your justification for believing other propositions, you will have *formed* or even *entertained* a belief in those latter propositions. Nor do I assume that you have *based* the first belief on a belief in those latter propositions.²⁴

[Which-misidentification] occurs when:

- (i) A subject has some grounds G that offer him knowledge of the existential generalization $\exists xFx$.
- (ii) Partly on the basis of G , the subject is also justified, or takes himself to be justified, in believing of some object a that *it* is F .
- (iii) But in fact a is not F . Some distinct object y is F , and it's because the grounds G "derive" in the right way from this fact *about* y that they offer the subject knowledge that $\exists xFx$.²⁵

The definition of which-misidentification is then revised in the following way:

[W]e can say that a singular proposition about a to the effect that it is F is *vulnerable to wh-misidentification when justified by grounds G* just in case:

When that proposition is believed on grounds G , it is possible for those grounds to be defeated by undercutting evidence in such a way that the following two conditions hold:

- (i) the combination of G and that undercutting evidence *no longer* justifies you in believing of a that it is F ; but

²⁴Pryor 1998: 274-276.

²⁵Pryor 1998: 282.

- (ii) the combination of G and that undercutting evidence *could*, by itself, offer you knowledge that $\exists xFx$.²⁶

We saw in Ch. 1 that, according to Pryor, while *de re* EM involves a mistaken identification component, which-misidentification does not. However, we maintained that the latter position is the result of a mistaken view about the “ingredients” which can occur in the identification component. We also claimed that once the identification component is liberalised in such a way as to include not only indexicals such as ‘I’, or proper names whose use is based on acquaintance with their bearers, but also rigidified descriptions, then also which-EM can be seen as a result of a mistaken identification component. Therefore, we concluded that it is wrong to claim that there are two kinds of EM (and consequently of IEM). Rather, there are different kinds of identification components, which can be the subject’s rational grounds for her judgement. Moreover, we claimed that there is no need for the subject actually to entertain such identification components, so long as the latter rationally ground her judgement.

Let us now consider Pryor’s subsequent claims. *De re* IEM occurs when the subject’s justification for believing [X is F] does not rest on her justification for believing an identity judgement. By contrast, *wh*-IEM occurs when the defeat of one’s grounds is not compatible with the possibility for one’s grounds to survive as grounds for the existential generalisation [Someone is F].

²⁶Pryor 1998: 284. Notice that ‘wh-EM’ and ‘wh-IEM’ are Pryor’s own abbreviations for which-EM and which-IEM respectively.

Pryor also holds that wh-IEM entails *de re* IEM. This is so because, if the judgement does not rest on an identification component then it follows that it is *de re* IEM. By contrast, he claims that there could be *de re* IEM, but it would not follow that there is wh-IEM.²⁷ That is to say, the absence of an identification component is not sufficient to guarantee wh-IEM. As an example of the latter possibility, he considers Evans' treatment of memory-based first-person thoughts. Since these thoughts are identification-free for Evans, they are *de re* IEM (according to Pryor's definition). However, Pryor argues that this does not entail that they are wh-IEM. He then goes on to discuss the case of a subject who quasi-remembers some event and self-ascribes the relevant property. Pryor argues that, in this case, the subject's judgement would be affected by wh-EM, although no identification component would be involved.²⁸

Some points are worth-making in this connection. First, although Pryor is right in saying that the actual judgement could fail to be formed or based on any identity belief or judgement, this, by his own lights, does not exclude the fact that one's justification for a certain judgement depends on believing some other identity proposition (cf. definition of *de re* EM above). Secondly, we saw that *de re* EM and wh-EM are not really two different kinds of EM. Hence, we can conclude that if EM arises this is due to a mistaken identification component. Putting these considerations together, we do in fact have our own understanding of EM. Namely, EM occurs iff the subject's judgement is causally

²⁷Cf. Pryor 1998: 285-286.

²⁸Cf. Pryor 1998: 288-297.

and rationally triggered by a pair of (not necessarily occurrent) beliefs in a predication component and in a mistaken identification component. As a result, we have it that if Evans' treatment of memory-based I-thoughts is open to EM, this must depend on the subject's belief in a mistaken identification component. As we saw at length in Ch. 5, discussing Evans' treatment of I-thoughts based on the deliverances of somatic proprioception, in a possible world where the subject is hooked up to someone else's body and she knew (or had reasons to believe) that she might be receiving information from someone else's body she could entertain EM I-thoughts. These, in turn, would be triggered by a (not necessarily occurrent) belief in the identity judgement [My body = the body I am feeling now]. We concluded that this shows that I-thoughts based on the deliverances of somatic proprioception are only *de facto* IEM. For, in a different possible world, they could be liable to EM. Similarly, we should conclude that also memory-based I-thoughts are only *de facto* IEM (relative to the state of information a subject may be in in a logically possible deviant causal world) and for the same reasons as before.²⁹

Finally, recall that in Ch. 2 we introduced two definitions of IEM. The first one mentioned the subject's commitment to an identification component, which, however, could not be mistaken. By contrast, the second mentioned the absence of such a commitment to an identification component. Here they are:

²⁹Pryor himself stresses the parallel between memory-based I-thoughts and I-thoughts based on the deliverances of somatic proprioception. Cf. Pryor 1998: 304, fn. 54.

IEM_{df1} A singular judgement of the form [a is F], formed on the basis of a recognition or an identification, is IEM, relative to the subject, iff it is not possible for the thinker to apply the particular singular concept [a] to any other object but the relevant object *a*.

IEM_{df2} A singular judgement of the form [a is F] is IEM relative to the subject, iff it is not formed on the basis of a recognition or an identification of the object thought about and, consequently, iff it does not involve an identification component.

The difference between Pryor's account and ours is that he thinks, mistakenly, that two ways in which a judgement can be IEM correspond to two ways in which it can be liable to EM. By contrast, we have maintained that there is just one kind of EM, i.e. that which involves a mistaken identification component, which is the equivalent of Pryor's *de re* EM. In Ch. 8, we will claim that there is just one kind of IEM, namely IEM_{df2} , which is the equivalent of Pryor's *de re* IEM. Hence, it follows that IEM is due to the fact that no identification component is involved. However, we must still explain why non-inferential mental self-ascriptions are IEM in this sense, and, moreover, why they are logically so. Yet, Pryor does not give us a clue as to why certain judgements, and, in particular, I-judgements are *de re* IEM. Hence, we are left without any explanation of logical IEM of non-inferential mental self-ascriptions and, consequently, without any satisfactory and complete response to the trilemma about the self.

4. Conclusions

In this chapter we have seen a number of ways of understanding EM and IEM. We have maintained that Shoemaker's definition of EM is basically right and that his distinctions between *de facto*/logical IEM and circumstantial/absolute IEM, once properly qualified, are relevant. However, we have claimed that he does not provide us with an explicit definition of IEM and with an explanation of why non-inferential mental self-ascriptions are logically so. Hence, he has failed to provide us with a revisionist account of IEM and, although he has rightly argued for compatibility between IEM and the fact that the self is embodied, he has failed to provide us with a fully-fledged response to the trilemma. We then moved to Wright's definitions of EM and IEM. We have claimed that they are sound ones and allow for compatibility between IEM and the fact that 'I' refers to an embodied self. This latter fact, however, is not enough to rule out the Idealist position, for the latter is motivated by epistemological considerations. In particular, it is motivated by Hume's insight that we are not presented with an object in introspection (at least when self-ascriptions of non-bodily sensations and of propositional attitudes are at stake). We have also claimed that although Wright gestures in the right direction for an understanding of IEM of non-inferential mental self-ascriptions, when he points out that it has to do with the grounds on which they are based, he does not substantiate it in detail. Consequently, he fails to provide us with a fully-fledged response to the trilemma. Finally, we have considered Pryor's position. We have argued that it neither provides us with a sound revisionist account of IEM, nor with a satisfactory explanation of IEM of non-inferential mental self-ascriptions. Thus, it fails to provide us with a deflationist response to the trilemma. Hence, we still need a sound account of IEM of non-inferential mental self-ascriptions. However, in order to provide one, we had better look at an account of how our

self-ascriptions of mental properties can be justified, although—as Hume noticed—there is no object in view when one is aware of one's own mental properties (at least in some cases). Hence, we will now turn to Peacocke whose recent work addresses this very question.

Chapter Seven

I-Thoughts and Δ -Theories: Peacocke

We concluded the last chapter by saying that a proper explanation of IEM of non-inferential mental self-ascriptions has to be looked for in the domain of the *epistemology* of these first-person thoughts. In this chapter we will consider Peacocke's account of first-person thoughts which was first presented in *Sense and Content*. This will allow us to supplement Evans' account of this class of thoughts with the quite intuitive idea that the possession of the first-person concept requires a subject to be able to form I-thoughts on the basis of her psychological evidence. Then we will claim that it is because the relevant self-ascriptions are IEM that Peacocke's account of first-person thoughts is correct, but that he does not fully explain why these self-ascriptions are IEM. Finally, we will turn to Peacocke's more recent work on the epistemology of non-inferential mental self-ascriptions presented in *Being Known*, in order to see whether it can shed some light on the reason why they are IEM. We will agree with Peacocke that in introspection no subject of thought is manifested as an object (at least in some cases). Yet, we will argue that Peacocke's account of non-inferential mental self-ascriptions in terms of what he labels "delta-accounts" can explain only in part why the subject is nevertheless entitled to *self*-ascribe the relevant properties, i.e. why the subject is not only entitled to ascribe the relevant

property to *someone*, but also why she is entitled to ascribe it to *herself*. For delta-accounts *presuppose* an explanation of this latter fact. In turn, only an explanation of this fact will allow us to see why direct awareness of one's own mental properties can lead to the corresponding self-ascriptions that are both representation-independent and IEM. Hence, Peacocke's account of the epistemology of some I-thoughts is very instructive but, in order to provide an exhaustive solution to the trilemma about the self, it needs to be supplemented with an account of why, when one is introspectively aware of a mental state, then one cannot be aware of someone else's. An explanation of this fact is what an explanation of logical IEM of non-inferential mental self-ascriptions requires. For, whether or not they are based on the presentation of the self as an object, it is the fact that one cannot be introspectively aware of someone else's mental state that guarantees that the subject who self-ascribes the mental property on those bases is the same as the one who enjoys the mental state. Hence, it is an explanation of this fact that guarantees that the relevant self-ascriptions are not based on taking someone else to have a given mental state and mistake that person for oneself.

1. *I-Thoughts*

Peacocke aims to offer a broadly Fregean account of I-thoughts. Recall the distinctive features of I-thoughts we individuated in Frege's approach and introduced in Ch. 4. Here they are: (a) I-thoughts are primitive; (b) they are private to their thinkers; (c) they are incommunicable; (d) they are objective.

Hence, we will expect Peacocke's account to satisfy these conditions. Here is his proposal:

There is some initial plausibility in the claim that when someone thinks of himself in the [self] type of way, his way of thinking of himself is specified by the mixed descriptive-demonstrative 'the person who has *these* conscious states', where the demonstrative picks out his token conscious states at the time of thinking.¹

How does this proposal fare with respect to (a)-(d)? Now, it looks *prima facie* hopeless with respect to (a) for the sense or mode of presentation of the first-person is given by a mixed description. In fact, we have already seen (cf. Ch. 4) that the application of Frege's test for difference of sense shows that the sense of the first-person cannot be equivalent to the sense of any definite description co-referring with the first-person. So here is Peacocke's way out of this *impasse*:

[W]e shall need to draw a distinction between two ways in which a description may determine an m.p. It may do so simply by giving the content of an m.p., as it does with descriptive m.p.'s. 'The person who has these conscious states' certainly does not give the content of the [self] m.p. Some can have *first-person-thoughts* without having the intellectual sophistication actually to think about his own experiences and thoughts, either demonstratively or in any other way. But giving the content is not the only means by which a description can determine an m.p. It can also specify what I shall call the *constitutive role* of an m.p.²

¹Peacocke 1983: 109.

²Peacocke 1983: 109-110. Note that "m.p." is an abbreviation for "mode of presentation".

So let us see how Peacocke characterises the notion of constitutive role:

I suggest that associated with each type of demonstrative m.p. is a kind of evidence which disposes a thinker to judge thoughts containing constituents of that type.

[W]ith [self] is associated evidence concerning the person who has the token experiences and thoughts which the thinker in fact has.³

Hence, we see that, according to Peacocke, a subject will have the first-person concept if she is disposed to judge [I am F] in light of the appropriate kind of evidence, i.e. when the person with *these* conscious states is F.⁴

A first minor qualification is needed here, however. For Peacocke claims that the mode of presentation of the first-person is given by 'the person who has *these* conscious states'. Yet, in fact, there is no guarantee that a person could keep track of her own mental states. Nor is there any reason why, in order to have the first person concept, this should be the case, at least *prima facie*.⁵ Rather, it seems enough, for a subject to be credited with the first-person concept, that she is disposed to judge [I am F] when she has got *one* particular conscious mental state. So the proposal is to consider the first-person mode of presentation as given by 'the person with *this* conscious mental state'. Thus, a

³Peacocke 1983: 110.

⁴In Peacocke 1992 concepts are equated with modes of presentation and constitutive conditions are replaced by possession conditions specified in a canonical form (*A(C)* form). So we will use these terms interchangeably.

⁵At least, Peacocke does not argue that one should be able to keep track of her own mental states in order to have the first-person concept and I do not see why we should impose this condition.

subject will have the first-person concept if she is disposed to judge [I am F] on the basis of an occurrence of a particular mental state. In this way, a possible objection to Peacocke's account, according to which it is surreptitiously presupposing some version of the bundle theory of the self would be avoided from the start.⁶ For his account would not presuppose the subject's ability to keep track of her own mental states across time.

Another minor qualification has to do with the fact that the sortal 'person' enters into the characterisation of the first-person mode of presentation, according to Peacocke. It is important to stress that, given his account of the mode of presentation in terms of constitutive role, this does not mean that in order to have the first-person concept one must master the use of the sortal 'person'. Rather, one must just be disposed to form I-judgements on the basis of the evidence that the person with this particular conscious mental state is *F*.

One further clarification has to do with the notion of evidence employed by Peacocke. He has in mind a notion of *canonical evidence*, which, in general, need not be conclusive. However, in the case of the first person, the sort of evidence he mentions is, after all, *conclusive*. For, if there is an *a priori* guarantee that one can only be aware of one's own mental states, and hence, that one can only demonstratively refer to one's own mental states, then it will be *a priori* guaranteed that by using the first-person on that kind of evidence, then one will refer to oneself. So, as we will see, it is only on the assumption that the subject who judges [I am F] and the subject who has got the relevant mental state are one and the same that the kind of evidence on which [I] is used

⁶Peacocke takes into account this possible objection. Cf. Peacocke 1983: 120.

in the judgement—specified by ‘the person with this conscious mental state’—will be conclusive evidence that it is the subject who judges [I am F] to be *F*. Indeed, this assumption is sound, but, as we will see, Peacocke offers no defence of it. We will come back to this point shortly.

Notice, moreover, that the constitutive role of [I] mentioned by Peacocke is compatible with the fact that one can be disposed to judge [I am F] on the basis of a *different* kind of evidence from the one required by the constitutive role of [self]. For instance, one can form the judgement [I am F] on the basis of one’s proprioceptive information, or on the basis of one’s perception of the environment. As we have seen with Evans, these kinds of judgements will be IEM. Hence, they give one conclusive evidence that the person who has the property *F* is indeed oneself. Still, Peacocke claims that his specification of the constitutive role of [self] is “meant to capture what is *essential* (italics mine) to the evidential pattern of first-person thoughts even when the collateral information is absent”.⁷ We will come back to this claim. For the time being, it suffices to say that, first, Peacocke has managed to find a way out of the charge that his account of the first-person concept would pass Frege’s test for difference of sense and would thus be inadequate. Moreover, he has found a way out of the obvious problem that we would like to grant subjects with first-person thoughts even when they are not actually able to articulate the mixed description which would give the content of the first-person concept. Finally, he has provided us with a constitutive role for [I] which specifies (at least in part)

⁷Peacocke 1983: 117.

the canonical evidence on which a judgement would count as a first-person one, which is compatible with the presence of other evidence but which is also, in this specific case, conclusive evidence.⁸

Let us now consider how Peacocke's account fares with respect to the Fregean claim (b) that the mode of presentation or sense of the first-person is private to the thinker. We have already seen how Evans distinguished between two possible readings of this claim (Cf. Ch. 4). On the one hand it can be thought that there are as many types of first-person concepts as there are thinkers. By contrast, on the other hand, it can be thought that there is just one first-person concept as a type that is exemplified in as many tokens as there are thinkers. Peacocke adopts this second reading and illustrates it in a way we have already mentioned in Ch. 4, but which is worth-repeating in this connection:

Suppose Peter thinks 'I am hungry' and Paul thinks 'I am hungry'. Since we are following the Fregean model, the modes of presentation they each express by 'I' must be different: for they determine different objects, Peter and Paul respectively. But of course when they think these thoughts, Peter and Paul think of themselves in the same *type* of way. It is this type which is denoted by '[self]'. The constituent of all Peter's first-person thoughts is called a token mode of presentation; it can be taken to consist of the type [self] indexed by Peter, that person himself. This token m.p. will be denoted by '[self_{Peter}]'. So it is token m.p.'s which are constituents of thoughts and which pick out particular objects. But one should not be misled by the label 'token': there is nothing relevantly unrepeatable about [self_{Peter}]. On many

⁸By contrast, according to Peacocke, the specification of the constitutive role of a demonstrative concept will not mention conclusive evidence.

different occasions, Peter may have attitudes to thoughts which contain it as a constituent.⁹

Hence, Peacocke's account of the concept [I] in terms of possession conditions is an elucidation of that concept as a type. But what is relevant is that each subject will employ a token of it that nobody else could entertain, for only the subject is immediately aware of her own mental states. Consequently, only the subject can be disposed to form a first-person judgement in light of that specific piece of evidence. This, as we will see, however, is a fact in need of explanation.

Notice, moreover, that this strategy allows Peacocke to avoid Husserl's conclusion (possibly endorsed by Frege as well, as we saw in Ch. 4) that the meaning of 'I' is always altering, from person to person and according to the circumstances of utterance. For the pattern of canonical evidence to which first-person judgements are respondent in any given occasion can be specified as a type. Consequently, it is true that each time a given first-person judgement is respondent to the occurrence of a specific mental state, but it is equally true that the judgement always responds, at least essentially, to occurrences of mental states.

As far as (c), i.e. incommunicability is concerned, Peacocke argues, on the one hand, that I-thoughts can find linguistic expression. Moreover, the meaning of 'I' as a linguistic device is correctly and fully specified by the token-reflexive rule (TRR) according to which each token of 'I' refers to whoever produces it. On the other hand, he claims that this should not lead to the dispensability of

⁹Peacocke 1983: 108.

the mode of presentation of the first-person and that the latter is, as such, incommunicable. Here is the relevant passage:

The position I am endorsing, then, is one according to which: (a) Barwise and Perry are right in thinking that the reference rule for 'I' fully determines its meaning in English; (b) The first-person way of thinking or mode of presentation exists and it is important for understanding 'I' that one realize that this word is used to express thoughts containing m.p.'s of this type; (c) Points (a) and (b) are compatible because there is an argument that in this case the reference rule determines what m.p. is expressed.¹⁰

The idea is this: the token-reflexive rule makes it the case that intentional, sincere utterances of 'I' do not only refer to the speaker, but express a first-person thought as well. Still, TRR cannot specify the sense or mode of presentation of the first-person. For not all first-person thoughts are uttered and, consequently, it would not make sense to say that in those occasions one thinks of oneself as the producer of a certain token of 'I'. Secondly, as Evans suggested (cf. Ch. 4), it would not be plausible to say that all there is to self-consciousness is one's ability to think of oneself as the producer of a certain token of 'I'. Nor would it be plausible to hold this position once TRR is amended in such a way as to be applicable to tokens of the concept [I], whereby one would be thinking of oneself as the producer of that token of [I]. Finally, if one were to defend this last version of TRR as specifying our first-person concept, one should also be prepared to hold that one would be thinking about oneself

¹⁰Peacocke 1983: 138.

as the producer of a certain token of [I].¹¹ But, in this case, the exercise of the concept of [I] would be presupposed. That is to say, one should already be able to entertain [I] in order to think of oneself as the producer of that token of the first-person concept. Hence, TRR cannot be seen as an explanation of what it takes for one to be able to have first-person thoughts or be self-conscious because it would already presuppose the occurrence of first-person self-conscious thoughts.¹²

From this we can conclude that Peacocke is right in holding that TRR does not specify the concept [I]. Consequently, he is also right in saying that TRR does not allow us to dispense with the mode of presentation of the first person. For utterances of 'I' would count as intentional just in case they were the expression of a first-person thought. So we do need an account of the first-person concept in order to have a complete account of our mastery of the first-person pronoun.

Finally, we have seen with Evans that the issue of (d) the objectivity of first-person thoughts is paramount. In fact, Evans claimed that first-person thoughts must be objective, in the sense of being about publicly accessible persons who think of themselves as such. Yet, given Peacocke's account of the first-person concept, it does not seem to be necessarily the case that, by being able to entertain the first-person concept, subjects should think of themselves as embodied. Consequently, they would not (necessarily) think of themselves as

¹¹Campbell 1995 defends the view that the token-reflexive rule can specify the sense of the first-person.

¹²Cf. Bermúdez 1998, Ch. 1.

physical entities on a par with all others. In this sense Peacocke's account of the first-person concept is not objective. But is this a real problem? In fact, it is not, for Peacocke's position is entirely *compatible* with the fact that selves are embodied and that they do think of themselves as such.

In fact, amongst the mental states which a subject could demonstratively refer in thought and which would be the evidence on which her use of the first-person is grounded, there are *bodily sensations*. However, bodily sensations give one a representation of oneself as embodied. Hence, even if the kind of evidence of the constitutive role of [I] is psychological, it also gives one a representation of oneself as embodied.

One may then argue that Peacocke is looking for the *canonical evidence* on which our use of the first-person is grounded. Hence, he is looking for necessary conditions for our use of the first person. Since, however, we can think of a subject who, despite being anaesthetised and hence unable to form bodily I-judgements on the basis of her proprioceptive feelings and of outer perception, is still able to entertain I-thoughts, then the latter must be grounded only on psychological evidence. Thus, one may conclude that the ability to form I-judgements on the basis of proprioception and perception is not necessary for self-conscious thought, while the ability to form I-judgements on the basis of evidence that the person with this particular mental state is *F* is.

However, recall that Evans does not require the relevant bodily self-ascriptions to be always in place (cf. Ch. 5). Rather, he argues that they must be *dispositionally* in place for one to be self-conscious. Hence, a subject who is anaesthetised (and also amnesiac) may have retained the disposition to think of

herself as embodied on the basis of the kind of evidence she may be prevented from having at a particular time.

Moreover, one may say that after all Peacocke is looking for *conclusive* evidence, i.e. evidence that guarantees that the person who appears to be *F* is indeed oneself. Hence, if it is somehow admitted that somatic proprioception and perception of one's environment may give one information relative to someone else's body, then it would no longer be guaranteed that the person who appears to be *F* is oneself. But, we have seen at length how bodily self-ascriptions made on the relevant bases are at least *de facto* IEM, hence the person who appears to be *F* on those bases is oneself.

Finally, notice that even if some deflationists like Shoemaker and Wright (cf. Ch. 6) were right in saying that the relevant bodily self-ascriptions are ultimately based on "avowable psychological matters", it would not follow that such psychological evidence does not give one a conception of oneself as embodied. For, as we have already remarked, it is precisely the fact that one enjoys *bodily* sensations that gives one a representation of oneself as embodied, e.g. as the person whose legs feel as bent.

Moreover, Peacocke's account if integrated with Evans' acceptance of the Generality Constraint will imply that for one to be able to self-ascribe a mental property one must be able to think of that property as satisfiable also by other persons and of oneself as someone who can be supposed by others to have such a property (cf. Ch. 5). We discussed the implications of this latter strand of GC in chapter five, but it is worth-repeating that, although not all mental properties would be self-attributed on the basis of a presentation in

introspection of oneself as an object, the ability to self-ascribe them would go hand in hand with the ability to think of oneself as embodied.

To conclude: the kind of evidence mentioned by Peacocke, although psychological, is not favourable to either the Cartesian or the Idealist conception of the self. Rather, it supplements Evans' position with respect to the possession conditions of the first-person concept in important ways. For surely part of what it takes for a subject to be self-conscious is to be able to think of oneself as an author of thoughts and as a possessor of sensations and experiences. Yet this does not mean that one would be thinking of oneself as a possessor of *just* those properties. We will now turn to Peacocke's account of logical IEM.

2. Peacocke's Account of IEM of Non-Inferential Mental Self-Ascriptions

The question is: what is it that makes it the case that 'I'm in pain' is logically IEM? Peacocke's answer is this:

[T]he possession by the first-person m.p. of the constitutive role 'the person who has *these* experiences' can explain why it is that in the first-person present-tense ascriptions of conscious states, one does not have to apply any tests of identity to check that the person in pain is oneself.

He then goes on to say:

For any given particular token conscious states and thoughts a thinker enjoys, only he and no one else has those particular token states and thoughts (and this is not a contingent fact). So he alone

can make demonstrative reference in thought to them in ways someone is capable of making such reference in virtue of having those particular states.¹³

The argument seems to be as follows. In order to make demonstrative reference to a mental state one must be conscious of it, i.e. one must be introspectively aware of it. Yet, (a) one can only be introspectively aware of one's own mental states, hence (b) only the subject who has the mental state can demonstratively refer to it, therefore (c) the use of 'I' on the basis of the evidence specified by 'the person with this conscious mental state' will make it the case that one's use of 'I' necessarily refers to the owner of the mental state. Hence, when 'I' is used on the basis of the evidence specified by 'the person with this conscious mental state' one will not have to apply any tests of identity in order to know that the person who has the property of being in pain is oneself.

Notice, however, that Peacocke's account of why [I am in pain] is identification-free, and hence IEM, rests on assumption (a). For if it were possible to be introspectively aware of someone else's mental states, then it would become possible to refer to them demonstratively, contrary to (b). But then, the evidence that the person with this conscious mental state is *F* would no longer give one conclusive evidence that one is *F oneself*. The consequences of this fact are twofold: first, one's non-inferential mental self-ascriptions would not be IEM, for one could have evidence that someone is in pain, wrongly take that person to be oneself and hence judge [I am in pain],

¹³Peacocke 1983: 121. Cf. Peacocke 1983: 175.

when in fact someone else is. Secondly, 'the person with this conscious mental state' would not be conclusive evidence that the [I] used in the judgement based on that evidence picks out the same subject which is identified through the mixed-definite description. Hence, that description would no longer be guaranteed to give the correct specification of the constitutive role of [I]. Thus, if we want an explanation of logical IEM (and motivate the fact that 'the person with this conscious mental state' is a good specification of the constitutive role of [I]) we need to rule out the possibility that someone could be introspectively aware of someone else's mental states and demonstratively refer to them. However, this is what Peacocke writes on the topic:

No one can think of another's states demonstratively where the demonstrative way is dependent upon being in those conscious states. To say this is not to be committed in any objectionable sense to the privacy of another's sensations and conscious thoughts: it is quite consistent with these claims about the constitutive role of [self] that one can on occasion know what type of experience and what thoughts another is having. All that is entailed is that one cannot think of the other's conscious states in the particular ways in which that other person can think of them.¹⁴

Of course what Peacocke is saying here is perfectly correct, but the question still remains: why is it that we cannot be directly or introspectively aware of someone else's mental states? So long as no answer is given to this question, then no satisfactory account of IEM of non-inferential mental self-ascriptions would be given either. We will now turn to Peacocke's more recent work on the

¹⁴Peacocke 1983: 121-122.

first-person where the issue of IEM is addressed again, although from a different perspective.

3. *Peacocke's Account of the Epistemology of First-Person Thoughts: Δ -Theories*

In "Self-Knowledge and the Illusions of Transcendence" Peacocke addresses the question of the epistemology of first-person thought.¹⁵ In this connection he distinguishes three phenomena that concern some I-thoughts: (i) immunity to error through misidentification, (ii) infallibility, (iii) representation-independence. He deems the third of these responsible for "the illusions of transcendence" with regard to the first-person. This claim is defended on theoretical grounds rather than on historical ones (although some Kantian exegesis is offered). The supporters of the transcendental view about the self mentioned by Peacocke are Kant, Schopenhauer, Husserl and Wittgenstein. In short, all of them have held the view that the subject who does the experiencing and the thinking, i.e. the one who is able to think [I see that p] and [I think that p], cannot be an empirical subject, for no such subject is manifested in experience and thought. It then becomes mysterious why these thoughts should in fact be in the first-person, rather than being impersonal, e.g. [There is a perception] or [There is thinking], since no empirical subject is manifested. Hence, in order to guarantee one's entitlement to the self-ascription they assume that the self referred to is a

transcendental subject, i.e. they assume that the self shrinks to a condition of possibility of thought and experience. Peacocke aims to offer a diagnosis of their mistake and to provide an account of the epistemology of this kind of thoughts which does not make mysterious the fact that they are indeed *knowledgeable self*-ascriptions. Moreover, he wants to show that a correct explanation of these self-ascriptions is compatible with the fact that the subject referred to be an empirical subject. In fact, he thinks that their mistake calls for what he labels “the integration challenge”.¹⁶ That is, the challenge to integrate considerations having to do with the *content* and the *epistemology* of the relevant class of thoughts, in such a way that one will no longer feel the need to appeal to peculiar metaphysical positions in order to account for the problematic phenomena. Let us consider Peacocke’s claims in more detail.

The basic question addressed by Peacocke is this: how is it possible for the judgement [I ψ that p],¹⁷ when based on direct awareness of one’s own mental states to amount to knowledge? This question can be broken down into two parts: (a) how can a subject be justified in judging [I ψ that p] when her ground for the judgement is awareness of a mental state of type ψ and content p ? That is to say, how can the subject be justified in *self*-ascribing that property if no subject of thought is manifested in introspection? Moreover, (b) how can the judgement [I ψ that p] amount to knowledge? Obviously, given our purposes,

¹⁵Peacocke 1999, Ch. 6.

¹⁶Cf. Peacocke 1999, Ch. 1.

¹⁷ ψ ranges over psychological predicates.

Peacocke's answer to question (a) is more relevant. Hence, in order to evaluate it, we will have to consider his claim that non-inferential mental self-ascriptions are representation-independent in more detail. Moreover, we will consider whether, despite the endorsement of the claim that no self as an object is presented in introspection, Peacocke manages to give an adequate account of the fact that one is nevertheless entitled to ascribe the mental state not only to someone, but also to oneself. Finally, we will see whether Peacocke manages to account for the fact that in introspection one can only be aware of one's own mental states so that there is no possibility of having reasons to believe that someone is *F* and be wrong about the fact that it is oneself who is *F*. This, in turn, will explain why 'the person with this conscious mental state' is a good specification of the constitutive role of [I].

3.1 *Representation-Independence*

First of all Peacocke introduces the notion of *representation-dependence* (hereafter, RD) thus:

We can say that a use of the first person, in a particular belief with the content 'I am *F*', is *representationally dependent* if

- (i) 'I am *F*' is the content of one of the thinker's current mental states, a state which represents that content as correct; and
- (ii) the thinker forms the belief 'I am *F*' by taking the mental state mentioned in (i) at face value, in respect of this content. That is,

(...) the thinker is operating in a mode in which such states are taken at face value.¹⁸

Here is an example of RD use of the first person:

The way the visual experience represents the world as being is one which justifies his [i.e. the subject's] acceptance of the first-person content endorsed in his belief 'I am in front of a door' (...). More generally, when a person forms a perceptual belief 'I am *F*', he does so because his experience itself has the content 'I am *F*', or has some content which justifies the content 'I am *F*'.¹⁹

So the idea is that, as in Evans, one's perceptual experience gives one egocentric information, i.e. information of the environment relative to the subject. Consequently, its content comes, as it were, in the first-person way, i.e. it includes a representation of the subject. Hence, if one takes such a content at face value, then one is immediately in a position to form a first-person judgement. For this reason, the relevant judgement will contain a representation-dependent use of the first-person. Let us now turn to Peacocke's characterisation of *representation-independent* uses of the first-person:

A representationally independent use of the first person, in a belief 'I am *F*', is a use in a belief which the subject forms for a reason, but whose content 'I am *F*' does not meet the conditions (i) and (ii) in the definition of representational dependence.²⁰

The examples given by Peacocke of RI uses of the first person are the following ones:

¹⁸Peacocke 1999: 265.

¹⁹Peacocke 1999: 264.

²⁰Peacocke 1999: 266.

I am thinking about Pythagoras's Theorem.
I see the phone is on the table.
I remember attending the party.
I remember that Russell was born in 1872.
I am beginning to dream.
I fear that the motion will not be carried.²¹

Peacocke then makes it clear that RI first-person judgements are held for reasons that are quite special, though. For instance:

[w]hen our thinker comes to judge 'I am thinking about Pythagoras's Theorem', his reason for his judgement is not a mental event or state whose representational content is that he is thinking about Pythagoras's Theorem. His reason is rather his particular occurrent thoughts about Pythagoras's Theorem.²²

This much seems to be clear, RI uses of the first-person are those which are not grounded in mental states that contain a representation of *oneself as having that particular mental state*. This qualification is needed in order to avoid at least the following objection.

Consider one of the examples of RI use of the first-person mentioned by Peacocke. That is to say [I see the phone to be on the table]. Here is Peacocke's explanation of it:

When our thinker judges 'I see the phone to be on the table', his reason for making his judgement is the occurrence of his visual perception. That perception is a seeing that the phone is on the table; but that perception is not one which has the representational content that he is seeing the phone. An experience in a given modality does

²¹*Ibid.*

²²Peacocke 1999: 267.

not normally have the representational content that the subject is having an experience in that modality.²³

It is then clear that for this example to count as RI, it must be the case that its grounds do not involve the representation of *oneself as seeing* the table. Otherwise, if the point were just that the grounds should not contain a representation of the subject, this would contrast with Peacocke's example of RD use of the first-person, namely [I am in front of a door], we discussed above. For, in that case, he argued that the experience of seeing a door does indeed have a first-person content.

What we can learn from this discussion is the following: the *content* of the experience or of the mental state is relevant to the determination of which sort of use of [I] is made in the judgement *only in the case of RD uses of the first-person*. By contrast, RI uses of the first person can be grounded in any sort of experience or mental state, irrespective of its content, because the relevant judgements are the result of the subject's sensitivity to the *occurrence of the experience or of the mental state itself*. Similar points can be made with reference to Peacocke's example [I remember attending the party], where the content of one's memory is, in normal conditions, about oneself and yet the use of the first-person is said to be RI. So we can generalise and conclude that the content of the mental state is relevant to the determination of the RD use of [I] in the judgement, but not to the determination of RI uses of [I]. For the latter ones are just the result of one's sensitivity to the occurrence of one's own mental states, irrespective of the latter's content. Hence, Peacocke is drawing

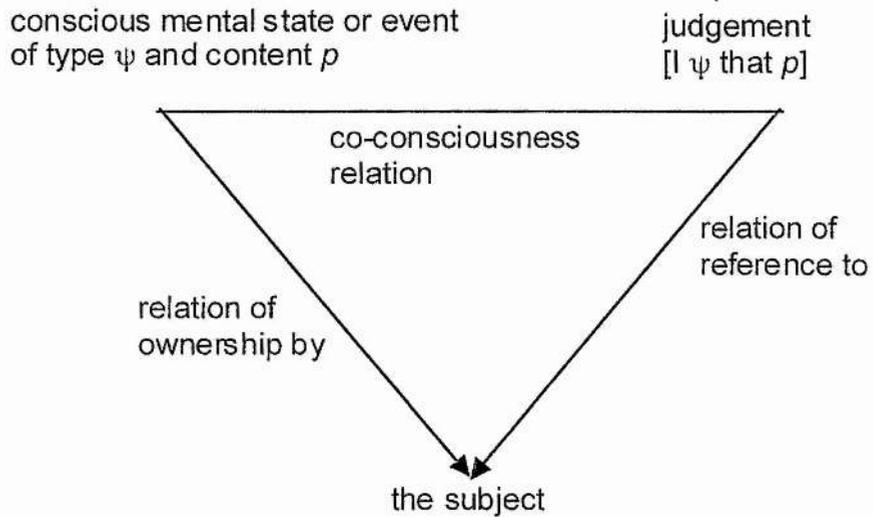
²³*Ibid.*

attention to the fact that mental states, whatever their content may be, do not as such manifest who their owner is. Yet, they can lead to the corresponding *self*-ascription, in suitably conceptually endowed thinkers, thus giving rise to the use of [I] as subject. Hence, what Peacocke aims to explain is the passage from one's awareness (which can either be perceptual or in any given propositional attitude) that *p* (which can also contain a representation of oneself) to the *corresponding self-ascription of the relevant mental state*. Let us now turn to Peacocke's account of the epistemology of these I-thoughts in terms of what he labels the "delta-account".

3.2 Δ -Theories

Δ -theories are introduced by Peacocke in order to explain how awareness of one's own mental states can lead to self-ascriptions of the relevant mental properties which amount to *knowledge*.²⁴ Yet, Δ -theories can help us get clear on the role that IEM plays in the epistemology of non-inferential mental self-ascriptions. For this reason it is worth considering them in detail. Here is the delta-model.

²⁴Cf. Peacocke 1999: 272, 273.



According to this model, a conceptually endowed creature, by enjoying a mental state of type ψ and content p is immediately in a position to form the judgement [I ψ that p]. In fact, her possession of the concepts [I], [ψ], and the concepts which are needed for entertaining the content p is shown by her disposition to judge [I ψ that p] in light of the appropriate evidence, i.e. in light of being enjoying the relevant mental state.

According to Peacocke, first-person beliefs formed according to this model are true ones. That is so because (a) the owner of the conscious state is the reference of [I] in the judgement; (b) the person is right about the kind of mental state she is enjoying, because this is not inferred from the content of some other representational state. Moreover, these true beliefs amount to knowledge (where knowledge is not just a matter of having reliably formed true beliefs), because (a) is *a priori* true and (b) holds for the following reason:

The transition (...) from an event with a certain representational content [a personal memory] to one with a corresponding content 'I

remember doing so-and-so' should not be construed as inferential. Rather, there is a rational sensitivity to the distinction between those of one's states which are memories, and those which are not. This sensitivity is already employed in ordinary first-level conscious thought and practical inference. This pre-existent sensitivity is exploited by someone who has the concept of memory, and self-ascribes personal memories in the manner given in a delta account. Similar remarks apply to other conscious attitudes.²⁵

What is relevant to our purposes, however, is to understand in what sense Δ -theories can account for the subject's entitlement to *self*-ascribe a given mental state. Basically, what Δ -theories assume is that mental states are *owned* and that one can be aware only of *one's own* mental states, so that a suitably conceptually endowed creature is actually entitled to *self*-ascribe them, without any need to make sure whose mental states these are. Here is Peacocke's acknowledgement of the situation:

[A] delta account cannot amount to an elucidation of the relation of ownership, nor of its close relative, the relation of co-consciousness. On the contrary, delta accounts simply presuppose those relations. The most that can be said is that if a delta account is correct, a good explication of these relations must leave room for a delta account.²⁶

And again,

[T]he delta account offers no positive support for 'no-subject' theories of mental states. By 'no-subject theories' I mean theories according to which mental phenomena have no subjects—and a fortiori no owners—at all. Perhaps Hume held a no-subject view in some moods; it is also suggested in some early and middle period writings

²⁵Peacocke 1999: 277.

²⁶Peacocke 1999: 277.

of Wittgenstein; some Buddhist texts may also express it (...). A delta account explains the status as knowledge of beliefs containing representationally independent uses of the first person, while agreeing that conscious states are owned. So, if anything, delta accounts are rather in the first instance ammunition for the opponent of the no-subject theorist. These highly distinctive phenomena can be explained without resort to a no-subject account.²⁷

Indeed, in so far as the relations of ownership and co-consciousness are not explained, then it is quite mysterious how Δ -theories can so much as *explain* the subject's entitlement to *self*-ascribe specific mental states.

However, Peacocke offers an explanation of why mental states are necessarily owned. Here it is:

The very idea of a conscious state or event involves the existence of a subject of that state or event. A conscious state, as Nagel said, is one such that there is something it is like to be in that state. This must mean: something it is like for the subject. Reference to the subject of conscious mental states is essential in elucidating what it is for them to be conscious.²⁸

Hence, according to Peacocke, conscious events have phenomenal content. Thus, there must be a subject for whom there is something like being in that state. Notice, however, that whether or not we endorse the thesis that all

²⁷Peacocke 1999: 278. Notice the tension between Peacocke's initial claims that the early Wittgenstein was a supporter of the transcendental self and the claim that in his early stage he held a version of the no-subject view. In fact, as we have maintained all along, Wittgenstein's no-subject period coincides with his middle writings.

²⁸Peacocke 1999: 292. Peacocke himself acknowledges to be indebted to Galen Strawson in this connection. Cf. Strawson 1994: 129-134.

conscious events have phenomenal content, still it is the very idea that a mental state be conscious which involves the fact that that state is occurring in *someone's* consciousness, i.e. that there is someone who is introspectively aware of it. That is, we are considering mental states as objects of awareness and this presupposes that there is someone who is aware of them.²⁹ For 'being aware of' is a relational property which involves the existence of someone who is aware of something. Hence, it is only a misconstrual of this relation which may lead one to think that there can be mental states which are objects of awareness without there being someone who is aware of them.

However, part of the explanation is still missing. For a complete account of the subject's entitlement to *self*-ascribe a certain mental state should comprise also an explanation of why the subject is right not only in ascribing those mental states to someone, but, above all, of why she is right in ascribing them to *herself*. We have seen as part of the explanation may go: since the constitutive role of [I] is specified by 'the person with this conscious mental state', then when one is aware of a mental state then one is entitled to self-ascribe it.

However, as we have seen, this account of our first-person concept is correct only in so far as it is assumed that the subject who has the mental state and the one who self-ascribes it in the judgement are one and the same. But only an

²⁹ However, by saying that mental states are objects of awareness, we do not mean to say that they are like objects, given in one's mental theatre, of which one is introspectively aware. In fact, mental states could be taken as acts, e.g. the act of thinking that *p*. Such an act, in turn, can be object of awareness for someone. That is to say, the subject may be introspectively aware of thinking that *p*.

account of the relation of *co-consciousness* will make it evident why whenever a subject is immediately aware of a mental state she cannot be wrong or even wonder about the fact that *she herself* is having it, so that it is a priori guaranteed that the subject who is aware of the mental state is also and necessarily the one, who, being aware of it, is in a position to self-ascribe it in the judgement. Hence, we still need an explanation of why, whenever a subject is immediately aware of one of her mental states, she is *ipso facto* entitled to self-ascribe it without finding out whose mental state it is.

4. Conclusions

In this chapter we have argued that Peacocke's account of first-person thoughts is essentially correct and can be profitably combined with Evans'. The result is that our first-person concept is a concept of an entity who thinks of herself as embodied and as a possessor of mental properties. Indeed, this is a vindication of the intuitive thought that we think of ourselves as persons, i.e. as irreducible entities with both physical and psychological properties.

We have then moved on to Peacocke's account of IEM, in terms of the constitutive role of [self]. We have made explicit that it rests on the assumption that one cannot be introspectively aware of someone else's mental states. We have then maintained that it is an explanation of this assumption which is needed in order to account fully for logical IEM of non-inferential, introspection-based mental self-ascriptions.

We have then turned to Peacocke's more recent work on the epistemology of first-person thoughts. We have argued that he is right in claiming that conscious mental states, which may fail to present their bearer, are nevertheless owned. Yet, we have also argued that Peacocke's explanation of why from the awareness of a mental state one is entitled to ascribe it to *oneself* is partially wanting. For it presupposes that the subject who is in a position to judge [I am F] on the basis of the awareness of the mental state is the same subject who has it. It has to be stressed that it is not our intention to deny the soundness of this assumption, but to defend it on rational grounds, i.e. we want to explain why it is *logically impossible* for someone to be aware of someone else's mental states and demonstratively refer to them. We will now turn to this task.

SECTION IV

THE EXHAUSTIVE SUBSTANTIVE RESPONSE TO THE TRILEMMA

Summary

In this section we summarise our previous findings and provide an *exhaustive substantive response to the trilemma*. To such an end,

- we offer a proper revisionist account of IEM and arguments in order to maintain that there *must* be IEM uses of demonstrative concepts and of the first person in order for judgements to be possible at all (§ 1);
- we point out that IEM so understood is an epistemological phenomenon, which has to do only with the grounds for one's judgement (§ 1.1);
- we summarise our findings according to which it is not true that all and only mental self-ascriptions are IEM, for there are some mental self-ascriptions which are not IEM, and also because there are some bodily self-ascriptions which are at least *de facto* IEM (§ 2);
- still we must explain why introspection-based mental self-ascriptions are *logically* IEM, i.e. they are such that they could never involve as part of their grounds a belief in an identification component. In order to do so, we recapitulate the gist of our discussion of Peacocke's proposal and we point out that it rests on an assumption, i.e. that one can only be introspectively aware of one's own mental states (§ 3). We contend that an explanation of why this is *a priori* true is the missing element of both an explanation of logical IEM and of why the constitutive role of [I] is given by 'the owner of this mental state';

- we provide such an explanation (§ 4): if one is immediately aware of a mental state, then this means that that mental state is occurring in one's consciousness. Hence, it follows that each mental state one is introspectively aware of is one's own;

- we consider and dismiss a possible counter-example to this view which has been recently put forward by John Campbell (§ 5);

- we conclude that we have managed to provide an *exhaustive substantive response* to the trilemma about the self (§ 6).

Chapter Eight

Immunity to Error through Misidentification of Mental Self-Ascriptions and the Exhaustive Substantive Response to the Trilemma

Our inquiry into EM/IEM and the nature of our first-person concept started with a passage from Wittgenstein's *The Blue Book*. There, as we saw, Wittgenstein was trying to suggest several claims: (a) EM/IEM are genuine phenomena which are characteristic of some uses of the first-person; (b) whether an I-judgement is EM or IEM seems to be a function of its subject-matter. In particular, (b. 1) all and only mental self-ascriptions seem to be IEM; finally, (c) uses of 'I' as subject (i.e. those which are IEM) are not about a particular person. Hence these uses of 'I' are not genuinely referential. We also saw how a possible consequence of Wittgenstein's claim is that mental states lack an owner. In this chapter we will summarise our discussion of (a)-(c) supplementing it with some further arguments where necessary. Finally, we will provide an explanation of logical IEM of introspection-based mental self-ascriptions, which will allow us to put forward an exhaustive substantive response to the trilemma.

1. EM and IEM

In Ch. 1 and 2 we defended and explained the first point made by Wittgenstein, namely (a) EM/IEM are genuine phenomena which are characteristic of some uses of the first-person. We defined EM as follows:

A singular judgement of the form [a is F] is affected by EM, relative to the subject, iff the thinker has the singular concept [a] of an object *a* and applies [a] to the distinct object *x* as a consequence of either a misrecognition or a misidentification.

We pointed out that EM has to do with the *grounds* on which the judgement is based. That is to say, a judgement is EM iff the subject bases her judgement on a pair of (not necessarily occurrent) beliefs in a predication component and in an identification one and the latter component is wrong.

As far as IEM is concerned, we saw that there are two possible definitions. They are as follows:

IEM_{df1} A singular judgement of the form [a is F], formed on the basis of a recognition or an identification, is IEM, relative to the subject, iff it is not possible for the thinker to apply the particular singular concept [a] to any other object but the relevant object *a*.

IEM_{df2} A singular judgement of the form [a is F] is IEM relative to the subject, iff it is not formed on the basis of a recognition (or identification) of the object thought about and, consequently, iff it does not involve an identification component.

We saw that the attempt to comply with the first definition leads to either the Cartesian or the Idealist position about the self; or else to the impossibility—on the view that the self is identical with a human being—of explaining IEM of (at least) mental self-ascriptions. By contrast, the second definition does not lead to these difficult positions and is indeed compatible with primitivist views about the self, according to which the self is identical with a human being. However, that mental self-ascriptions in particular comply with it still requires showing.

One could think that our preference for IEM_{df2} over IEM_{df1} is *ad hoc*. That is to say, since IEM_{df1} necessarily leads to unbearable conclusions, IEM_{df2} is obviously preferable because, at least *prima facie*, it does not incur those problems. But there does not seem to be any independent argument in favour of IEM_{df2} , apart from its *prima facie* aptness to avoiding the dilemma.¹ However, it is worth bearing in mind that for misrecognition and misidentification to be possible at all, and hence, for error through misidentification to occur at all, the exercise of two singular concepts must be involved. In particular, the identification concept must be possessed *already*. Hence, on pain of an infinite regress, there must be a separate exercise of these two concepts and, in particular, an exercise of the identification concept which is not itself based on recognition (or identification), i.e. an exercise of it outside an identification component. That is to say, according to the terminology we introduced in the first chapter, there must be some uses of those concepts, which can eventually feature in an identification component, which figure only in an *individuation component*, i.e. as subject in a judgement that is not based on recognition (or

¹Cf. Ch. 2.

identification), but just on mere individuation of a particular. Hence, these judgements will be IEM according to the second definition introduced above.

Therefore, it is obvious, recalling our examples in Ch. 2, that many uses of demonstratives will be IEM. But it is equally obvious that also (at least some) uses of indexicals such as 'here', 'now' and 'you' will turn out to be IEM. That is, any judgement containing these concepts, when formed on the basis of a mere individuation of a particular, a person, a place or a particular time, will be IEM. In fact, it is the primary function of indexical concepts to be used on the basis of an individuation of either a particular—persons included, of course—or a place or a time. For, without such a kind of concepts, exercised on the basis of an individuation, we would lack the building-blocks of our judgements.

However, if we stick to the first definition of IEM, then demonstrative judgements could not be IEM. For, in the case of empirical objects, it can never be the case that the subject be infallible in recognising them. This is why it has been thought that perceptual demonstrative judgements are *always* liable to EM. Thus, we can better appreciate why Cartesianism and Idealism have been thought to be the only possible ways of explaining IEM of mental self-ascriptions: since ordinary empirical objects could not be at the origin of a IEM judgement, given IEM_{df1}, the same has been taken to be true of persons. In other words, a peculiar metaphysics of the self would be necessary, given IEM_{df1}, to explain IEM of some first-person judgements.

¹Cf. Ch. 2.

But why is it that demonstrative judgements have been thought to involve identification, thus necessarily involving an identification component? The answer to this question depends on the fact that demonstrative concepts have been considered necessarily to involve a *general* concept. In other words, demonstratives have often been taken *not to be sufficient*, on their own, to pick out an object in the context of utterance and in need of a sortal completion.² Although this is obviously true in the case of demonstratives as *linguistic* devices, it has been taken to be true of demonstrative *concepts* as well. Hence, all demonstrative concepts have been taken to have the form [This/That F]. However, it has also been thought that the application of a general concept—like the sortal concept [F]—would always involve some form of identification. That is to say, it would always involve taking a particular object as an instance of a certain sort or kind. Accordingly, it has been deemed that the general scheme activated for recognition which consists in taking an object to be an instance of an identification concept is already applicable at the level of individuation, although the identification concept would be a general and not a singular one. Thus, all cases of individuation have been treated as cases of identification. However, we can always be mistaken in taking a particular object

²This is clearly stated by Anscombe, for instance, but it can also be seen at work in Frege and Wittgenstein. Anscombe 1990: 142 writes: "Assimilation [of 'I'] to a demonstrative will not (...) do away with the demand for a conception of the object indicated. For, even though someone may say just 'this' or 'that', we need to know the answer to the question 'this *what?*' if we are to understand him; and he needs to know the answer if he is to mean anything".

to be an instance of a sortal concept and this is why—it has been argued—demonstrative judgements are always liable to EM.

This model, however, is wrong for two reasons. First, in order to bring any object under a sortal concept we must have already formed a concept relative to the object. Such a base concept would be [This] or [That], i.e. the concept of that object, whatever it is, which is perceptually given to us. Hence, for complex demonstrative concepts to be possible at all, we do need simple ones, based on the individuation of a particular before its sort has been determined. Therefore, there are uses of demonstratives, which are indeed fundamental to the very possibility of judgement, which are IEM. But only IEM_{df2} is compatible with this claim. Therefore, IEM_{df2} ought to be preferred over IEM_{df1}.

Secondly, this model obviously rests on a conflation between identification and *characterisation*. For, when we individuate a particular like *that F*, this is not the result of an identification, i.e. it is not the result of an *identity judgement* involving two singular concepts, to the effect that they are taken to be coinstantiated. Rather, it is the result of taking a particular object to be an instance of a certain *sort* or *kind*, which could be instantiated also by other individuals. Hence, our complex demonstrative concept [That F] is, in fact, the result of a *predication* or a *characterisation*. Of course we can be wrong about taking a given particular to be an instance of a certain sort or kind, but that would be a mistake of predication, not of identification proper. Therefore, it could not count as a case of error through misidentification properly so-called, since it would not involve the exercise of a pair of *singular* concepts.

In the case of the first person we have seen that in order for [I] to be used in a judgement that can properly be said to involve error through misidentification, it must already be available to the subject. In particular, on pain of an infinite regress, it must already be available as a concept that does not feature in an identification component. Therefore, there must be uses of [I] which are IEM_{df2} . In fact, both Evans and Peacocke, as we have seen, aim to specify a pattern of evidence which would immediately (i.e. non-inferentially) license an I-judgement, thus leading to an IEM_{df2} use of the first-person concept. This, in turn, provides us with the concept [I] which, in an episode of misrecognition or misidentification, we wrongly apply in such a way as to produce an EM I-judgement. Once again, this result is compatible only with IEM_{df2} . Hence, given our distinction between conservative and revisionist accounts of IEM (cf. Ch. 2), drawn on the basis of whether they comply with IEM_{df1} or IEM_{df2} , we can conclude that a proper solution to the trilemma can only be *revisionist*.

Finally, notice that since an IEM judgement is compatible with its being false, our reasons for saying that there must be IEM uses of demonstratives and indexicals differ, at least in part, from Evans' (Ch. Ch. 3 and 5). That is to say, we cannot argue that IEM judgements are necessary because, on pain of an infinite regress, *knowledge* must rest on judgements which are not themselves the result of an identification. Rather, our argument stresses that the very possibility of there being identification *judgements* at all requires that the identification concepts be possessed independently of their occurrence within identification components, since, in order for them to occur in identification components, they must be *already* available to the subject. Notice, however,

that our reasons are compatible with Evans', for *true* IEM judgements will be at the basis of knowledge.

1.1 IEM as a Feature of the Grounds of One's Judgement

It is a consequence of our way of understanding EM/IEM that IEM_{df2} has to be slightly amended in light of the fact that we have taken EM to arise in connection with judgements in which the identification concept is applied to the wrong object because the *grounds* on which the judgement is based would involve a wrong identification component. Hence, IEM_{df2} must be replaced by the following definition:

IEM_{df2*} A singular judgement of the form [a is F] is IEM relative to the subject, iff it is not formed on the basis of a recognition (or identification) of the object thought about and, consequently, iff it does not contain an identification component as a *rational ground* for the judgement.

Let us expound on IEM_{df2*} before going on. A problem with IEM_{df2} is that it might suggest the thought that IEM has to do with the *content* of one's judgement. One could in fact think that, in so far as we do not bring identification concepts to bear in the judgement—thus sticking to, for instance,

demonstratives—then we are *ipso facto* guaranteed that the judgement will be IEM.³

However, we can think of counter-examples to IEM_{df2} *so understood*. For instance, we could think that in so far as we are sticking to the use of a demonstrative expression like 'That woman', our statement 'That woman is flying to Italy tomorrow' will always be IEM. However, this conclusion would be mistaken. For, if the judgement has been made on the basis of collateral information, like [That woman = NN], [NN told me she is going to fly to Italy tomorrow], therefore [That woman is going to fly to Italy tomorrow], then, of course, there is room for misidentification to occur, because of an error in the identification component, although the judgement contains a demonstrative concept. Hence, it would be wrong to conclude that EM is a function of the *kind* of concepts involved in the judgement. That is why it is important to stress that the judgement is IEM iff it has not been formed on the basis of a recognition (or identification), i.e. iff it does not involve an identification component. Indeed, this is a point we have already stressed in connection with Evans' discussion of identification-free judgements in the narrow sense and of the example of [That man is the chairman of the Department] (cf. Ch. 3).

³Pryor makes the same point when he writes: "However, we should note that it's not—or not just—the *propositional content* of a belief that determines the belief's scope for error through (...) misidentification. Whether or not a belief is immune to (...) misidentification *depends on its grounds*". Pryor 1998:279. Cf. also Evans 1982, Wright 1998, Peacocke 1999.

But, then, we can ask at which level we should look for the presence (or the absence) of the identification component. As we saw in Ch. 1, the idea is that it is not necessary for the subject to form the judgement on the basis of an occurrent inferential process involving an identification component. Rather, we are saying that the belief on the subject's part in an identification component causally and rationally triggers the judgement. Hence, the identification component must be available to the *subject* as part of the *grounds* for her judgement. Similarly, in the case of IEM judgements, they are such because they do not involve as part of their grounds a belief in an identification component on the subject's part. Thus, we have agreed with Evans in so far as he himself stresses that EM and IEM are a consequence of either identification-dependence or identification-freedom respectively, and that they have to do with the *grounds* on which a judgement is based. Thus, we have agreed with Evans on the fact that those grounds are the *subject's grounds* for her judgement and must be available to her.

To conclude: given IEM_{df2}, we can now see what an explanation of IEM of I-judgements really amounts to. Namely, it amounts to an explanation of *which* I-judgements will not involve an identification component as part of their grounds and of *why* this is so. We will now summarise our previous findings and we will supplement them with new arguments when necessary.

2. *The Extension of IEM I-judgements*

In Ch. 5 we turned to a discussion of Wittgenstein's second claim, i.e. (b) the fact that an I-judgement is EM or IEM seems to be a function of its subject-matter, and (b. 1) all and only mental self-ascriptions are IEM. We did so by introducing Evans' criticism to the idea (b. 1) that all and only mental self-ascriptions are IEM. As we saw, not only does he show that when mental self-ascriptions are based on inference, they are liable to EM, but, more importantly, he also shows that there are some bodily self-ascriptions which, if made on appropriate grounds, are IEM. For instance, when a bodily self-ascription like [My legs are bent] is based on the deliverances of somatic proprioception, it does not involve as part of its grounds a belief on the subject's part in a predication and in an identification component such as [The legs I am feeling now are bent] and [The legs I am feeling now = my legs]. For, in this world, subjects receive somatic proprioceptive information from just one body, i.e. the body in which their brain is located. Hence, there is no reason to base one's judgement on a belief in an identification component such as the one mentioned above.

However, this, as we saw in Ch. 5 and 6, is compatible with holding that in a logically possible world, where there are deviant causal conditions and the subjects are *appraised* of the fact that they may be receiving information from someone else's body, their bodily self-ascription would be based on an identification component which, if wrong, would lead to a EM I-judgement.

Nevertheless, in this world, some bodily self-ascriptions conform to IEM_{df2}. This, in turn, shows that IEM is compatible with the fact that the self be embodied. Moreover, if the relevant self-ascriptions are based on “avowable psychological matters”, i.e. if they are based on a belief such as [I feel my legs as bent], as Shoemaker and Wright maintain (cf. Ch. 6), this means that some mental states like bodily sensations far from just giving one a conception of oneself as a mere author of thought and a possessor of sensations give one a representation of oneself as embodied, contrary to the Cartesian and the Idealist proposals.

So here we have the first elements of a *substantive* response to the trilemma, which aims to show that our concept of ourselves is anti-Cartesian and anti-Idealist. To recapitulate: (i) there are bodily self-ascriptions that are IEM (and must be at least dispositionally in place). These are based on discriminative knowledge of ourselves and allow us to meet RP. In turn, RP states a fundamental condition upon concept-possession (according to Evans). Moreover, if (following Peacocke), we hold that the manifestation of the possession of a concept is shown by being able to form a judgement containing that concept on bases which do not involve inferential components—for those, in turn, would require the subject’s possession of that concept—then being able to form the relevant bodily self-ascriptions would count as a manifestation of one’s possession of the first-person concept. Finally, (ii) even if the relevant bodily self-ascriptions are based on self-ascriptions of bodily sensations, this does not mean that the former are based on an identification of one’s body, or that one would just be thinking of oneself as a mere possessor of sensations. In

fact, precisely because those sensations give one a representation of oneself as embodied, one would thereby acquire a conception of oneself as a physical entity. Hence, our first person concept is anti-Cartesian and anti-Idealist.

However, in order to dispense fully with the Cartesian and the Idealist, a theorist who wants to give an exhaustive substantive response to the trilemma must show *why* non-inferential mental self-ascriptions are *logically* IEM, according to IEM_{df2^*} , i.e. why they are IEM_{df2^*} in any logically possible world. As we saw, however, Evans does not provide us with such an explanation and the deflationists do so only in part. Looking for such an explanation, we then turned to Peacocke.

3. *Logical IEM of Non-Inferential Mental Self-Ascriptions*

In Ch. 7 we considered Peacocke's claim that IEM of non-inferential mental self-ascriptions is explained by the fact that the constitutive role of [I] is given by the mixed definite description 'the person with this conscious mental state'. Since, according to Peacocke, one can only be aware of and demonstratively refer to one's own mental states, then the constitutive role of [I] explains why, when one judges e.g. [I am in pain] on the basis of feeling pain, one need not apply any tests of identity in order to find out who is in pain. That is to say, one does not have to apply any tests of identity in order to find out that the person who has that conscious mental state is the one to whom the [I] in the judgement refers.

We did not take issue with this explanation, but we simply pointed out that it is based on the *assumption* that one can only be aware of and demonstratively

refer to one's own mental states. In turn, we made it clear that it is not our aim to deny the soundness of this assumption. Rather, our aim is to *explain why* this assumption is *logically* and not just empirically sound. In other words, our aim is to show why there is indeed an *a priori* guarantee that the mental states one can be immediately aware of and demonstratively refer to in thought are one's own, so that it is right to characterise the constitutive role of [I] in terms of 'the person with this conscious mental state'. This in turn explains why, when one judges [I am F] on those bases, one does not have to find out that the subject referred to by [I] and the subject who has the mental state are one and the same.

We also saw that the interest of Peacocke's proposal is not just that it gives us part of an explanation of IEM of non-inferential mental self-ascriptions, but that it helps us to vindicate the idea that in order for one to have the first-concept one must have a conception of oneself as a person, i.e. as an entity with both physical *and psychological* properties. Hence, we argued for the compatibility between Evans' and Peacocke's accounts of the conditions of possession of the first-person concept.

In order to see whether Peacocke could offer us an explanation of the *a priori* guarantee that the mental states that one is immediately aware of and can self-ascribe in thought are one's own, we turned to his more recent work on the epistemology of non-inferential mental self-ascriptions. The interest of the views presented in *Being Known* is twofold, given our purposes. On the one hand, Peacocke's account of the epistemology of this class of I-thoughts draws on the assumption that the subject who is immediately aware of a mental state and the

one who has the mental state are one and the same. So, we looked at such an account hoping to find an explanation of why that assumption is *a priori* true. On the other hand, according to Peacocke, self-ascriptions of propositional attitudes are not based on the presentation in introspection of the self as an object. However, this requires explaining why one is, after all, not only entitled to ascribe the relevant mental states to *someone*, but also to *oneself*. That is to say, once it is taken on board that at least some mental self-ascriptions are not based on awareness of the self as an object, the problem arises of avoiding Wittgenstein's conclusion (c) that the uses of 'I' as subject (i.e. those mental self-ascriptions which are IEM) are not about a particular person. This conclusion implies, in turn, that the uses of 'I' in the relevant cases are not genuinely referential and, possibly, that mental states lack an owner. Hence, if Peacocke's account of the epistemology of the relevant class of self-ascriptions is successful, it would provide an alternative to the Idealist view about the self.

In order to understand how representation-independent uses of the first-person are possible (recall that uses of the first person are representation-independent if they are not based on a mental state which represents the subject as enjoying that particular mental state), we introduced Peacocke's account of the epistemology of this class of thoughts in terms of what he labels "delta-theory". According to such a model (1) mental states are owned; (2) the subject who has the mental state is the same as the one who is immediately aware of it; (3) it is guaranteed by the constitutive role of [I] in terms of 'the person with this conscious mental state', that the person to which [I] refers is the same as the one who has the mental state.

We considered Peacocke's explanation of (1), which appeals to the fact that conscious mental states have *phenomenal content* and that this, in turn, requires the existence of a subject for whom there is something like being in that state. We supplemented this explanation by noticing that, whether a conscious mental state has phenomenal content or not, it is a mental state which is an object of awareness. But it would take a misconstrual of the relation 'being aware of' to think that there are states which are objects of awareness but which are not objects of awareness for *someone*. Hence, in so far as a mental state is a conscious one, then it has an owner. However, we could not find an explanation of (2). But, an explanation of (2) is both the missing piece of an explanation of why non-inferential mental self-ascriptions are IEM and a necessary element for making the delta-account of the epistemology of the relevant self-ascriptions a workable model.

We will now provide an explanation of why it is *a priori* guaranteed that the mental states one is introspectively aware of and can demonstratively refer to in thought are one's own. By so doing, we will provide a reason for holding Peacocke's account of the constitutive role of [I] (which can be integrated with Evans'), which, in turn, will explain why one need not apply any tests of identity in order to find out that the person the [I] in the judgement refers to and the person who has the mental state and who is introspectively aware of it are one and the same.

4. *Why There Is An A Priori Guarantee That the Mental States One is Immediately Aware of Are One's Own*

Let us consider a conscious mental state like feeling pain. For it to be conscious it means that there is a subject who is immediately (or introspectively) aware of it. That is to say, for it to be a conscious mental state there must be a subject who is aware of it non-observationally and non-inferentially.⁴ This is indeed a consequence of the fact that conscious mental states cannot exist un-owned and of how “immediate awareness” should be understood. But if a subject is immediately aware of pain this just means that she herself is feeling pain. If we then suppose that the subject could be immediately aware of someone else’s mental states, e.g. someone else’s pain, then it would follow that the subject herself would be feeling pain. Similarly, if X is immediately aware of the belief that it is raining, this means that she is believing that it is raining. So, even if, *per impossibile*, she could be immediately aware of Y’s belief that it is raining, this would just mean that X herself would be believing that it is raining.

Notice here the difference with the case of somatic proprioception. For, from the fact that one could be immediately aware of bodily sensations which are located in someone else’s *body*, it does not follow that that body is one’s own. This is so because we have independent non-phenomenological criteria of identity for what has to count as one’s body. But we do not have criteria of identity for what has to count as *one’s own* conscious mental state which are

⁴A positive characterisation of what introspection amounts to falls beyond the scope of this work.

independent of one's being immediately aware of it. Hence, there is an *a priori* guarantee, namely a *conceptual* guarantee, that if someone is immediately aware of a mental state, then this just means that that mental state is her own. Therefore, in the case of direct awareness of a mental state, even if, *per impossibile*, one could be directly aware of someone else's mental state, this would just mean that the very subject would be enjoying that mental state. Thus, the subject would be entitled to self-ascribe that mental state since that mental state would rightly be said to be *her own*.

This indeed vindicates Shoemaker's observations, partially discussed in Ch. 6, that "in being aware that one feels pain one is, tautologically, aware, not simply that the attribute *feel(s) pain* is instantiated, but that it is instantiated in *oneself*".⁵ For the "tautology" has to do with the fact that we do not have an independent grasp of what it means for a conscious mental state to be one's own, but by reference to the fact that there is a subject who is immediately/introspectively aware of it. Similarly, our explanation of why it is *a priori* guaranteed that the subject who is introspectively aware of a mental property is the one who has that property explains why Shoemaker was right in saying that there are certain predicates (e.g. mental ones) "each of which can be known to be instantiated in such a way that knowing it to be instantiated *in that way* (italics mine) is equivalent to knowing it to be instantiated in oneself".⁶ For if one knows that someone is in pain by being directly aware of that mental

⁵Shoemaker 1994a: 89.

⁶Shoemaker 1994a: 90.

state, this means that one oneself is in pain and this, in turn, explains why one would know that that pain is instantiated in oneself.

Therefore, we can conclude that it is *a priori* true that for each mental state one is immediately aware of and is in a position to self-ascribe in judgement on the basis of introspection, that mental state is one's own. To repeat, this is due to the fact that immediate awareness of a mental state logically guarantees that the subject herself is enjoying that mental state. Hence, saying that a subject is immediately aware of a certain mental state guarantees that the relevant mental state is her own.

This also explains why Peacocke is right to characterise the first-person concept in the way proposed in Ch. 7. That is to say, the first-person concept is correctly characterised as that concept which is used in judgements on the basis of the evidence that the person with a certain mental state is *F*. For, there must be such an owner and it cannot be but oneself. Hence, [I] refers to the same subject who has the relevant mental state and who is necessarily the same subject as the one who could self-ascribe that mental state on the basis of introspection.

To conclude, whether or not a non-inferential self-ascription of a mental state is based on the presentation of the self as an object, it would be logically IEM because there is an *a priori* guarantee that any mental state one could be immediately aware of would be *ipso facto* one's own and there is no possibility of mistaking someone else's mental states for one's own. Consequently, subjects are not only entitled to ascribe these mental properties to a subject, but

they are also—and without carrying out an investigation—entitled to *self*-ascribe them.

It has to be stressed that nothing said so far is favourable to the Cartesian position about the self. For, even if conscious mental states are necessarily owned (contrary to the Idealist position), and they are owned by the very subject who is immediately aware of them and can form an I-judgement on the basis of the evidence that the owner of that mental state is *F*, it does not follow that the subject is a mental substance which is immediately present to the mind every time a mental property is manifested.

In fact, this conclusion would be based on several mistakes. First of all, it would conflate the claim that mental states are necessarily owned with the claim that their *owner* is manifested any time such properties, which belong to her, are. Yet, we have agreed that when one is aware of a mental state, at least in some cases, one is not aware of a subject who has it.

However, the Cartesian could avoid the conclusion that a subject is manifested any time one is aware of a mental property by holding the more plausible view that by being aware of one's properties one also acquires knowledge about the object which has that property. Hence, by being aware of our mental properties we acquire knowledge of ourselves as entities that have those properties. But then, given that we are also aware of bodily sensations which give us a conception of ourselves as embodied, the Cartesian would not be justified in concluding that awareness of mental properties just gives us a conception of ourselves as mere owners of sensations and thoughts.

Moreover, even if it were right to say that awareness of mental states just gives us a conception of ourselves as owners of sensations and authors of thoughts, the Cartesian would not be warranted at all in concluding from the kind of property the subject is immediately aware of to the very nature of the subject these properties belong to. For, as Shoemaker noticed,⁷ that would involve an unwarranted passage from considerations pertaining to the epistemological level to considerations pertaining to the metaphysical level. Hence, the Cartesian would be wrong in concluding that the nature of the self could be *entirely* determined by the kind of properties one is aware of in introspection. Therefore, the Cartesian has no reason to maintain that the self is just a mental substance.

5. *Are There Counter-Examples to the View that Non-Inferential Mental Self-Ascriptions are Logically IEM?*

John Campbell has recently maintained that “the phenomenon of thought insertion as described by schizophrenic patients (...) seems to involve an error of identification”.⁸ For, he writes:

A patient who supposes that thoughts have been inserted into his mind by someone else is right about which thoughts they are, but wrong about whose thoughts they are. So thought insertion seems to be a counterexample to the thesis that present-tense introspectively

⁷Cf. Ch. 6.

⁸Campbell 1999: 609.

based reports of psychological state cannot involve errors of identification.⁹

First of all, let us clarify the phenomenon of thought insertion. The following report made by a schizophrenic patient is used by Campbell as an example of the phenomenon at hand:

Thoughts come into my head like 'Kill God'. It's just like my mind working, but it isn't. They come from this chap, Chris. They're his thoughts.¹⁰

Notice, however, that for this report to be a counterexample to the thesis that present-tense introspectively based *self*-ascriptions of mental states are logically IEM, it should be expressed in the *first person* way. But, interestingly, the subject is not saying 'I am thinking 'Kill God''. Rather, she is saying that Chris is thinking the thought, which is occurring in her mind. So, if anything, this report would show that it is *conceivable* to make *third-person* psychological ascriptions on the basis of direct knowledge of someone else's mental states. But this would not show yet that the relevant *self*-ascriptions are liable to EM. The case would be similar to alleged cases of telepathy, whereby subjects maintain that they are introspectively aware of someone else's mental states. Let us consider this possibility in more detail.

If, on the one hand, we take it seriously, then it would mean that Y's mental states are occurring in X's consciousness. For instance Y's pain or belief that it is raining would be occurring in X's consciousness. But, then, X herself would

⁹ Campbell 1999: 609-610.

¹⁰ Frith 1992: 66 quoted in Campbell 1999: 609.

be in pain or thinking that it is raining. Thus, that pain or that belief would be X's own. Hence, under this reading, telepathy would *not* represent a counterexample to the view that mental states one is introspectively aware of are (logically) one's own. However, telepathy so understood would not license third-person psychological claims either. For the subject does not have any reason to suppose that the thought she is currently introspectively aware of is someone else's. Hence, if X were to self-attribute the relevant mental state, then she would be right and her self-ascription would be logically IEM.

If, on the other hand, telepathy is to be understood in a loose sense as a kind of awareness of someone else's mental states, which however, does not make it the case that those states are occurring in one's consciousness, then it would be more a case of *empathy*. That is to say, X can be aware, though not really introspectively, of the fact that Y is suffering or believing that it is raining. But this does not mean that X is in pain herself (unless Y's pain causes X to be in pain herself. In this latter case, however, if X were to self-attribute pain she would be right). Nor does this mean that X herself believes that it is raining. Moreover, X's report would be either 'I'm feeling that Y is in pain' or, equivalently, 'I'm aware that Y is in pain' or, in the case of the belief, 'I am aware that Y believes that it is raining'. So X would not be erroneously self-attributing Y's pain or belief, for she would just be saying that she is aware of someone else's pain or belief. Hence, neither the first understanding of telepathy nor the second are counterexamples to the view that mental states one is introspectively aware of are (logically) one's own and if one were to self-attribute them, then the self-ascription would be logically IEM.

Let us go back to the case of thought-insertion and let us suppose that a genuine first-person claim be made. So X says: 'I am thinking 'Kill God''. Now, this report is perfectly in order. For the schizophrenic subject *is* actually thinking that thought and she is self-ascribing it. If she goes on to wonder whether it really is she who is thinking that thought, then this does not show that she is or can be introspectively aware of someone else's thoughts and thus rationally wonder whether a certain thought is her own. For, after all, the phenomenon of thought-insertion is taken to be a symptom of schizophrenia. So, we do not want to end up saying that she is actually introspectively aware of someone else's mental states nor that, after all, she could be and would thus be justified in wondering whether that thought is her own. Nor, by the same token, does this show that her first-person report is affected by EM or that it could be.

Rather, what the subject's wondering about the possessor of the thought just shows is *her loss of a "sense of ownership" over her thoughts*.¹¹ So the idea is that the schizophrenic patient, for whatever reason, disavows her own thoughts.¹² In fact, Campbell himself writes:

¹¹ Cf. Campbell 1999: 617.

¹² Campbell gives an explanation of the loss of the sense of ownership at a sub-personal level (cf. Campbell 1999: 611-622). It is not our concern to assess it. Rather, our aim is to see whether the phenomenon of thought-insertion is a counterexample to the view that non-inferential mental self-ascriptions are logically IEM. Notice, however, that one could give an alternative explanation of the loss of sense of ownership over one's thoughts at the personal level along the following lines: schizophrenic patients do have certain thoughts which, however, they cannot recognise as their own because, for

The thought inserted into the subject's mind is indeed in some sense his, just because it has been successfully inserted into his mind; it has some special relation to him. He has, for example, *some especially direct knowledge of it* (italics mine). On the other hand, there is, the patient insists, a sense in which the thought is not his, a sense in which the thought is someone else's (...).¹³

So, by Campbell's own lights, the fact that the subject has "some specially direct knowledge" of the thought is enough to make it her own. Still, Campbell seems to think that the best way of explaining the subject's attribution of the thought to someone else would be in terms of an error of identification. But, further on, he himself writes:

(...) [W]hen things go wrong, as with the schizophrenic (...). The sense of ownership of the occurrent thought will be disturbed too.¹⁴

And finally:

[T]he form of words I am recommending—that the schizophrenic has introspective knowledge of a thought of which *he does not recognise himself to be the agent* (italics mine)—does best elucidate the content of the illusion of thought insertion (...).¹⁵

instance, they would feel too ashamed of them. Hence, they *pretend* that someone else has put them in their head.

¹³ Campbell 1999: 610. Notice that when Campbell writes that the thought has been successfully inserted into the subject's mind, he must not be taken literally, otherwise the schizophrenic patient would, after all, be right in saying that the thought is not her own. But, then, she would not be schizophrenic but sane.

¹⁴ Campbell 1999: 617.

¹⁵ Campbell 1999: 619.

However, a loss of a sense of ownership, or of agency (if we are to follow Campbell's terminological suggestion) over one's thoughts does not make the thoughts one is introspectively aware of someone else's, nor is it sufficient for making them not to be one's own. Nor, and this is the real point, does a loss of a sense of ownership over one's thoughts show that it is *logically* possible to be introspectively aware of someone else's thoughts or that it is *logically* possible that the mental states one is introspectively aware of are not one's own. For the alleged counterexample just shows that there is more to the *phenomenology* of conscious thought than the mere having a thought in one's stream of consciousness. Still a phenomenological difference does not produce a logical one. So, even if there were a community whose members were all prone to disavow the thoughts which occur in their consciousness because they do not feel them as their own (or because they do not recognise themselves as their agents), this would not mean that these thoughts would not be theirs, i.e. thought by them and introspectively accessible only to them. Therefore, we can conclude that the phenomenon of thought-insertion is not a counterexample to the view that if someone makes a psychological self-ascription on the basis of introspection, then her judgement is logically IEM.

6. *Conclusions*

We have provided a sound *revisionist* account of IEM and we have offered arguments in its favour. However, we have also showed how both some physical and psychological self-ascriptions can comply with it either *de facto* or

logically. Moreover, we have endorsed and supplemented in important respects the substantive response to the trilemma which was first presented by Evans.

Hence, our reflections on IEM show that it is not true that only a peculiar metaphysics of the self—like the Cartesian or the Idealist—could account for this phenomenon. Rather, IEM is a phenomenon that is entirely compatible with the fact that the 'I' be always a genuine referring expression which refers to a *person*, i.e. to an entity with both physical as well as psychological properties. Moreover, if IEM I-thoughts are the ones which must be in place for a subject to be credited with the possession of the first-person concept, then our considerations, drawing on Evans' and Peacocke's, show that selves, i.e. persons, have a concept of themselves as such.

Thus, the considerations developed in this work have shown not only that the Cartesian and the Idealist positions are problematic for both *metaphysical* and *semantic* reasons. For they would lead us to maintain either the existence of mental substances or the absence of a bearer of conscious mental states, together with the denial of the genuinely referential role of 'I'. Nor have they just shown that the Cartesian and the Idealist views fail to be required in order to account for the *epistemology* of some I-thoughts. For, even if some mental self-ascriptions are not based on a presentation in introspection of the self as an object, there is an *a priori* guarantee that conscious mental states are owned and they are owned by the very subject who is immediately aware of them and who is in a position to self-ascribe them in thought; where none of this—as we saw—is compatible with the Cartesian claim that the self be a mental substance. Rather and foremost, our reflections have shown that the Cartesian

and the Idealist positions are inconsistent with our use of the first-person *concept*. For the various kinds of evidence on which our first-person concept is deployed are such as to give us knowledge of ourselves as embodied and located in space and as persons that exist through time and can be perceived and made sense of by others.

Hence, not only is there *no need to be either Cartesian or Idealist* in order to explain IEM, but in fact *we cannot be either Cartesian or Idealist*. For Cartesianism and Idealism are problematic in their own right, both at the metaphysical and at the semantic level. Moreover, they fail to be required for epistemological reasons and, finally, they are incompatible with our first-person concept. The considerations developed in this work have thus allowed us to provide an *exhaustive substantive response to the trilemma about the self*.

Bibliography

- Anscombe, G. E. M. 1990 "The First Person", in Yourgrau (ed.) 1990, pp. 135-153.
- Bermúdez, J. L. 1994 "Peacocke's Argument Against the Autonomy of Non-Conceptual Content", *Mind and Language*, 9, pp. 402-418.
- 1995a "Nonconceptual Content: From Perceptual Experience to Subpersonal Computational States", *Mind and Language*, 10, pp. 333-369.
- 1995b "Aspects of the Self", *Inquiry*, 38, pp. 489-501.
- 1998 *The Paradox of Self-Consciousness*, Cambridge (Mass.) - London, MIT.
- Bermúdez, J. L., Marcel, A. J., Eilan, N. (eds.) 1995 *The Body and the Self*, Cambridge (Mass.), MIT.
- Burge, T. 1994 "Individualism and Self-Knowledge", in Cassam (ed.) 1994, pp. 65-79.
- Campbell, J. 1995 *Past, Space and Self*, Cambridge (Mass.) - London, MIT.
- 1997 "Sense, Reference and Selective Attention", *The Aristotelian Society Supplement*, LXXI, pp. 55-74.

— 1999 "Schizophrenia, the Space of Reasons, and Thinking as a Motor Process", *The Monist*, 82: 4, pp. 609-625.

Cassam, Q. (ed.) 1994 *Self-Knowledge*, Oxford, OUP.

— 1997a "Subjects and Objects", *Philosophy and Phenomenological Research*, 57, pp. 643-648.

— 1997b *Self and World*, Oxford, Clarendon Press.

Castañeda, H. N. 1966 "'He': A Study in the Logic of Self-Consciousness", *Ratio*, 8, pp. 130-157.

— 1967 "Indicators and Quasi-Indicators", *American Philosophical Quarterly*, 4, pp. 85-100.

— 1994 "On the Phenomeno-Logic of the I", in Cassam (ed.) 1994, pp. 160-166.

Christofidou, A. 1995 "First-Person: the Demand for Identification-Free Self-Reference", *Journal of Philosophy*, 92: 4, pp. 223-234.

Crane, T. (ed.) 1992 *The Contents of Experience*, Cambridge, CUP.

Dancy, J. (ed.) 1997 *Reading Parfit*, Oxford, Blackwell.

Davidson, D. 1994 "Knowing One's Own Mind", in Cassam (ed.) 1994, pp. 43-64.

- Donnellan, K. 1966 "Reference and Definite Descriptions", *Philosophical Review*, 75, pp. 281-304.
- 1970 "Proper Names and Identifying Descriptions", *Synthese*, 21, pp. 335-358.
- Dummett, M. 1973 *Frege. Philosophy of Language*, London, Duckworth.
- Egidi, R. (ed.) 1995 *Wittgenstein: Mind and Language*, Dordrecht, Kluwer.
- Evans, G. 1982 *The Varieties of Reference*, Oxford, Clarendon Press.
- 1985 "Understanding Demonstratives", in Evans, G. 1985 *Collected Papers*, Oxford, Clarendon Press.
- Frege, G. 1892 "Über Sinn und Bedeutung", transl. in Harnish (ed.) 1994, pp. 142-160.
- 1914 "Letter to Jourdin", transl. in Frege 1980, pp. 78-80.
- 1918 "Der Gedanke", transl. in Harnish (ed.) 1994, pp. 517-535.
- 1980 *Philosophical and Mathematical Correspondence*, Oxford, Basil Blackwell.
- French, P. et al. (eds.) 1979 *Contemporary Perspectives in the Philosophy of Language*, Minneapolis, Minnesota University Press.

- Frith, C. D. 1992 *The Cognitive Neuropsychology of Schizophrenia*, Hillsdale NJ, Erlbaum.
- Garrett, B. 1995 "Wittgenstein and the First Person", *Australasian Journal of Philosophy*, 73, pp. 347-355.
- 1997 "Anscombe on 'I'", *The Philosophical Quarterly*, 47, pp. 507-511.
- Gibson, J. J. 1979 *The Ecological Approach to Visual Perception*, Boston, Houghton Muffin.
- Glock, H. 1996 *A Wittgenstein Dictionary*, Oxford, Blackwell.
- Hacker, P.M.S. 1997 *Insight and Illusion*, Bristol, Thoemmes Press.
- Hale, B., Wright, C. (eds.) 1997 *A Companion to the Philosophy of Language*, Oxford, Blackwell.
- Hamilton, A. 1991 "Anscombian and Cartesian Scepticism", *The Philosophical Quarterly*, 41, pp. 39-54.
- Harnish, R. M. (ed.) 1994 *Basic Topics in the Philosophy of Language*, New York - London, Harvester - Wheatsheaf.
- Heal, J. 1994 "Moore's Paradox: A Wittgensteinian Approach", *Mind*, 103, pp. 5-24.

- Hume, D. 1978 *Treatise of Human Nature*, edited by L. A. Selby-Brigge and P. H. Nidditch, Oxford, Clarendon Press.
- Husserl, E. 1970 *Logical Investigations*, London, Routledge.
- Kripke, S. 1979 "Speaker's Reference and Semantic Reference", in French, *et al.* (eds.) 1979.
- 1980 *Naming and Necessity*, Cambridge (Mass.), Harvard University Press.
- Malcolm, N. 1995 "Disentangling Moore's Paradox", in Egidi (ed.) 1995, pp. 195-205.
- Martinich, A. P. (ed.) 1996 *The Philosophy of Language*, Oxford, OUP.
- McDowell, J. 1977 "On the Sense and Reference of a Proper Name", *Mind*, 86, pp. 159-185.
- 1984 "De Re Senses", *Philosophical Quarterly*, 34, pp. 283-294.
- 1986 "Singular Thought and the Extent of Inner Space", in Pettit, McDowell (eds.) 1986, pp. 137-168.
- 1990 "Peacocke and Evans on Demonstrative Content", *Mind*, XCIX, pp. 255-266.
- 1994 *Mind and World*, Harvard, Harvard University Press.

— 1997 "Reductionism and the First Person", in Dancy (ed.) 1997, pp. 230-250.

McGinn, C. 1983 *The Subjective View*, Oxford, Clarendon Press.

Mellor, 1989 "I and Now", *Proceedings of the Aristotelian Society*, 89, pp. 79-94.

Millikan, R. 1993 "Knowing What I Am Thinking Of", *The Aristotelian Society Supplement*, LXVII, pp. 91-108.

— 1994 "On Unclear And Indistinct Ideas", *Philosophical Perspectives*, 8, pp. 75-100.

Noonan, H. W. 1991 "Object-Dependent Thoughts and Psychological Redundancy", *Analysis*, 51, pp. 1-9.

Nozick, R. 1981 *Philosophical Explanations*, Oxford, Clarendon Press.

O'Brien, L. 1995 "Evans on Self-Identification", *Noûs* 29: 2, pp. 232-247.

Parfit, D. 1984 *Reasons and Persons*, Oxford, Clarendon Press.

Peacocke, C. 1983 *Sense and Content*, Oxford, Clarendon Press.

— 1992 *A Study of Concepts*, Cambridge (Mass.) - London, MIT.

— 1999 *Being Known*, Oxford, Clarendon Press.

- Pears, D. 1987 *The False Prison*, Vol. I, II, Oxford, Clarendon Press.
- Perry, J. 1990 "Frege on Demonstratives", in Yourgrau (ed.) 1990, pp. 50-70.
- 1994 "The Problem of the Essential Indexical", in Cassam (ed.) 1994, pp. 167-183.
- 1997 "Indexicals and Demonstratives", in Hale, Wright (eds.) 1997, pp. 586-612.
- Pettit, P., McDowell, J. (eds.) 1986 *Subject, Thought, and Context*, Oxford, Clarendon Press.
- Picardi, E. 1995 "Malcolm on Moore's Paradox", in Egidi (ed.) 1995, pp. 207-209.
- Pryor, J. 1998 "Immunity to Error through Misidentification", *Philosophical Topics*, 26: 1, pp. 271-304.
- Rovane, C. 1987 "The Epistemology of First-Person Reference", *Journal of Philosophy*, 84, pp. 147-167.
- Russell, B. 1905 "On Denoting", in Harnish (ed.) 1994, pp. 161-173.
- 1912 *The Problems of Philosophy*, London, OUP.
- Ryle, G. 1966 *The Concept of Mind*, Harmondsworth, Penguin.

- Sacchi, E. 1999 "Why Is Russell's Principle So Elusive?", *Lingua e Stile*, XXXIV/1, pp. 27-50.
- Sainsbury, M. 1985 "Critical Notice: *The Varieties of Reference*", *Mind*, 94, pp. 120-142.
- Shoemaker, S. 1968 "Self-Reference and Self-Awareness", *Journal of Philosophy*, 65, 19, pp. 555-567, reprinted in Cassam (ed.) 1994, pp. 80-93. Referred to as Shoemaker 1994a.
- 1986 "Introspection and the Self" in French, P. *et al.* 1986 (eds.) *Studies in the Philosophy of Mind, Midwest Studies in Philosophy*, 10, pp. 101-120, reprinted in Cassam (ed.) 1994, pp. 118-139. Referred to as Shoemaker 1994b.
- 1996 *The First Person Perspective and Other Essays*, CUP.
- Stanley, J. 1998 "Persons and Their Properties", *Philosophical Quarterly*, 48, pp. 159-175.
- Strawson, G. 1994 *Mental Reality*, Cambridge (Mass.), MIT Press.
- Strawson, P. 1959 *Individuals*, London and New York, Routledge.
- 1994 "The First Person-And Others", in Cassam (ed.) 1994, pp. 210-215.
- Taschek, W. W. 1985 "Referring to Oneself", *Canadian Journal of Philosophy*, 15, pp. 629-652.

Wittgenstein, L. 1921/61, *Tractatus Logico-Philosophicus*, translated by D. F. Pears and B. F. McGuinness, London, Routledge and Kegan Paul.

— 1953 *Philosophical Investigations*, Oxford, Basil Blackwell. Referred to as PI.

— 1958 *The Blue and the Brown Books*, Oxford, Basil Blackwell. Referred to as Blb.

Wright, C. J. G. 1989 "Wittgenstein's Later Philosophy of Mind: Sensation, Privacy and Intention", *The Journal of Philosophy*, 86, pp. 622-634.

— 1998 "Self-Knowledge: The Wittgensteinian Legacy", in Wright, Smith, Macdonald (eds.) 1998, pp. 13-45.

Wright, C., Smith, B., Macdonald, C. (eds.) 1998 *Knowing Our Own Minds*, Oxford, Clarendon Press.

Yourgrau, P. (ed.) 1990 *Demonstratives*, Oxford, OUP.